

Bevezetés a mátrixelméletbe

EGERVÁRY JENŐ

emlékének ajánlom e könyvet

halálának ötvenedik évfordulója alkalmából

Elméleti matematika

A sorozat kötetei:

Bagyinszky János – György Anna: Diszkrét matematika főiskolásoknak
Bolla Marianna – Krámlí András: Statisztikai következtetések elmélete
G. Horváth Ákos – Szirmai Jenő: Nemeuklideszi geometriák modelljei
Járai Antal: Modern alkalmazott analízis
Kiss Emil: Bevezetés az algebrába
Komornik Vilmos: Valós analízis előadások I–II.
Kostrikin, A. I. – Safarevics, I. R.: Algebra
Lovász László: Kombinatorikai problémák és feladatok
Lovász László – Pelikán József – Vesztegombi Katalin: Diszkrét matematika
Prasolov, Viktor Vasziljevics: Lineáris algebra
Simonovits András: Válogatott fejezetek a matematika történetéből
Stoyan Gisbert: Numerikus matematika mérnököknek és programozóknak
Tóth János – Simon L. Péter: Differenciálegyenletek
Weeks, Jeffrey R.: A tér alakja

RÓZSA PÁL

BEVEZETÉS A MÁTRIXELMÉLETBE



TYPOTEX

Budapest, 2009

© Rózsa Pál, Typotex, 2009

ISBN 978 963 279 028 2

ISSN 1788-1811

Témakör: *lineáris algebra*

Kedves Olvasó!

Önre gondoltunk, amikor a könyv előkészítésén munkálkodtunk. Kapcsolatunkat szorosabbra fűzhetjük, ha belép a *TypoKlubba*, ahonnan értesülhet új kiadványainkról, akcióinkról, programjainkról, és amelyet a *www.typotex.hu* címen érhet el. Honlapunkon megismerkedhet kínálatunkkal is, egyes könyveinknél pedig új fejezeteket, bibliográfiát, hivatkozásokat találhat, illetve az esetlegesen előforduló hibák jegyzékét is letöltheti.

Észrevételeiket a *velemeney@typotex.hu* e-mail címen várjuk.

Kiadja a Typotex kiadó, az 1795-ben alapított

Magyar Könyvkiadók és Könyvterjesztők Egyesülésének tagja.

Felelős kiadó: Votisky Zsuzsa

A könyvet gondozta: Oláh Judit

Borítóterv: Tóth Norbert

Terjedelem: 33,3 (A/5 ív)

Nyomta és kötötte: László András és Társa Nyomdaipari Bt.

Felelős vezető: László András

TARTALOMJEGYZÉK

Előszó	9
1. Mátrixalgebra	13
1.1 Elnevezések és jelölések	13
1.2 Műveletek mátrixokkal	17
1.2.1 Mátrixok összeadása és számmal való szorzása	17
1.2.2 Mátrixok szorzása	19
1.2.3 Speciális mátrixszorzatok	26
1.2.4 Az inverz mátrix	33
1.3 A mátrix rangja	38
1.3.1 A rang fogalma; mátrixok minimális diadikus felbontása	38
1.3.2 Vektorok lineáris függetlensége	46
1.3.3 Rangra vonatkozó tételek	50
1.3.4 Elemi transzformációk, ekvivalens transzformációk, mátrix normálalakja	52
1.3.5 A nullitás fogalma; a Sylvester-féle nullitási tétel	60
1.4 Speciális tulajdonságú mátrixok	62
1.4.1 Speciális mátrixok	62
1.4.2 Szimmetrikus egyenletes kontinuáns mátrix invertálása	70
1.4.3 Nilpotens és ciklikus mátrix polinomjának invertálása	75
1.5 Hiper mátrixok	77
1.5.1 Hiper mátrixok szorzása és faktorizációja	77
1.5.2 Szimmetrikusan particionált másodrendű hiper mátrix faktorizálása, determinánsa, inverze	80
1.5.3 Módosított mátrix és minormátrix inverze	85
1.6 Projektorok	100
1.6.1 Projektorokra vonatkozó tételek	101
1.6.2 Mátrixok általánosított inverze	108
1.7 Lineáris egyenletrendszerek	113
1.7.1 Homogén lineáris egyenletrendszer	114
1.7.2 Inhomogén lineáris egyenletrendszer	117
1.7.3 Lineáris egyenletrendszer kvadratikus együtthatómátrixszal	127
2. A lineáris algebra alapjai	145
2.1 A lineáris tér	146
2.2 Az euklideszi tér	152
2.3 Lineáris függvények, bilineáris és kvadratikus alakok	166
2.4 Lineáris transzformációk	174

2.5 A bázisvektorok transzformációja	186
2.5.1 A koordináták transzformációja új bázisra való áttérés esetén	186
2.5.2 Bilineáris alak mátrixának transzformációja új bázisra való áttérés esetén (kongruens transzformáció)	191
2.5.3 Az \mathbf{x} és \mathbf{Ax} vektorok koordinátái közötti összefüggés	193
2.5.4 A lineáris transzformáció mátrixának transzformációja új bázisra való áttérés esetén (hasonlósági transzformáció)	194
2.6 Lineáris transzformáció sajátvektorai és sajátértékei	206
2.7 Adjungált lineáris transzformációk	214
2.8 Diagonalizálható transzformációk, transzformációpárok, általánosított sajátérték-feladat	216
2.8.1 Önadjungált transzformációk	216
2.8.2 Unitér transzformációk	218
2.8.3 Felcserélhető és normális transzformációk	220
2.8.4 Pozitív definit transzformációk	226
2.8.5 Főtengelytétel és általánosítása	230
2.8.6 Sajátértékek extrémális tulajdonsága	235
2.9 Lineáris transzformációk a valós lineáris térben	239
2.9.1 Lineáris transzformáció normálalakja	239
2.9.2 Szimmetrikus és ortogonális transzformációk	242
2.9.3 Kvadratikus alakok	251
3. Mátrixfüggvények	255
3.1 Egyszerű struktúrájú mátrixok spektrális tulajdonságai	256
3.1.1 Mátrix spektrálfelbontása	256
3.1.2 Projektormátrix spektrálfelbontása	260
3.1.3 Unitér transzformációval diagonalizálható mátrixok	262
3.1.4 Mátrixok szinguláris értékek szerinti felbontása	267
3.2 A mátrixfüggvény fogalma és előállítása	
a minimálpolinom egyszeres gyökei esetén	272
3.2.1 A Cayley–Hamilton-tétel és élesítése	272
3.2.2 A mátrixfüggvény értelmezése és redukciója mátrixpolinomra	275
3.2.3 A Lagrange-féle mátrixpolinomok tulajdonságai	278
3.2.4 Mátrixfüggvény spektrálfelbontása	280
3.2.5 Lagrange-féle mátrixpolinomok előállítása a karakterisztikus mátrix adjungáltjával	287
3.3 Kommutatív blokkokból álló hipermátrixok	298
3.3.1 A hipermátrix determinánsa	298
3.3.2 Mátrixok direkt szorzata	301
3.3.3 Hipermátrix spektrálfelbontása	304
3.3.4 Kronecker-polinomok	307
3.4 Mátrixfüggvény előállítása a minimálpolinom többszörös gyökei esetén	311
3.4.1 Mátrixfüggvény előállítása Hermite-féle mátrix-polinomok segítségével	311
3.4.2 Az Hermite-féle mátrixpolinomok tulajdonságai	318
3.4.3 Mátrixok kváziagonalizálása	322

3.4.4 Nilpotens mátrixok transzformációja Jordan-féle normálalakra	326
3.4.5 Mátrixfüggvények kanonikus előállítása	341
3.5 Elemi osztók elmélete	347
3.5.1 A determinánsosztó invarianciája	347
3.5.2 A determinánsosztó invarianciája speciális esetben	349
3.5.3 A determinánsosztó invarianciája általános esetben	351
3.5.4 Az elemi osztók és a Jordan-féle normálalak	353
3.6 Lineáris differenciálegyenletrendszerek	358
3.6.1 Explicit alakban megadott lineáris elsőrendű közönséges differenciálegyenlet-rendszerek	359
3.6.2 A differenciálegyenlet-rendszer megoldása a rezolvensmátrix ismeretében	362
3.6.3 A rezolvensmátrix meghatározása	364
3.6.4 A rezolvensmátrix előállítása a felcserélhetőségi reláció teljesülése esetén	366
3.6.5 Állandó együtthatómátrixú differenciálegyenlet-rendszerek megoldása	369
3.6.6 Elsőrendű közönséges differenciálegyenlet-rendszer periodikus megoldása	378
3.6.7 Rezgő rendszerek stabilitásvizsgálata	380
3.6.8 Explicit alakban megadott másodrendű rendszerek	385
4. Nemnegatív elemű mátrixok	401
4.1 Irreducibilis mátrixok	402
4.1.1 Út a Frobenius-tételekhez	402
4.1.2 A Frobenius-tételek	407
4.1.3 A Frobenius-tételek következményei	416
4.2 Reducibilis mátrixok	421
4.2.1 A reducibilis mátrixok alaptétele	421
4.2.2 Reducibilis nemnegatív elemű mátrix normálalakja	424
4.3 Primitív és imprimitív mátrixok	428
4.4 Sztochasztikus mátrixok	430
4.4.1 Alapfogalmak és alapvető tételek	431
4.4.2 Markov-láncok ergodicitása; a sztochasztikus mátrixok osztályozása	437
4.4.3 Bolyongási feladatok	443
4.4.4 Spektrálfelbontás meghatározása generátorfüggvény segítségével	456
Irodalomjegyzék	465
Névmutató	471
Tárgymutató	473

ELŐSZÓ

*A természet nagy könyvében csak az tud olvasni, aki ismeri azt a nyelvet,
amelyen e könyv írva van, és ez a nyelv: a matematika.
(G. Galilei: Párbeszédek a két legnagyobb világrendszerről,
a ptolemaiosziról és a kopernikusziról)*

E könyv előzményének tekinthető „Lineáris algebra és alkalmazásai” című könyvem, amely 1973-ban és 1975-ben a Műszaki Könyvkiadó, majd átdolgozás után 1991-ben a Tankönyvkiadó gondozásában jelent meg. A könyv címének megváltoztatásával egyrészt arra utalok, hogy ez a kiadás a lineáris algebrának a korábbinál szűkebb területére korlátozódik, másrészt a címében is kifejezésre szeretném juttatni azt a törekvésemet, hogy olyan könyvet adjak az Olvasó kezébe, amelynek segítségével az elméleti ismeretek megalapozása mellett kellő rutint szerezhet a mátrixokkal való számítások elvégzésében is. Ezért néhány fejezetet, amelyek alkalmazására a tapasztalat szerint kisebb az igény, elhagytam, és ugyancsak kihagytam a lineáris algebra numerikus módszereivel foglalkozó önálló fejezetet. Ennek anyagát részben beépítettem a korábbi fejezetekbe, illetve úgy ítélt meg, hogy az óriásira duzzadt anyag már nem olvasható egybe az alapozó fejezetekkel. Helyette inkább megnőveltem a kidolgozott példák számát, amelyek hozzásegítik az Olvasót az anyag jobb megértéséhez.

A könyv négy fejezetre tagolódik. Az első fejezet a mátrixalgebra elemeit tartalmazza és egyúttal itt ismerkedhet meg az Olvasó olyan speciális tulajdonságú, ún. strukturált mátrixokkal, amelyek az alkalmazások számos területén előfordulnak. A fejezetet a mátrixalgebra legfontosabb alkalmazási területe, a lineáris egyenletrendszerek elmélete és megoldása zárja. A második fejezet a lineáris algebra alapjaival foglalkozik, amely azt az elméleti háttérrel szolgáltatja, amelybe a mátrixelmélet beágyazható. Ez a fejezet tartalmazza a sajátérték-feladatot és a lineáris transzformációk elméletét. A harmadik fejezet a mátrixfüggvények definíciójával és előállításával foglalkozik. Ennek keretében tárgyalja a mátrixok felosztását diagonalizálható és nemdiagonalizálható mátrixok osztályára és ennek kapcsán a Jordan-féle normálalakra való

transzformálásukat. Itt kapott helyet mátrixok szinguláris értékeinek az értelmezése és szinguláris értékek szerinti felbontásuk. Végül a mátrixfüggvények legfontosabb alkalmazási területe, a lineáris differenciálegyenlet-rendszerek elmélete és megoldása zárja a fejezetet. A negyedik fejezet a mátrixelmélet egy speciális területével, nemnegatív elemű mátrixokkal foglalkozik, ezen belül sztochasztikus mátrixokkal és ezek alkalmazásával bolyongási feladatok megoldására. A fejezet végén egy speciális bolyongási feladattal kapcsolatos Sylvester–Kac-mátrix sajátérték-feladatának generátorfüggvény segítségével nyerhető megoldása található.

A könyv egy egységes szemléletű és felépítésű bevezetést kíván nyújtani a mátrixelméletbe, amely tartalmában és módszereiben figyelembe veszi a tárgyalt anyag műszaki és természettudományokban való alkalmazásának az igényét. Az anyag felépítésében központi helyet foglalnak el a diádok és a projektorok. Mátrixok minimális diadikus előállításával vezeti be a rang fogalmát, és ezen keresztül jut a lineáris egyenletrendszerek elméletéhez. Projektorok minimális diadikus előállítása automatikusan biortogonális vektorrendszert szolgáltat, és ez vezet a mátrixok spektrálfelbontásához. Mátrixok függvényének értelmezésével és előállításával mutat utat a mátrixok osztályozásához és lineáris differenciálegyenlet-rendszerekre való alkalmazásához. Itt felismerhető az Egerváry-iskola által megteremtett módszer, amely ma már hagyományosnak tekinthető a hazai mátrixelméleti kutatásokban. A könyv ebben eltér a legtöbb külföldi szakirodalomban követett – klasszikusnak mondható – iránytól. A bizonyítások során, amikor csak lehet, konstruktív módszerek szerepelnek, ezzel lehetővé válik, hogy egyúttal a megoldási módszereket is megismerje az Olvasó. Példa erre a Jordan-féle normálalak bevezetése, vagy a Kronecker-polinomok spektrálfelbontása. A 103 kidolgozott példa megválasztásánál fontos szempont volt, hogy a legkülönbözőbb alkalmazási területeken előforduló típusok forduljanak elő. Sok esetben a vizsgált rendszer szabályos tulajdonságai strukturált mátrixokra vezetnek. Ilyenek pl. a tridiagonális és a ciklikus mátrixok, perturbált mátrixok, vagy a kommutatív blokkokból álló hipermátrixok. Ezek invertálása, illetve spektrálfelbontásuk előállítása során számos ötletet ismerhet meg az Olvasó és ezzel egyúttal rutint szerezhet bonyolultabb feladatok megoldásához.

A könyv elsősorban egyetemi és főiskolai hallgatóknak, doktoranduszoknak, tudományos kutatóknak és azoknak az érdeklődőknek kíván bepillantást nyújtani a mátrixok világába, akiknek korábbi tanulmányaik során erre nem volt alkalmuk, vagy fel kívánják frissíteni régebben szerzett ismereteiket. Véleményem szerint eredményesen forgathatják a könyvet azok a matematikusok, fizikusok, mérnökök, informatikusok és közgazdászok, akik munkájuk során olyan matematikai kérdésekkel találkozhatnak, amelyek kapcsolatba hozhatók mátrixelméleti problémákkal. Az anyag összeállítása és felépítése

során az volt a cél, hogy szinte a nulláról kiindulva fokozatosan ismerkedjék meg az Olvasó a fogalmakkal és tételekkel, és ezek bizonyítása során elsajátítsa azt a gondolkodásmódot, amely elősegíti az újabb szakirodalom megértését, esetenként pedig új problémák felvetését és azok megoldását. Az anyag megértéséhez az Olvasónak szüksége van egy minimális ismeretre a matematika egyéb területeiről, így az analízisből ismernie kell függvények hatványsorba fejtését, algebrából a determináns fogalmát, az algebra alaptételét és a polinomok elméletének alapjait, a komplex számok algebráját és a lineáris differenciálegyenletek elemeit. A negyedik fejezetben szereplő sztochasztikus mátrixok tárgyalásánál jó, ha tisztában van a valószínűség-számítás alapfogalmaival és ismeri a komplex változós függvényekre vonatkozó reziduum-tételt.

A könyv formai szerkezetének kialakításában a matematikai irodalomban szokásos módszer követhető nyomon, amely az alapvető ismereteket „definíció – tétel – bizonyítás” hármas tagolásban közli; a bizonyítás végét a ■ jel, a példák végét pedig három csillag, * * * jelzi. Az egyes fogalmakat a könnyebb megkülönböztethetőségük céljából egymástól eltérő betűtípusok jelölik, így

a mátrixokat félkövér, álló nagybetűk: **A**, **B**, ...

az oszlop- és sorvektorokat félkövér, álló kisbetűk: **a**, **b**, ...

az absztrakt lineáris tér vektorait félkövér, dőlt kisbetűk: **a**, **b**, ...

a geometriai tér vektorait félkövér, groteszk kisbetűk: **a**, **b**, ...

a lineáris tér transzformációit groteszk nagybetűk **A**, **B**, ...

Remélhetőleg ezek a formai megoldások megnövelik az anyag áttekinthetőségét.

Ezen a helyen szeretnék köszönetet mondani mindazoknak, akik kritikai észrevételeikkel és megjegyzéseikkel segítettek kijavítani a könyv korábbi kiadásában található hibákat. Mindenekelőtt Lee Annának, a könyv korábbi lektorának szeretném megköszönni gondos munkáját és értékes észrevételeit, amelyekkel sokat segített a könyv szerkezetének javításában és számos következetlenség kiküszöbölésében. Köszönet illeti volt tanítványaimat, akik 40 év alatt aktív jelenlétükkel, az előadásra való reagálásukkal segítettek az anyag didaktikai felépítésének formálásában – nem is beszélve arról a segítségéről, amelyet a konkrét hibák összegyűjtésével nyújtottak. Nagy segítségemre volt Stubnya Gusztávné, elsősorban a példák kiválasztásával és megoldásukkal. Külön köszönet illeti Erő Zsuzsát az anyag gondos nyomdai előkészítéséért, dr. Tegze Juditot az ábrák precíz elkészítéséért, Oláh Juditot, aki az előző kiadáshoz hasonlóan elvállalta a könyv szerkesztését és számos hasznos tanácsával hozzájárult az egységes szerkezet kialakításához. Végül köszönetemet fejezem ki a Typotex Könyvkiadó valamennyi munkatársának, elsősorban Votisky Zsuzsának, lelkiismeretes munkájukért és támogatásukért, amellyel lehetővé tették a könyv megjelenését.

1. fejezet

MÁTRIXALGEBRA

*Az egyenlet leír egy összefüggést, de önmagában nem ad megoldást.
Az egyenletet meg kell oldani!
(Amir D. Aczel: Isten egyenlete)*

Ebben a fejezetben bevezetjük azokat az alapfogalmakat és jelöléseket, amelyek az egész könyvben végigkísérik az Olvasót. Értelmezzük a mátrixokkal végzett műveleteket, és ennek során bemutatjuk a legfontosabb speciális szerkezetű mátrixokat, amelyek a későbbi fejezetekben ismételten előfordulnak. Már kezdettől fogva kiemeljük azt a központi szerepet, amelyet az elmélet fokozatos kiépítése során a *diádoknak* szánunk. Viszonylag nagy helyet szentelünk speciális szerkezetű mátrixok invertálásának, és kidolgozott feladatok segítségével kívánjuk segíteni az Olvasót, hogy kellő rutint szerezhesen a mátrixok kezelésében. A *rang* fogalmát összekapcsoljuk a mátrixok faktORIZÁCIÓJÁNAK feladatával és ennek kapcsán bevezetjük a mátrixok *minimális diadikus előállítását*, amely alapvető fontosságú feladat lesz a későbbiek során is. Külön pontban foglalkozunk a négy blokkra particionált *hipermátrixokkal*, valamint a *projektorok* legfőbb tulajdonságaival, végül alkalmazzuk a mátrix-algebra elemeit a lineáris egyenletrendszerek megoldásának elméletében.

1.1 ELNEVEZÉSEK ÉS JELÖLÉSEK

Tekintsük az a_{ij} komplex számoknak egy m sorból és n oszlopból álló sémáját:

$$\begin{bmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \dots & \dots \\ a_{m1} & a_{m2} \dots a_{mn} \end{bmatrix}.$$

Ezt a sémát $m \times n$ típusú, vagy $m \times n$ -es mátrixnak nevezzük és a következőképpen jelöljük:

$$\mathbf{A}_{(m,n)} = [a_{ij}] \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, n),$$

vagy egyszerűen

$$\mathbf{A} = [a_{ij}] \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, n).$$

Az a_{ij} számok a mátrix elemei. Annak feltüntetésére, hogy az \mathbf{A} mátrix ij indexű eleme a_{ij} , gyakran a következő jelölést alkalmazzuk: $(\mathbf{A})_{ij} = a_{ij}$. Ha $m = n$, akkor a mátrixot *n-edrendű kvadratikus mátrixnak* nevezzük. Jelölése:

$$\mathbf{A}_n = [a_{ij}] \quad (i, j = 1, 2, \dots, n),$$

vagy egyszerűen

$$\mathbf{A} = [a_{ij}] \quad (i, j = 1, 2, \dots, n).$$

Mivel ezek szerint *n*-edrendű mátrix csak kvadratikus lehet, ezért a kvadratikus jelzőt ilyenkor elhagyhatjuk. Ugyanígy elhagyhatjuk annak feltüntetését is, hogy milyen típusú mátrixról van szó, ha a tárgyalás során felírt műveletekből vagy összefüggésekből ez egyértelműen kiderül.

A sorok és oszlopok felcserélésével nyert mátrixot az eredeti mátrix *transzponáltjának* nevezzük és T felső indexszel jelöljük:

$$\mathbf{A}^T = [a_{ij}^T] = [a_{ji}].$$

Ebből következik, hogy egy $m \times n$ típusú mátrix transzponáltja $n \times m$ típusú. Ha egy mátrix elemei helyébe ezeknek a konjugáltját tesszük, akkor az eredeti mátrix *konjugáltját* kapjuk, ezt felülvonással jelöljük:

$$\overline{\mathbf{A}} = [\overline{a_{ij}}].$$

A mátrix *transzponált konjugáltját* H felső indexszel jelöljük:

$$\overline{\mathbf{A}}^T = [\overline{a_{ji}}] = \mathbf{A}^H.$$

Két mátrix akkor *egyenlő* egymással, ha azonos típusúak és megfelelő (azonos indexű) helyen álló elemeik egyenlők:

$$\mathbf{A} = \mathbf{B}, \text{ ha } a_{ij} = b_{ij} \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, n).$$

Ha az \mathbf{A} mátrix megegyezik a konjugáltjával:

$$\mathbf{A} = \overline{\mathbf{A}},$$

akkor \mathbf{A} valós elemű, vagy röviden *valós mátrix*; ha

$$\mathbf{A} = [-\overline{a_{ij}}] = -\overline{\mathbf{A}},$$

akkor \mathbf{A} *képzetes (imaginárius) mátrix*. Ha egy (kvadratikus) mátrix megegyezik a transzponáltjával:

$$\mathbf{A} = \mathbf{A}^{\top},$$

akkor *szimmetrikus*, ha pedig

$$\mathbf{A} = -\mathbf{A}^{\top},$$

akkor *ferdén szimmetrikus**.

Ha egy (kvadratikus) mátrix megegyezik transzponáltjának konjugáltjával:

$$\mathbf{A} = \mathbf{A}^{\mathrm{H}},$$

akkor *hermitikus*** , ha pedig

$$\mathbf{A} = -\mathbf{A}^{\mathrm{H}},$$

akkor \mathbf{A} *ferdén hermitikus*.

Ebből következik, hogy speciálisan a valós elemű hermitikus mátrix mindig *szimmetrikus*, a valós elemű ferdén hermitikus mátrix pedig mindig *ferdén szimmetrikus*.

Az \mathbf{A} kvadratikus mátrix ii indexű elemei a mátrix *főátlójában* (*fődiagonálisában*) helyezkednek el. Ha egy kvadratikus mátrixnak a főátlóján kívüli valamennyi eleme zérus, akkor ún. *diagonálmátrixhoz* jutunk. Jelölése:

$$\mathbf{D} = \langle d_1 d_2 \dots d_n \rangle \text{ vagy } \mathbf{D} = \langle d_k \rangle \quad (k = 1, 2, \dots, n),$$

ahol d_k az a_{kk} elemet jelöli. \mathbf{A}

$$\delta_{ij} = \begin{cases} 1, & \text{ha } i = j, \\ 0, & \text{ha } i \neq j. \end{cases}$$

Kronecker***-féle szimbólumot használva, a $[\delta_{ij}]$ mátrix az \mathbf{E} egységmátrix:

$$\mathbf{E} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}.$$

*Az irodalomban gyakran találkozunk a ferdén szimmetrikus mátrixokra vonatkozóan az „antiszimmetrikus mátrixok” elnevezéssel. Tekintettel arra, hogy ez tisztán nyelvi szempontból nem kifogástalan szóképzés, másrészt pedig az orosz, angol, ill. német nyelvű matematikai irodalomban a *koco*, *skew*, illetve *schief* kifejezés terjedt el, ezért könyvünkben a „ferdén” elnevezést fogjuk használni.

**Ch. Hermite (ejtsd: ermit) francia matematikus (1822–1901) nevééről kapta az elnevezést.

***L. Kronecker (1823–1891) német matematikus.

Tehát az egységmátrix mindig kvadratikus. Ha a rendjét is fel akarjuk tüntetni, akkor az \mathbf{E}_n jelölést használjuk.

Az olyan mátrixot, amelynek egyetlen eleme 1, a többi pedig 0, *mátrix-egységnek* nevezzük és ha az ij indexű elem 1, akkor az \mathbf{E}_{ij} jelölést alkalmazzuk. A csupa 0 elemből álló mátrixot *zérusmátrixnak* nevezzük és $\mathbf{0}$ -val jelöljük.

A mátrixok körében kitüntetett szerepük van az egyoszlopos, ill. az egy soros mátrixoknak. Ezeket *oszlop-*, ill. *sorvektoroknak* nevezzük és kisbetűkkel jelöljük:

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}; \quad \mathbf{a}^T = [a_1 \quad a_2 \quad \dots \quad a_n].$$

Mivel a sorvektor a megfelelő oszlopvektor transzponáltja, megegyezés szerint a sorvektort a T felső indexszel látjuk el. Az olyan oszlop-, ill. sorvektort, amelynek egyetlen eleme 1, a többi pedig 0, *egységvektornak*, mégpedig *egység-oszlopvektornak*, ill. *egység-sorvektornak*, ha minden eleme 0, nullvektornak nevezzük; ha az i -edik elem 1, akkor az \mathbf{e}_i , ill. \mathbf{e}_i^T jelölést alkalmazzuk; a nullvektor jele $\mathbf{0}$.

A mátrixot vízszintes és függőleges vonalakkal részekre oszthatjuk. Az ilyen részekre osztást *particionálásnak*, a particionálással adódó részmatrixokat *blokkoknak* nevezzük. A legegyszerűbb és legtermészetesebb részekre osztás, ha a mátrixot csupa függőleges vonallal oszlopvektoraira, vagy csupa vízszintes vonallal sorvektoraira particionáljuk:

$$\mathbf{A} = \left[\begin{array}{c|c|c|c} \vdots & \vdots & \vdots & \vdots \\ \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \\ \vdots & \vdots & \vdots & \vdots \end{array} \right], \quad \text{illetve} \quad \mathbf{A} = \left[\begin{array}{c} \mathbf{a}^1 \\ \hline \mathbf{a}^2 \\ \hline \vdots \\ \hline \mathbf{a}^m \end{array} \right].$$

Itt a sorvektorokat felső indexekkel jelöljük, mivel azok nem a megfelelő indexű oszlopvektorok transzponáltjai!

Általánosabb particionálás a következő:

$$\mathbf{A} = \left[\begin{array}{cc} \overbrace{\mathbf{A}_{11} \quad \mathbf{A}_{12}}^r & \overbrace{\mathbf{A}_{21} \quad \mathbf{A}_{22}}^s \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{array} \right] \left. \vphantom{\begin{array}{cc} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{array}} \right\} \begin{matrix} p \\ q \end{matrix}$$

ahol az \mathbf{A}_{11} blokk $p \times r$ típusú, az \mathbf{A}_{12} blokk $p \times s$ típusú, \mathbf{A}_{21} típusa $q \times r$ és \mathbf{A}_{22} típusa $q \times s$.

Ha kiválasztjuk az \mathbf{A} mátrix i_1, i_2, \dots, i_r -edik sorainak j_1, j_2, \dots, j_s -edik elemeit és ezeket ismét mátrixként írjuk, akkor az eredeti mátrix egy *minor-mátrixát* kapjuk. Jelölése:

$$(1.1.1) \quad \mathbf{A}_{j_1, j_2, \dots, j_s}^{i_1, i_2, \dots, i_r} = \begin{bmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \dots & a_{i_1 j_s} \\ a_{i_2 j_1} & a_{i_2 j_2} & \dots & a_{i_2 j_s} \\ \dots & \dots & \dots & \dots \\ a_{i_r j_1} & a_{i_r j_2} & \dots & a_{i_r j_s} \end{bmatrix}.$$

Az \mathbf{A} kvadratikus mátrix elemeiből alkotott determinánst \mathbf{A} *determinán-sának* nevezzük. Jelölése: $|\mathbf{A}|$, vagy $\det \mathbf{A}$, tehát – a determináns definíciója szerint –

$$|\mathbf{A}| = \sum_{(n!)} (-1)^I a_{1i_1} a_{2i_2} \dots a_{ni_n},$$

ahol I az $1, 2, \dots, n$ számok i_1, i_2, \dots, i_n permutációjában szereplő inverziók* számát jelenti, az összegzést pedig ki kell terjeszteni az $1, 2, \dots, n$ számok valamennyi permutációjára. Adott \mathbf{A} mátrix valamely kvadratikus minor-mátrixának determinánsát röviden *minornak* fogjuk nevezni.

Ha az \mathbf{A} kvadratikus mátrix i -edik sorát és j -edik oszlopát elhagyjuk és az így nyert $(n-1)$ -edrendű minormátrix determinánsát $(-1)^{i+j}$ előjellel ellátjuk (ezt a szabályt szokás a determinánselméletben *sakktáblaszabálynak* nevezni), akkor az A_{ij} *előjeles aldeterminánst* kapjuk. Ezt másképpen az a_{ij} elemhez tartozó *algebrai komplementumnak* (az angol irodalomban *cofactor*-nak) is szokták nevezni. (A determinánsokra vonatkozóan lásd pl. [1].)

1.2 MŰVELETEK MÁTRIXOKKAL

1.2.1 Mátrixok összeadása és számmal való szorzása

1.2.1 definíció. *Azonos típusú mátrixok összegén olyan mátrixot értünk, amelynek elemei az egyes mátrixok megfelelő (azonos indexű, azonos helyen álló) elemeinek összege.*

Legyen $\mathbf{A} = [a_{ij}]$ és $\mathbf{B} = [b_{ij}]$ egy-egy $m \times n$ típusú mátrix. Összegük az a $\mathbf{C} = [c_{ij}]$ mátrix, amelyre

$$(1.2.1) \quad c_{ij} = a_{ij} + b_{ij}.$$

*Emlékeztetünk arra, hogy az $1, 2, \dots, n$ számok egy i_1, i_2, \dots, i_n permutációjában az i_k és i_l elem inverziót alkot, ha $k < l$ és $i_k > i_l$.

E definíció alapján közvetlenül belátható, hogy a mátrixok összeadása kommutatív és asszociatív művelet:

$$\begin{aligned}\mathbf{A} + \mathbf{B} &= \mathbf{B} + \mathbf{A}, \\ \mathbf{A} + (\mathbf{B} + \mathbf{C}) &= (\mathbf{A} + \mathbf{B}) + \mathbf{C}.\end{aligned}$$

1.2.2 definíció. *Mátrix szorzása komplex számmal olyan mátrixot ad, amelynek minden eleme az adott mátrix megfelelő elemének szorzata a komplex számmal.*

Legyen $\mathbf{A} = [a_{ij}]$ tetszőleges mátrix és c egy komplex szám; ekkor a $c\mathbf{A}$ szorzat eredménye olyan $\mathbf{G} = [g_{ij}]$ mátrix, amelynek elemeire

$$(1.2.2) \quad g_{ij} = ca_{ij}.$$

E definíció alapján közvetlenül beláthatók az alábbi disztributív és asszociatív tulajdonságok:

$$\begin{aligned}c(\mathbf{A} + \mathbf{B}) &= c\mathbf{A} + c\mathbf{B}, \\ (a + b)\mathbf{C} &= a\mathbf{C} + b\mathbf{C}, \\ (ab)\mathbf{C} &= a(b\mathbf{C}).\end{aligned}$$

Itt \mathbf{A} , \mathbf{B} , \mathbf{C} mátrixokat, a , b , c pedig komplex számokat jelent.

Hasonlóképpen belátható, hogy a transzponálás és a konjugálás tagonként elvégezhető:

$$\begin{aligned}(\mathbf{A} + \mathbf{B})^T &= \mathbf{A}^T + \mathbf{B}^T, \\ \overline{(\mathbf{A} + \mathbf{B})} &= \overline{\mathbf{A}} + \overline{\mathbf{B}},\end{aligned}$$

továbbá

$$(c\mathbf{A})^T = c\mathbf{A}^T \quad \text{és} \quad \overline{(c\mathbf{A})} = \overline{c}\overline{\mathbf{A}}.$$

A fentiek alapján két azonos típusú mátrix különbségének értelmezése:

$$(\mathbf{A} - \mathbf{B}) = \mathbf{A} + (-1)\mathbf{B}.$$

A determinánsoknak abból a tulajdonságából, amely szerint egy determináns egyetlen sorának (vagy oszlopának) elemeit egy adott számmal megszorozva, az egész determináns szorozódik az adott számmal, következik, hogy n -edrendű \mathbf{A} mátrixra

$$|\alpha\mathbf{A}_n| = \alpha^n |\mathbf{A}_n|.$$

A fentiekből következik, hogy azonos módon particionált mátrixok összeadása és számmal való szorzása blokkonként elvégezhető.

Mátrixok összeadásának és számmal való szorzásának a definíciója lehetővé teszi a következő egyszerű, de igen fontos tétel bizonyítását.

1.2.1 tétel. *Bármely komplex elemű kvadratikus mátrix felírható*

$$\mathbf{A} = \mathbf{H}_1 + i\mathbf{H}_2$$

alakban, ahol \mathbf{H}_1 és \mathbf{H}_2 hermitikus mátrix.

Bizonyítás. Tekintsük az alábbi azonosságot:

$$\mathbf{A} = \frac{1}{2} (\mathbf{A} + \mathbf{A}^H) + \frac{1}{2} (\mathbf{A} - \mathbf{A}^H).$$

Közvetlenül belátható, hogy ezen azonosság jobb oldalának első tagja hermitikus, második tagja pedig ferdén hermitikus mátrix, vagyis

$$\begin{aligned} \left\{ \frac{1}{2} (\mathbf{A} + \mathbf{A}^H) \right\}^H &= \frac{1}{2} (\mathbf{A}^H + \mathbf{A}) = \frac{1}{2} (\mathbf{A} + \mathbf{A}^H), \\ \left\{ \frac{1}{2} (\mathbf{A} - \mathbf{A}^H) \right\}^H &= \frac{1}{2} (\mathbf{A}^H - \mathbf{A}) = -\frac{1}{2} (\mathbf{A} - \mathbf{A}^H). \end{aligned}$$

Hasonlóképpen belátható, hogy ha \mathbf{F} ferdén hermitikus mátrix, akkor $\frac{1}{i}\mathbf{F}$ hermitikus mátrix, ahol i az imaginárius egység; ugyanis $\mathbf{F}^H = -\mathbf{F}$ miatt

$$\left(\frac{1}{i}\mathbf{F} \right)^H = -\frac{1}{i}\mathbf{F}^H = \frac{1}{i}\mathbf{F}$$

valóban fennáll. Bevezetve a

$$\mathbf{H}_1 = \frac{1}{2} (\mathbf{A} + \mathbf{A}^H), \quad \mathbf{H}_2 = \frac{1}{2i} (\mathbf{A} - \mathbf{A}^H)$$

jelölést, a tétel állítása adódik. ■

1.2.2 Mátrixok szorzása

1.2.3 definíció. Az $m \times p$ típusú \mathbf{A} mátrix és a $p \times n$ típusú \mathbf{B} mátrix szorzatán azt az $m \times n$ típusú mátrixot értjük, amelynek ij indexű elemét az \mathbf{A} mátrix i -edik sorának és a \mathbf{B} mátrix j -edik oszlopának kompozíciója* adja.

Legyen $\mathbf{A} = [a_{ij}]$ és $\mathbf{B} = [b_{ij}]$; akkor az $\mathbf{AB} = \mathbf{C} = [c_{ij}]$ szorzatmátrix elemei:

$$(1.2.3) \quad c_{ij} = \sum_{k=1}^p a_{ik} b_{kj}.$$

*Az a_1, a_2, \dots, a_n és b_1, b_2, \dots, b_n szám- n -esek kompozícióján az $a_1 b_1 + a_2 b_2 + \dots + a_n b_n$ szorzatösszeget értjük.

A mátrixok szorzása a legjellegzetesebb mátrixművelet; nyilvánvaló sajátossága, hogy *nem kommutatív*.

Az összeszorozhatóság feltétele: a bal oldali tényező oszlopai számának meg kell egyeznie a jobb oldali tényező sorainak a számával. Márpedig a tényezők sorrendjének felcserélése esetén nincs mindig biztosítva ez a feltétel, tehát az egyik sorrendben esetleg elvégezhető a szorzás, a másik sorrendben pedig nem. Két azonos rendű kvadratikus mátrix vagy egy $m \times p$ és egy $p \times m$ típusú mátrix mindkét sorrendben összeszorozható ugyan, de ez a szorzás sem kommutatív, mivel – pl. n -edrendű mátrixokra – általában

$$\sum_{k=1}^n a_{ik} b_{kj} \neq \sum_{k=1}^n b_{ik} a_{kj}.$$

A szorzás kommutativitásának hiánya a legtöbb probléma forrása, amelyet a mátrixelmélet felvet; ez adja viszont az elmélet szépségét is.

Természetesen vannak olyan speciális mátrixok, amelyekre a szorzás kommutatív. Az ilyen mátrixokat *felcserélhető*nek nevezzük.

Várható, hogy ilyen esetekben sokkal többet állíthatunk a szóban forgó mátrixokról, mint általában, ezért a következőkben mindig gondosan hangsúlyozni fogjuk, ha azt találjuk, hogy a mátrixok bizonyos osztályai körében a szorzás kommutativitása fennáll.

Könnyű meggyőződni arról, hogy az eddig megismert mátrixokra a felcserélhetőség a következő esetekben érvényes:

(1) Bármely kvadratikus mátrix felcserélhető a vele azonos rendű zérusmátrixszal és egységmátrixszal:

$$\begin{aligned} \mathbf{A}\mathbf{E} &= \mathbf{E}\mathbf{A} = \mathbf{A}, \\ \mathbf{A}\mathbf{0} &= \mathbf{0}\mathbf{A} = \mathbf{0}. \end{aligned}$$

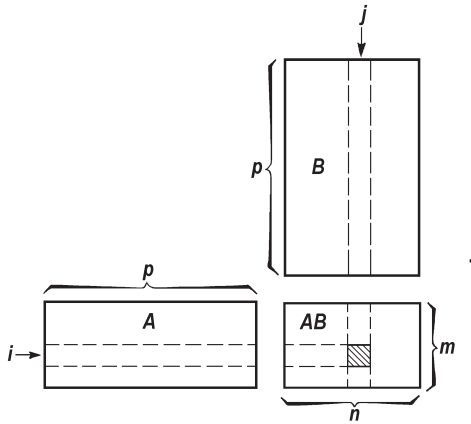
Ezekből az összefüggésekből látható, hogy az egységmátrix és a zérusmátrix ugyanazt a szerepet tölti be a mátrixok körében, mint az 1, ill. 0 a számok körében; ez az elnevezésüket is indokolja.

(2) Diagonálmátrixok mindig felcserélhetőek:

$$\mathbf{D}_1 \mathbf{D}_2 = \mathbf{D}_2 \mathbf{D}_1 = \left\langle d_k^{(1)} d_k^{(2)} \right\rangle.$$

(Diagonálmátrixok szorzata szintén diagonálmátrix!) A szorzás definíciójából következik, hogy ha az \mathbf{A} mátrixot balról szorozzuk a \mathbf{D} diagonálmátrixszal, akkor az \mathbf{A} mátrix *sorait* kell szorozni a diagonálmátrix megfelelő elemeivel, ha pedig jobbról szorozzuk, akkor az *oszlopait*.

Mátrixok szorzásának begyakorlására – különösen a kezdőknek – ajánlható a következő elrendezés:



Eszerint a bal oldali tényező egyes sorait és a jobb oldali tényező egyes oszlopait meghosszabbítva, ezek metszéséhez kerülnek a szorzatmátrix megfelelő elemei. A tévedés így sokkal könnyebben elkerülhető, mintha egymás mellé írnánk a tényezőket.

1. Példa.

$$\begin{bmatrix} 2 & 3 & -1 \\ 4 & -5 & 6 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 5 & 3 & 5 \\ 7 & 6 & -2 \end{bmatrix} = \begin{bmatrix} 12 & 5 & 17 \\ 25 & 25 & -37 \end{bmatrix}$$

* * *

Ha a megfelelő összeszorozhatósági feltételek teljesülnek, akkor a szorzás definíciója alapján többtényezős mátrixszorzatok értelmezhetők; ezekre fennáll az *asszociativitás* törvénye. Ez közvetlenül belátható, ha felírjuk három mátrix szorzatának általános elemét először úgy, hogy az első két tényező szorzatát szorozzuk a harmadikkal, utána pedig úgy, hogy az első tényezőt szorozzuk a másik kettő szorzatával. Legyen

$$\mathbf{A}_{(m,r)} = [a_{ij}]; \quad \mathbf{B}_{(r,s)} = [b_{ij}]; \quad \mathbf{C}_{(s,n)} = [c_{ij}].$$

Ekkor az $\left(\mathbf{A}_{(m,r)} \mathbf{B}_{(r,s)} \right)_{(m,n)} \mathbf{C}_{(s,n)}$ mátrix ij indexű eleme

$$\sum_{q=1}^s \left(\sum_{p=1}^r a_{ip} b_{pq} \right) c_{qj},$$

az $\mathbf{A}_{(m,r)} \begin{pmatrix} \mathbf{B} & \mathbf{C} \\ (r,s) & (s,n) \end{pmatrix}$ mátrix ij indexű eleme pedig

$$\sum_{p=1}^r a_{ip} \left(\sum_{q=1}^s b_{pq} c_{qj} \right).$$

Minthogy skalárokra vonatkozóan véges összegezek esetén az összegezek sorrendje felcserélhető, így az ij indexű elemre kapott két kifejezés megegyezik egymással. A jobb oldalak egyenlőségéből következik, hogy a többtényezős mátrixszorzás valóban független a tényezők csoportosításától:

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}) = \mathbf{ABC}.$$

A háromtényezős mátrixszorzat általános elemét tehát kettős szumma alakjában nyerjük. Tegyük fel az egyszerűség kedvéért, hogy a szóban forgó mátrixok mind kvadratikusak. Ekkor

$$(1.2.4) \quad \mathbf{ABC} = \left[\sum_{q=1}^n \sum_{p=1}^n a_{ip} b_{pq} c_{qj} \right].$$

2. Példa. Határozzuk meg az \mathbf{ABC} szorzatot, ha

$$\mathbf{A} = \begin{bmatrix} 1 & -2 \\ 2 & 4 \\ 3 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 2 & -3 \\ 1 & -1 & 2 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

Megoldás.

$$\begin{bmatrix} 1 & -2 \\ 2 & 4 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & -3 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 2 & 4 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} -1 & 4 & -7 \\ 6 & 0 & 2 \\ 3 & 6 & -9 \end{bmatrix} \begin{bmatrix} 0 & -5 \\ 8 & 14 \\ 6 & 3 \end{bmatrix}.$$

$$\text{Tehát } \mathbf{ABC} = \begin{bmatrix} 0 & -5 \\ 8 & 14 \\ 6 & 3 \end{bmatrix}.$$

* * *

Az összeadás és a szorzás definíciójából következik, hogy érvényes a *kétoldali disztributivitás* törvénye:

$$(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}, \quad \mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}.$$

Közvetlenül belátható, hogy akárhány tényezős mátrixszorzat esetén a szorzatmátrixnak annyi sora van, mint az első tényezőnek és annyi oszlopa, mint az utolsó tényezőnek.

Mivel egy kvadratikus mátrix szorzás szempontjából önmagával mindig felcserélhető, a mátrixszorzás asszociativitása folytán egyértelműen definiálható a kvadratikus mátrixok pozitív egész kitevőjű hatványa:

$$\mathbf{A}^p = \overset{1}{\mathbf{A}} \overset{2}{\mathbf{A}} \dots \overset{p}{\mathbf{A}}.$$

Innen adódik, hogy a számok hatványainak szorzatára vonatkozó szabály értelemszerűen átvihető mátrixokra is: közös alapú hatványokat úgy szorzunk, hogy a közös alapot a hatványkitevők összegére emeljük, vagyis

$$\mathbf{A}^p \mathbf{A}^q = \mathbf{A}^{p+q}.$$

A zérus kitevőjű hatványt egységmátrixként értelmezzük: $\mathbf{A}^0 = \mathbf{E}$.

Ezek alapján tehát egyértelműen definiálható egy kvadratikus mátrix skalár együtthatós polinomja. Ha ugyanis tekintjük a

$$p_m(z) \equiv c_0 + c_1 z + c_2 z^2 + \dots + c_m z^m$$

m -edfokú polinomot (c_0, c_1, c_2, \dots komplex számok, $c_m \neq 0$), akkor a z skalárváltozó helyére az \mathbf{A} mátrixot helyettesítve, a

$$p_m(\mathbf{A}) \equiv c_0 \mathbf{E} + c_1 \mathbf{A} + c_2 \mathbf{A}^2 + \dots + c_m \mathbf{A}^m$$

mátrixpolinomot kapjuk. A mátrixpolinom értelmezéséből következik, hogy ugyanazon \mathbf{A} kvadratikus mátrix tetszőleges polinomjaira a szorzás a közönséges polinomokra szokásos módon végezhető el és a mátrixpolinomok a szorzásra nézve felcserélhetők. Továbbá, hogy bármely *egyváltozós racionális egész skalár azonosság változatlanul érvényben marad*, ha a skalár változó helyébe egy *kvadratikus mátrixot*, az egység helyébe pedig az egységmátrixot helyettesítjük. Például $(1+x)(1-x) \equiv 1-x^2$ miatt, tetszőleges kvadratikus \mathbf{A} mátrixra fennáll az $(\mathbf{E} + \mathbf{A})(\mathbf{E} - \mathbf{A}) \equiv \mathbf{E} - \mathbf{A}^2$ azonosság.

A mátrixszorzat transzponálására az ún. *fordított sorrend* szabálya érvényes, azaz

$$(1.2.5) \quad (\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top.$$

Ugyanis a transzponált mátrix elemeire bevezetve az

$$\mathbf{A}^\top = [a_{ij}^\top] = [a_{ji}], \quad \mathbf{B}^\top = [b_{ij}^\top] = [b_{ji}]$$

jelölést, az

$$\mathbf{AB} = \left[\sum_{k=1}^n a_{ik} b_{kj} \right]$$

szorzat transzponálja

$$(\mathbf{AB})^T = \left[\sum_{k=1}^n a_{ik} b_{kj} \right]^T = \left[\sum_{k=1}^n a_{jk} b_{ki} \right] = \left[\sum_{k=1}^n b_{ki} a_{jk} \right] = \left[\sum_{k=1}^n b_{ik}^T a_{kj}^T \right] = \mathbf{B}^T \mathbf{A}^T.$$

A fordított sorrend szabályának ismételt alkalmazásával bármely véges m tényező esetére

$$(\mathbf{A}_1 \mathbf{A}_2 \dots \mathbf{A}_m)^T = \mathbf{A}_m^T \mathbf{A}_{m-1}^T \dots \mathbf{A}_2^T \mathbf{A}_1^T.$$

3. Példa. Határozzuk meg az \mathbf{ABC} szorzat transzponáltját, ha

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 & -1 \\ -4 & 0 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 2 & 1 \\ 3 & -2 \end{bmatrix}.$$

Megoldás.

$$(\mathbf{ABC})^T = \mathbf{C}^T \mathbf{B}^T \mathbf{A}^T = \begin{bmatrix} 2 & 3 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 2 & -4 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & -8 \\ 4 & -4 \end{bmatrix} \begin{bmatrix} -15 & -29 \\ -4 & -4 \end{bmatrix}.$$

$$\text{Tehát } (\mathbf{ABC})^T = \begin{bmatrix} -15 & -29 \\ -4 & -4 \end{bmatrix}.$$

* * *

A szorzás definíciójából következik, hogy ha az \mathbf{A} és \mathbf{B} mátrixra az összeszorozhatóság feltétele teljesül és az \mathbf{A} mátrix ugyanúgy van partícionálva oszlopai szerint, mint a \mathbf{B} mátrix sorai szerint (\mathbf{A} sorok szerinti és \mathbf{B} oszlopok szerinti partícionálása tetszőleges lehet), akkor az \mathbf{AB} szorzat a partícionált mátrixok blokkjainak segítségével is számítható; ha

$$\mathbf{A} = [\mathbf{A}_{ik}] \quad (i = 1, 2, \dots, p; \quad k = 1, 2, \dots, q)$$

és

$$\mathbf{B} = [\mathbf{B}_{kl}] \quad (k = 1, 2, \dots, q; \quad l = 1, 2, \dots, r),$$

akkor

$$\mathbf{AB} = \left[\sum_{k=1}^q \mathbf{A}_{ik} \mathbf{B}_{kl} \right] \quad (i = 1, 2, \dots, p; \quad l = 1, 2, \dots, r).$$

4. Példa. Határozzuk meg az \mathbf{AB} szorzatot, ahol

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 3 & -1 & 0 & 1 \\ -2 & 0 & 3 & -1 \\ 0 & -2 & 2 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 0 & 3 & -1 \\ 0 & 1 & 1 & 2 \\ 1 & 4 & 1 & 0 \\ -2 & 1 & 0 & 1 \end{bmatrix}.$$

Megoldás. Célszerű a mátrixokat másodrendű blokkokra particionálni:

$$\begin{aligned}
 & \begin{bmatrix} 1 & 0 & 3 & -1 \\ 0 & 1 & 1 & 2 \\ \hline 1 & 4 & 1 & 0 \\ -2 & 1 & 0 & 1 \end{bmatrix} \\
 & \begin{bmatrix} 1 & 2 & 1 & 0 \\ 3 & -1 & 0 & 1 \\ \hline -2 & 0 & 3 & -1 \\ 0 & -2 & 2 & 1 \end{bmatrix} \left[\begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 4 \\ -2 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \end{bmatrix} \right. \\
 & \left. \begin{bmatrix} -2 & -2 \end{bmatrix} + \begin{bmatrix} 3 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ -2 & 1 \end{bmatrix} - 2 \begin{bmatrix} 3 & -1 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 3 & -1 \\ 2 & 1 \end{bmatrix} \right] \\
 & \text{Az eredmény: } \begin{bmatrix} 2 & 6 & 6 & 3 \\ 1 & 0 & 8 & -4 \\ \hline 3 & 11 & -3 & 1 \\ 0 & 7 & 0 & -3 \end{bmatrix} \\
 & \qquad \qquad \qquad * \quad * \quad *
 \end{aligned}$$

Két mátrix szorzatának definíciójából közvetlenül következik, hogy a szorzatmátrix (1.1.1) szerinti minormátrixát megkapjuk, ha a bal oldali tényező i_1, i_2, \dots, i_r indexű sorából és a jobb oldali tényező j_1, j_2, \dots, j_s indexű oszlopából alkotott minormátrixot összeszorozzuk, vagyis

$$(1.2.6) \quad (\mathbf{AB})_{j_1 j_2 \dots j_s}^{i_1 i_2 \dots i_r} = \mathbf{A}_{1 \dots n}^{i_1 i_2 \dots i_r} \mathbf{B}_{j_1 j_2 \dots j_s}^{1 \dots n}.$$

5. Példa. Határozzuk meg az \mathbf{AB} szorzat másodrendű jobb felső minor-mátrixát, ha

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 5 \\ 2 & -4 & 1 & 0 \\ 6 & 8 & 10 & 12 \\ 10 & -6 & 9 & 7 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 5 & 20 & 1 & 3 \\ 9 & -11 & -2 & 1 \\ 8 & 5 & 1 & 1 \\ 7 & 17 & 0 & -1 \end{bmatrix}.$$

Megoldás.

$$(\mathbf{AB})_{34}^{12} = \mathbf{A}_{1234}^{12} \cdot \mathbf{B}_{34}^{1234} = \begin{bmatrix} 1 & 2 & 3 & 5 \\ 2 & -4 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ -2 & 1 \\ 1 & 1 \\ 0 & -1 \end{bmatrix}.$$

Jelölje most a megfelelő sorvektorokat $\mathbf{y}^T = [y_j]$ ($j = 1, 2, \dots, m$), ill. $\mathbf{c}^T = [c_l]$ ($l = 1, 2, \dots, n$); ekkor a szorzat a következő alakban írható fel:

$$(1.2.9) \quad \mathbf{y}_{(m,n)}^\top \mathbf{A}_{(m,n)} = \mathbf{c}^\top.$$

Részletesen kiírva a szorzatvektor elemeit:

$$(1.2.10) \quad \begin{aligned} y_1 a_{11} + y_2 a_{21} + y_3 a_{31} + \cdots + y_m a_{m1} &= c_1 \\ y_1 a_{12} + y_2 a_{22} + y_3 a_{32} + \cdots + y_m a_{m2} &= c_2 \\ &\vdots \\ y_1 a_{1n} + y_2 a_{2n} + y_3 a_{3n} + \cdots + y_m a_{mn} &= c_n. \end{aligned}$$

Mint látható, ismét lineáris algebrai egyenletrendszert kapunk, ennek együtthatómátrixa az (1.2.7) egyenlet együtthatómátrixának transzponáltja. Tehát az (1.2.10) egyenletrendszer

$$\mathbf{A}^\top \mathbf{y} = \mathbf{c}$$

alakba írható, ami egyébként az (1.2.9) összefüggés transzponálásával is megkapható.

A lineáris egyenletrendszereket részletesen az 1.7 szakaszban tárgyaljuk.

Ha egyetlen sorvektornak egyetlen oszlopvektorral való szorzatát tekintjük, akkor ez egy „elsőrendű” mátrixot, azaz egyetlen elemet ad, ami a két vektor *skaláris szorzatának* felel meg.

Vázlatosan:

$$\frac{\text{Area of rectangle}}{\text{Area of square}} = \frac{\text{Side of rectangle}}{\text{Side of square}}$$

részletesen:

$$(1.2.11) \quad \mathbf{a}^\top \mathbf{b} = \sum_{k=1}^n a_k b_k.$$

6. **Példa.** Határozzuk meg az \mathbf{a} és \mathbf{b} vektor skaláris szorzatát, ha $\mathbf{a}^\top = [1 \ 2 \ 3]$, $\mathbf{b}^\top = [3 \ -2 \ 1]$.

Megoldás.

$$\mathbf{a}^T \mathbf{b} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}.$$

Megjegyezzük, hogy mivel a skaláris szorzat eredménye egy szám, amelynek a transzponáltja önmaga, ezért a skaláris szorzat egyenlő a transzponáltjával:

$$(\mathbf{a}^\top \mathbf{b})^\top = \mathbf{b}^\top \mathbf{a} = \begin{bmatrix} 3 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

* * *

Fordított sorrendben véve a vektortényezőket, a továbbiak szempontjából alapvető jelentőségű szorzathoz, az ún. *diadikus szorzathoz* vagy röviden *diád-hoz* jutunk. Az m elemű \mathbf{a} oszlopvektornak az n elemű \mathbf{b}^\top sorvektorral való szorzata ugyanis egy $m \times n$ típusú mátrixhoz vezet, amelynek elemei az egyes vektortényezők egy-egy elemének páronként vett szorzatai.

Vázlatosan:

$$\boxed{} \boxed{} = \boxed{},$$

azaz

$$(1.2.12) \quad \mathbf{a} \mathbf{b}^\top = [a_i b_j] = \begin{bmatrix} a_1 b_1 & a_1 b_2 & \dots & a_1 b_n \\ a_2 b_1 & a_2 b_2 & \dots & a_2 b_n \\ \dots & \dots & \dots & \dots \\ a_m b_1 & a_m b_2 & \dots & a_m b_n \end{bmatrix}.$$

7. Példa. Határozzuk meg az előző példában szereplő vektorok $\mathbf{a} \mathbf{b}^\top$ diadikus szorzatát!

Megoldás.

$$\mathbf{a} \mathbf{b}^\top = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 3 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -2 & 1 \\ 6 & -4 & 2 \\ 9 & -6 & 3 \end{bmatrix}.$$

* * *

Adott mátrix oszlopai, sorai, elemei előállíthatók ilyen speciális mátrix-szorzatok alakjában. Ha ui. a mátrixszorzat egyik (vagy mindkét) tényezője egységvektor, akkor könnyen felírható a szorzat eredménye: az \mathbf{A} mátrixot jobbról megszorozva a j -edik egység-oszlopvektorral, a mátrix j -edik oszlopát kapjuk, ha viszont balról az i -edik egység-sorvektorral szorozzuk, akkor a mátrix i -edik sorát kapjuk:

$$(1.2.13) \quad \mathbf{A} \mathbf{e}_j = \mathbf{a}_j$$

Részletesen kiírva:

$$\begin{bmatrix} | & | & | & | \\ \mathbf{a}_1 & \dots & \mathbf{a}_j & \dots & \mathbf{a}_n \\ | & | & | & | \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{a}_j \end{bmatrix},$$

ill.

$$(1.2.14) \quad \mathbf{e}_i^\top \mathbf{A} = \mathbf{a}^i,$$

azaz

$$\begin{bmatrix} 0 \dots 1 & 0 \dots 0 \end{bmatrix} \begin{bmatrix} \mathbf{a}^1 \\ \vdots \\ \mathbf{a}^i \\ \vdots \\ \mathbf{a}^n \end{bmatrix} = \begin{bmatrix} \mathbf{a}^i \end{bmatrix}.$$

Hasonlóképpen könnyen beláthatók az alábbi egyszerű, de fontos összefüggések:

$$\begin{aligned} \mathbf{e}_i^\top \mathbf{a} &= a_i, \\ \mathbf{e}_i^\top \mathbf{e}_j &= \delta_{ij}, \quad (\text{Kronecker-féle szimbólum}) \end{aligned}$$

$$\mathbf{e}_i \mathbf{e}_j^\top = \mathbf{E}_{ij} = \begin{bmatrix} 0 & 0 \dots 0 \dots 0 \\ \vdots & \vdots \\ 0 & 0 \dots 1 \dots 0 \\ \vdots & \vdots \\ 0 & 0 \dots 0 \dots 0 \end{bmatrix} \overset{j}{\underbrace{\quad}} \overset{i}{\underbrace{\quad}} \quad (\text{mátrixegységek}),$$

valamint

$$(1.2.15) \quad \mathbf{e}_i^\top \mathbf{A} \mathbf{e}_j = \mathbf{e}_i^\top \mathbf{a}_j = \mathbf{a}^i \mathbf{e}_j = a_{ij}.$$

Az (1.2.15) összefüggés szerint tehát egy \mathbf{A} mátrix ij indexű elemét megkapjuk, ha az \mathbf{A} mátrixot az i -edik egység-sorvektorral balról és a j -edik egység-oszlopvektorral jobbról szorozzuk.

A diád fogalmának ismeretében lehetővé válik két mátrix szorzatának egy másik előállítás. Egy $m \times p$ és egy $p \times n$ típusú mátrix szorzatának az elemei ugyanis p tagú összegek, így a szorzatmátrix p számú mátrix összegeként írható fel. Az összeadandók a bal oldali tényező egy-egy oszlopából és a jobb oldali tényező egy-egy sorából alkotott diádok. Részletesen kiírva:

$$\begin{aligned}
 \mathbf{AB} &= && \text{a szorzat} \\
 &= \left[\sum_{k=1}^p a_{ik} b_{kj} \right] = && \text{mint olyan mátrix, melynek elemei } p \text{ tagú össze-} \\
 &= \sum_{k=1}^p [a_{ik} b_{kj}] = && \text{gekek,} \\
 (1.2.16) \quad &= \sum_{k=1}^p \mathbf{a}_k \mathbf{b}^k = && \text{ezt } p \text{ számú mátrix összegeként felírva,} \\
 &&& \text{a mátrixokról látható, hogy diádok, így } p \text{ számú} \\
 &&& \text{diád összegéhez jutunk;}
 \end{aligned}$$

$$\begin{aligned}
 \mathbf{A} &= \begin{bmatrix} | & | & | & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_p \\ | & | & | & | \end{bmatrix} && \text{a diádok oszlopvektorai az } \mathbf{A} \text{ mátrix oszlopai,} \\
 \mathbf{B} &= \begin{bmatrix} \mathbf{b}^1 \\ \hline \mathbf{b}^2 \\ \hline \vdots \\ \hline \mathbf{b}^p \end{bmatrix} && \text{sorvektorai pedig a } \mathbf{B} \text{ mátrix sorai.}
 \end{aligned}$$

Tehát két ($m \times p$ és $p \times n$ típusú) mátrix szorzata mindig felírható (p tagú) diádösszegeként. Ez egyébként a particionált mátrixok szorzásának szabályából közvetlenül is következik, ha a bal oldali tényezőt oszlopaira, a jobb oldalit pedig soraira particionáljuk.

Az (1.2.16) összefüggést fordítva olvasva belátható, hogy *n diád összege mindig felírható két mátrix szorzataként*; a bal oldali tényező az egyes diádok oszlopaiból, a jobb oldali tényező pedig a diádok soraiból alkotott mátrix:

$$\begin{aligned}
 (1.2.17) \quad & \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T & \vdots & \mathbf{v}_n^T \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} + \begin{bmatrix} \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} + \dots + \begin{bmatrix} \mathbf{u}_n \\ \vdots \\ \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \mathbf{v}_n^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} = \\
 &= \sum_{k=1}^n \begin{bmatrix} \mathbf{u}_k \\ \vdots \\ \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \mathbf{v}_k^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} = \begin{bmatrix} | & | & | & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_n \\ | & | & | & | \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \hline \mathbf{v}_2^T \\ \hline \vdots \\ \hline \mathbf{v}_n^T \end{bmatrix} = \mathbf{UV}^T.
 \end{aligned}$$

Ennek jelentőségét akkor látjuk majd, amikor egy mátrix szorzattá alakításának, azaz *faktorizációjának* feladatával kerülünk szembe. A mondottak értelmében ugyanis a faktorizáció feladata ekvivalens a mátrixnak diádösszegeként való felírásával.

A következőkben speciális hármas szorzatokat vizsgálunk meg. Tekintsük először azt az esetet, amikor az \mathbf{ABC} szorzat középső tényezője diagonálmátrix, azaz $\mathbf{B} = \langle b_k \rangle$. Az (1.2.4) összefüggésből következik, hogy ekkor

$$(1.2.18) \quad \mathbf{ABC} = \left[\sum_{k=1}^n a_{ik} b_k c_{kj} \right],$$

vagyis a hármas szorzat elemei most kettős szumma helyett *egyszeres szumma* alakjában adódnak. Vegyük ismét figyelembe, hogy az egyes elemek n tagú összegek; így a hármas szorzat n számú mátrix összegeként írható fel. Az n mátrix mindegyikéből most egy-egy skalár szorzó (b_1, b_2, \dots, b_n) rendre kiemelhető, az így megmaradó $[a_{ik} c_{kj}]$ mátrixok pedig az \mathbf{A} mátrix oszlopvektoraiból és a \mathbf{C} mátrix sorvektoraiból alkotott diádok. Ilyen módon belátuk, hogy *ha három olyan mátrix szorzatát tekintjük, ahol a középső tényező diagonálmátrix, akkor ez a szorzat az első tényező oszlopvektoraiból és a harmadik tényező sorvektoraiból alkotott diádoknak a második tényező elemeivel képezett lineáris kombinációja*.*

Részletesen kiírva:

$$\begin{aligned}
 \mathbf{ABC} &= && \text{a hármas szorzat} \\
 &= \left[\sum_{k=1}^n a_{ik} b_k c_{kj} \right] = && \text{mint olyan mátrix, melynek elemei } n \\
 &= \sum_{k=1}^n [a_{ik} b_k c_{kj}] = && \text{ezt } n \text{ számú mátrix összegeként felírva és} \\
 (1.2.19) \quad &= \sum_{k=1}^n b_k [a_{ik} c_{kj}] = && \text{mindegyik mátrixból a } b_k \text{ skalár szorzót} \\
 &= \sum_{k=1}^n b_k \mathbf{a}_k \mathbf{c}^k && \text{kiemelve,} \\
 & && n \text{ diád lineáris kombinációjához jutunk;} \\
 \mathbf{A} &= \left[\begin{array}{c|c|c|c} & & & \\ \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \\ & & & \end{array} \right] && \text{a diádok oszlopvektorai az } \mathbf{A} \text{ mátrix} \\
 & && \text{oszlopai,} \\
 \mathbf{C} &= \left[\begin{array}{c} \mathbf{c}^1 \\ \hline \mathbf{c}^2 \\ \hline \vdots \\ \hline \mathbf{c}^n \end{array} \right] && \text{sorvektorai a } \mathbf{C} \text{ mátrix sorai,} \\
 \mathbf{B} = \langle b_1 b_2 \dots b_n \rangle &= \left[\begin{array}{c} b_1 \\ b_2 \\ \ddots \\ b_n \end{array} \right] && \text{a lineáris kombináció együtthatói pedig} \\
 & && \text{a } \mathbf{B} \text{ diagonálmátrix elemei.}
 \end{aligned}$$

* Ismeretes, hogy az $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ vektorok lineáris kombinációján a $c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \dots + c_r \mathbf{a}_r$ összeget értjük, ahol c_i ($i = 1, \dots, r$) tetszőleges (komplex) számok.

Az (1.2.19) összefüggést fordítva olvasva, belátható, hogy diádok lineáris kombinációja mindig felírható három olyan mátrix szorzataként, ahol az első tényezőt a diádok oszlopvektorai, a harmadik tényezőt a diádok sorvektorai alkotják, a középső tényező pedig olyan diagonálmátrix, amelynek elemei a lineáris kombináció együtthatói.

Részletesen:

(1.2.20)

$$\sum_{k=1}^n \lambda_k \mathbf{u}_k \mathbf{v}_k^T = \begin{bmatrix} | & | & | & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_n \\ | & | & | & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} = \mathbf{U} \mathbf{\Lambda} \mathbf{V}^T.$$

Megjegyzés. Ennek a felírásnak a jelentőségét akkor fogjuk igazán megismerni, amikor mátrixok *sajátérték-feladatával* foglalkozunk majd.

Tekintsük most három olyan mátrix szorzatát, ahol az első tényező sorvektor, a harmadik tényező pedig oszlopvektor. A szorzás definíciójából következik, hogy a szorzat egyetlen skalár, amelyet kettős szumma alakjában tudunk felírni. Jelölje $\mathbf{A} = [a_{ij}]$ a középső tényezőt, $\mathbf{y}^H = [\bar{y}_i]$ és $\mathbf{x} = [x_j]$ pedig az első, ill. harmadik tényezőt (a továbbiakban – a 2.3 szakaszban – látjuk majd, hogy miért célszerű az első tényezőt valamilyen komplex elemű \mathbf{y}^T sorvektor konjugáltjaként felírni). Ekkor

$$(1.2.21) \quad \mathbf{y}^H \mathbf{A} \mathbf{x} = \sum_{j=1}^n \sum_{i=1}^n a_{ij} \bar{y}_i x_j.$$

Ha y_i és x_j változók, akkor az (1.2.21) kifejezés függvényt definiál. Az (1.2.21) alakú függvényt az y_i és x_j változók *bilineáris alakjának*, az \mathbf{A} mátrixot a *bilineáris alak mátrixának* nevezzük. Ha \mathbf{A} hermitikus mátrix, akkor a fenti kifejezést *hermitikus bilineáris alaknak* nevezzük és \mathbf{y} helyére az \mathbf{x} vektort írva, az így adódó

$$(1.2.22) \quad \mathbf{x}^H \mathbf{A} \mathbf{x} = \sum_{j=1}^n \sum_{i=1}^n a_{ij} \bar{x}_i x_j$$

kifejezést az x_i változók *hermitikus alakjának* nevezzük. Abban a speciális esetben, ha \mathbf{A} valós szimmetrikus mátrix és \mathbf{x} is valós, akkor az

$$(1.2.23) \quad \mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{j=1}^n \sum_{i=1}^n a_{ij} x_i x_j$$

kifejezést az x_i valós változók *kvadrátikus alakjának* nevezzük.

Könnyen belátható, hogy az hermitikus alak mindig *valós értékű függvény*. Ugyanis – tekintettel arra, hogy $\mathbf{A} = \mathbf{A}^H$ – fennáll

$$(\mathbf{x}^H \mathbf{A} \mathbf{x})^H = \mathbf{x}^H \mathbf{A}^H \mathbf{x} = \mathbf{x}^H \mathbf{A} \mathbf{x}.$$

A bilineáris, hermitikus és kvadratikus alakokkal részletesen majd a 2.3 szakaszban foglalkozunk.

8. Példa. Írjuk fel azt a kvadratikus alakot, amelynek a mátrixa

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 3 & -5 \\ 4 & -5 & 6 \end{bmatrix}.$$

Megoldás.

$$\begin{aligned} \mathbf{x}^T \mathbf{A} \mathbf{x} &= \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ 2 & 3 & -5 \\ 4 & -5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \\ &= x_1^2 + 3x_2^2 + 6x_3^2 + 4x_1x_2 + 8x_1x_3 - 10x_2x_3. \\ &\quad * \quad * \quad * \end{aligned}$$

9. Példa. Írjuk fel azt az hermitikus alakot, amelynek mátrixa

$$\mathbf{A} = \begin{bmatrix} 1 & 2+3i \\ 2-3i & 4 \end{bmatrix}.$$

Megoldás. Az hermitikus alak a következő:

$$\begin{aligned} \mathbf{x}^H \mathbf{A} \mathbf{x} &= \begin{bmatrix} \bar{x}_1 & \bar{x}_2 \end{bmatrix} \begin{bmatrix} 1 & 2+3i \\ 2-3i & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \\ &= \begin{bmatrix} \bar{x}_1 \cdot 1 + \bar{x}_2(2-3i) & \bar{x}_1(2+3i) + 4\bar{x}_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \\ &= \bar{x}_1x_1 + (2-3i)x_1\bar{x}_2 + (2+3i)\bar{x}_1x_2 + 4\bar{x}_2x_2 = \\ &= |x_1|^2 + 4 \operatorname{Re} x_1\bar{x}_2 + 6 \operatorname{Im} x_1\bar{x}_2 + 4|x_2|^2. \\ &\quad * \quad * \quad * \end{aligned}$$

1.2.4. Az inverz mátrix

Ebben a pontban megismerkedünk a mátrix inverzének, más elnevezéssel a reciprokmátrixnak a fogalmával; megvizsgáljuk létezésének szükséges és elégséges feltételét. Ehhez néhány új fogalom bevezetésére van szükség, mégpedig

az adjungált mátrix, valamint a szinguláris és a nonszinguláris mátrix fogalmára.

1.2.4 definíció. Az n -edrendű $\mathbf{A} = [a_{ij}]$ kvadratikus mátrix **adjungáltján** azt a mátrixot értjük, amelyet az \mathbf{A} mátrixból úgy nyerünk, hogy az a_{ij} elem helyére az \mathbf{A}^T transzponált ugyanazon helyén álló $a_{ij}^T = a_{ji}$ elemének algebrai komplementumát, vagyis az $(n-1)$ -edrendű A_{ji} előjeles aldetermináns értékét írjuk.

Tehát

$$(1.2.24) \quad \text{adj } \mathbf{A} = [A_{ji}].$$

10. Példa. Legyen

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 2 & -1 & 3 \end{bmatrix}.$$

Számítsuk ki ennek adjungáltját!

Megoldás. Az adjungáltnak pl. a jobb felső sarokban álló elemét úgy kapjuk, hogy meghatározzuk a bal alsó sarokban álló elemhez tartozó előjeles aldeterminánst:

$$A_{31} = (-1)^{3+1} \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} = +1(12 - 15) = -3.$$

Hasonló módon kiszámítva a többi elemet, az adjungáltra a következő adódik:

$$\text{adj } \mathbf{A} = \begin{bmatrix} 21 & -9 & -3 \\ 0 & -3 & 6 \\ -14 & 5 & -3 \end{bmatrix}.$$

Megjegyzés. A kezdő hajlamos arra, hogy a *transzponálásról*, valamint a *sakktábla-szabály* alkalmazásáról – tehát az előjelről – megfeledkezzék, ezért a hibás számolás forrását leggyakrabban itt kell keresni.

* * *

Az adjungált mátrix segítségével könnyen megfogalmazható a determinánssokra vonatkozó *kifejtési tétel*. Emlékeztetünk arra, hogy a determinánsok sorok, ill. oszlopok szerinti kifejtése a következő:

$$(1.2.25) \quad \sum_{k=1}^n a_{ik} A_{jk} = \delta_{ij} |\mathbf{A}|, \quad \text{bármely } i, j = 1, 2, \dots, n\text{-re,}$$

ill.

$$(1.2.26) \quad \sum_{k=1}^n a_{ki} A_{kj} = \delta_{ij} |\mathbf{A}|, \quad \text{bármely } i, j = 1, 2, \dots, n\text{-re,}$$

ahol δ_{ij} a Kronecker-féle szimbólum. Az adjungált mátrix segítségével (1.2.25) és (1.2.26) az

$$(1.2.27) \quad \mathbf{A} \cdot \text{adj } \mathbf{A} = |\mathbf{A}| \mathbf{E},$$

ill.

$$(1.2.28) \quad \text{adj } \mathbf{A} \cdot \mathbf{A} = |\mathbf{A}| \mathbf{E}$$

alakban írható. Ezekből két fontos összefüggés olvasható le:

- (1) bármely kvadratikus mátrix és adjungáltja felcserélhető,
- (2) bármely kvadratikus mátrixnak az adjungáltjával való szorzata a mátrix determinánsával szorzott egységmátrixot adja.

A 10. példán szemléltetve az (1.2.27) és (1.2.28) összefüggéseket:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} 21 & -9 & -3 \\ 0 & -3 & 6 \\ -14 & 5 & -3 \end{bmatrix} = \begin{bmatrix} -21 & & \\ & -21 & \\ & & -21 \end{bmatrix},$$

ill.

$$\begin{bmatrix} 21 & -9 & -3 \\ 0 & -3 & 6 \\ -14 & 5 & -3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 2 & -1 & 3 \end{bmatrix} = \begin{bmatrix} -21 & & \\ & -21 & \\ & & -21 \end{bmatrix},$$

ahonnan látható, hogy $|\mathbf{A}| = -21$.

A kapott összefüggések jelentőségének megvilágítása céljából bevezetünk néhány új fogalmat.

1.2.5 definíció. Az \mathbf{A} kvadratikus mátrixot **nemszinguláris mátrixnak** nevezzük, ha az elemeiből alkotott determináns zérustól különböző; ha a determináns zérussal egyenlő, akkor a mátrixot **szingulárisnak** nevezzük.

Ha az (1.2.27) és (1.2.28) azonosságban szereplő \mathbf{A} mátrix nemszinguláris, akkor az $|\mathbf{A}| \neq 0$ determinánssal oszthatunk, és így (1.2.27) és (1.2.28) helyett

$$(1.2.29) \quad \mathbf{A} \frac{\text{adj } \mathbf{A}}{|\mathbf{A}|} = \mathbf{E},$$

ill.

$$(1.2.30) \quad \frac{\text{adj } \mathbf{A}}{|\mathbf{A}|} \mathbf{A} = \mathbf{E}$$

írható. A kapott eredményt tételként is kimondhatjuk:

1.2.2 tétel. *Ha \mathbf{A} nemszinguláris, akkor létezik az*

$$(1.2.31) \quad \frac{\text{adj } \mathbf{A}}{|\mathbf{A}|}$$

mátrix és ezt az \mathbf{A} mátrixszal bármely sorrendben összeszorozva, az egység-mátrixot kapjuk.

1.2.6 definíció. *Ha az \mathbf{A} kvadratikus mátrixhoz hozzárendelhető olyan \mathbf{X} mátrix, amely kielégíti mind az*

$$(1.2.32) \quad \mathbf{AX} = \mathbf{E},$$

mind pedig az

$$(1.2.33) \quad \mathbf{XA} = \mathbf{E}$$

*egyenletet, akkor az \mathbf{A} mátrixot **invertálhatónak**, az \mathbf{X} mátrixot pedig **\mathbf{A} inverzének** nevezzük.*

Az 1.2.2 tétel alapján egy nemszinguláris \mathbf{A} mátrixnak mindig van inverze, mégpedig az $\frac{\text{adj } \mathbf{A}}{|\mathbf{A}|}$ mátrix. Felmerül azonban a kérdés, lehet-e több inverz, vagyis az (1.2.32) és (1.2.33) mátrixegyenleteknek lehet-e több közös megoldásuk? Erre a kérdésre válaszol a következő tétel.

1.2.3 tétel. *Ha az \mathbf{A} kvadratikus mátrixnak létezik inverze, akkor az egyértelmű.*

Bizonyítás. A tételt indirekt módszerrel bizonyítjuk. Tegyük fel, hogy állításunkkal ellentétben \mathbf{X} és \mathbf{Y} az \mathbf{A} mátrix két különböző inverze. Felhasználva a mátrixszorzás asszociativitását, az egységmátrix tulajdonságát és az inverz mátrix (1.2.32) és (1.2.33) definiáló egyenleteit,

$$\mathbf{Y} = \mathbf{YE} = \mathbf{Y}(\mathbf{AX}) = (\mathbf{YA})\mathbf{X} = \mathbf{EX} = \mathbf{X}$$

írható, ahonnan, a feltevéssel ellentétben, $\mathbf{Y} = \mathbf{X}$ adódik.*

*A bizonyításból kiolvasható, hogy a tétel a következő, erősebb alakban is érvényes: Ha \mathbf{X} kielégíti az (1.2.32), \mathbf{Y} pedig az (1.2.33) egyenletet, akkor $\mathbf{X} = \mathbf{Y}$ és \mathbf{A} inverze egyértelmű.

Ha tehát egy \mathbf{A} mátrix invertálható, akkor egyetlen inverze van, így ebben az esetben az (1.2.32)–(1.2.33) egyenletek megoldására bevezethetjük az $\mathbf{X} = \mathbf{A}^{-1}$ jelölést. A következő tétel azt mondja ki, hogy mikor invertálható az \mathbf{A} mátrix, azaz milyen feltételek esetén létezik \mathbf{A} inverze.

1.2.4 tétel. *Az \mathbf{A} kvadratikus mátrix akkor és csak akkor invertálható, ha nonszinguláris, azaz ha $|\mathbf{A}| \neq 0$.*

Bizonyítás. A feltétel szükségessége közvetlenül adódik a determinánsok szorzási szabályának alkalmazásából. Ha ugyanis tekintjük az (1.2.32) vagy az (1.2.33) egyenletben mindkét oldal determinánsát és kihasználjuk azt, hogy két mátrix szorzatának determinánsa egyenlő az egyes tényezők determinánsának szorzatával (ami a determinánsok sor-oszlop szerinti szorzásából közvetlenül következik), akkor akár az $|\mathbf{A}\mathbf{X}| = |\mathbf{E}|$, akár az $|\mathbf{X}\mathbf{A}| = |\mathbf{E}|$ összefüggésből $|\mathbf{A}||\mathbf{X}| = 1$ adódik, ami csakis $|\mathbf{A}| \neq 0$ esetén állhat fenn.

Az 1.2.2 tétel szerint nonszinguláris \mathbf{A} esetén az $\frac{\text{adj } \mathbf{A}}{|\mathbf{A}|}$ mátrix kielégíti az inverz mátrix definiáló egyenleteit, ami azt bizonyítja, hogy a nonszingularitás egyúttal elégséges feltétele is az invertálhatóságnak. ■

Ha \mathbf{A} nonszinguláris, akkor az 1.2.2 tétel egyúttal \mathbf{A} inverzének az előállítását is megadja:

$$(1.2.34) \quad \mathbf{A}^{-1} = \frac{\text{adj } \mathbf{A}}{|\mathbf{A}|}.$$

11. Példa. Határozzuk meg az előző példában szereplő mátrix inverzét!

Megoldás. Felhasználjuk, hogy \mathbf{A} determinánsát már kiszámítottuk: $|\mathbf{A}| = -21$, így az inverz mátrix

$$\mathbf{A}^{-1} = -\frac{1}{21} \begin{bmatrix} 21 & -9 & -3 \\ 0 & -3 & 6 \\ -14 & 5 & -3 \end{bmatrix}.$$

Mátrixszorzat inverzének képzésére is a „fordított sorrend” szabálya érvényes: *többtényezős mátrixszorzat inverze az egyes tényezők inverzének fordított sorrendben vett szorzata.* Ezt közvetlen szorzással igazolhatjuk, ugyanis

$$(\mathbf{ABC})(\mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}) = \mathbf{AB}(\mathbf{CC}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}) = \mathbf{A}(\mathbf{BB}^{-1})\mathbf{A}^{-1} = \mathbf{AA}^{-1} = \mathbf{E},$$

tehát valóban

$$(1.2.35) \quad (\mathbf{ABC})^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}.$$

* * *

Megjegyzés. Az inverz mátrix itt ismertetett alakjának elsősorban elméleti jelentősége van, hiszen nagyméretű (nagy rendszámú) mátrixok esetén az aldeterminánsok kiszámítása óriási munkát igényel, megfelelő numerikus módszerek alkalmazásával viszont az inverz meghatározásának munkája lényegesen lecsökkenthető.

Vannak olyan feladatok, amelyekben célszerű az (1.2.34) definiáló összefüggést használni. Elsősorban akkor jöhet ez számításba, ha az invertálandó mátrix szabályos szerkezetű, vagyis elemei szabályos elhelyezkedést mutatnak, és a rendszám függvényében kívánjuk az inverz mátrix elemeit megadni.

1.3 A MÁTRIX RANGJA

Az 1.2.4 pontban megmutattuk, hogy csak a nemszinguláris mátrixok invertálhatók. Tehát, ellentétben a skalármennyiségekkel, amelyek a zérus kivételével mind invertálhatók, a mátrixok körében a zérusmátrixon kívül még nagyon sokféle mátrix létezik, amelynek nincs inverze. Feltehetjük a kérdést, vajon vannak-e fokozatok a zérusmátrix és a nemszinguláris (tehát invertálható) mátrixok között, azaz mérhető-e valamilyen jellemző számmal az, hogy egy mátrix „mennyire” szinguláris. Erre a kérdésre ad választ a mátrix *rangjának* fogalma.

Az 1.3.1 pontban a mátrix rangjára három definíciót is adunk, és megmutatjuk, hogy ezek a definíciók ekvivalensek. Ennek során bevezetjük a vektorok lineáris függetlenségének, valamint a mátrixok minimális diadikus előállításának fogalmát. Megmutatjuk, hogy a minimális diadikus előállítás egyúttal a mátrixok egy lehetséges faktorizációját szolgáltatja. Látni fogjuk, hogy ha a rangot mint a minimális diadikus előállításban szereplő diádok számát definiáljuk, akkor könnyen bizonyítható a mátrixok rangjának becslésére vonatkozó több tétel.

1.3.1 A rang fogalma; mátrixok minimális diadikus felbontása

A rang fogalmát legegyszerűbb az aldeterminánsok segítségével definiálni.

1.3.1 definíció. Egy \mathbf{A} mátrix $\varrho(\mathbf{A})$ rangja az \mathbf{A} elemeiből kiválasztható legmagasabbrendű, zérustól különböző értékű aldeterminánsnak a rendszáma.

Tehát ha az \mathbf{A} mátrix elemeiből alkotott r -edrendű aldeterminánsok között van legalább egy, amelynek értéke zérustól különböző, de valamennyi $r + 1$ -edrendű és ennél magasabb rendű aldeterminánsának értéke zérus, akkor a mátrix rangja r .

Jelölése: $\varrho(\mathbf{A}) = r$.

A definíció közvetlen következménye, hogy n -edrendű nemszinguláris mátrix rangja n . A mátrix rangjának definíciójából az is következik, hogy

ha egy mátrix sorait vagy oszlopait átrendezzük, akkor rangja nem változik meg. Mivel a sorok, ill. az oszlopok átrendezése azt jelenti, hogy a mátrixot balról, ill. jobbról megfelelő permutáló mátrixszal szorozzuk, azt is mondhatjuk, hogy a *mátrix rangja permutáló mátrixszal való szorzással szemben invariáns*.

A rang 1.3.1 definíciója alapján könnyen belátható, hogy egy diád rangja 1, hiszen egy diád sorai (ill. oszlopai) egymástól csupán állandó szorzóban különböznek, tehát bárhogyan választunk is belőle másodrendű (vagy magasabbrendű) aldeterminánst, az zérus lesz. Eszerint $\mathbf{a} \neq \mathbf{0}$, $\mathbf{b} \neq \mathbf{0}$ esetén

$$(1.3.1) \quad \varrho(\mathbf{ab}^\top) = 1.$$

Ez a felismerés azt az ötletet sugallja, hogy a mátrix rangját valamiképpen a mátrixot alkotó diádok számával hozzuk összefüggésbe. Minden mátrix felírható ugyanis diádok összegeként. Egy ilyen triviális felírásmód adódik például, ha a mátrixot oszlopaira particionáljuk és megszorozzuk a soraira particionált egységmátrixszal:

$$\mathbf{A} = \mathbf{AE} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_n] \begin{bmatrix} \mathbf{e}_1^\top \\ \mathbf{e}_2^\top \\ \vdots \\ \mathbf{e}_n^\top \end{bmatrix} = \sum_{k=1}^n \mathbf{a}_k \mathbf{e}_k^\top.$$

Egy mátrix diádösszegeként való előállításai között kitüntetett szerepe van azoknak, amelyek a mátrixot a lehető legkevesebb számú diád összegeként állítják elő. Így jutunk a mátrix *minimális diadikus előállításának a fogalmához* (lásd [25], [26]).

1.3.2 definíció. *Ha egy mátrixot a lehető legkevesebb számú diád összegeként állítunk elő, akkor ezt a mátrix egy **minimális diadikus előállításának (felbontásának)** nevezzük.*

A következőkben egy eljárást mutatunk be, amellyel adott mátrix minimális diadikus előállítása megkonstruálható, majd bebizonyítjuk, hogy a minimális diadikus előállításban szereplő diádok száma egyenlő a mátrix rangjával.

Ehhez előbb bevezetjük az $\mathbf{A} = [a_{ij}]$ mátrix $a_{ij} \neq 0$ eleme által generált *diád* fogalmát: ezen a mátrix j -edik oszlopának és i -edik sorának szorzatát értjük, osztva még az a_{ij} elemmel:

$$\frac{\mathbf{a}_j \mathbf{a}_i^\top}{a_{ij}} = \frac{\mathbf{A} \mathbf{e}_j \mathbf{e}_i^\top \mathbf{A}}{\mathbf{e}_i^\top \mathbf{A} \mathbf{e}_j}; \quad \text{ahol} \quad a_{ij} = \mathbf{e}_i^\top \mathbf{A} \mathbf{e}_j \neq 0.$$

Tekintsük tehát az $\mathbf{A} = [a_{ij}]$ kvadratikus mátrixot és tegyük fel, hogy a bal felső sarokeleme nem zérus: $a_{11} \neq 0$. Ez a sorok és az oszlopok megfelelő

© Rózsa Pál

Mint látható, az $\mathbf{A}^{(3)}$ mátrix első két sora és első két oszlopa már csupa zérus elemet tartalmaz. Az eljárást addig folytatjuk, amíg a zérusmátrixhoz jutunk. Tegyük fel, hogy ezt s lépés után érjük el:

$$\mathbf{A}^{(s)} - \frac{\mathbf{A}^{(s)} \mathbf{e}_s \mathbf{e}_s^T \mathbf{A}^{(s)}}{\mathbf{e}_s^T \mathbf{A}^{(s)} \mathbf{e}_s} = \mathbf{0}.$$

A vázolt eljárás az alábbi rekurziós formulával jellemezhető:

$$(1.3.2) \quad \mathbf{A}^{(k+1)} = \mathbf{A}^{(k)} - \frac{\mathbf{A}^{(k)} \mathbf{e}_k \mathbf{e}_k^T \mathbf{A}^{(k)}}{\mathbf{e}_k^T \mathbf{A}^{(k)} \mathbf{e}_k}; \quad \mathbf{A}^{(1)} = \mathbf{A}, \quad \mathbf{A}^{(s+1)} = \mathbf{0}.$$

Összeadva az egyenlőségeket $k = 1, \dots, s$ -re és rendezve:

$$(1.3.3) \quad \mathbf{A} = \sum_{k=1}^s \frac{\mathbf{A} \mathbf{e}_k \mathbf{e}_k^T \mathbf{A}}{\mathbf{e}_k^T \mathbf{A} \mathbf{e}_k}$$

adódik, vagyis az \mathbf{A} mátrixot s diád összegére bontottuk. Egyszerűség kedvéért jelölje ezeket a diádokat a továbbiakban

$$\mathbf{u}_k \mathbf{v}_k^T \quad (k = 1, 2, \dots, s).$$

Az \mathbf{u}_k , ill. \mathbf{v}_k^T vektorokat – a fenti konstrukcióból következően – az jellemzi, hogy a k index növekedésével egyre több zérus elemet tartalmaznak. Vázlatosan szemléltetve:

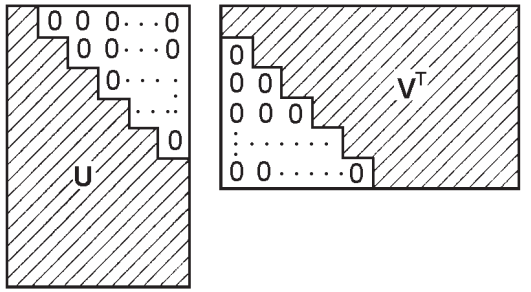
$$\mathbf{A} = \begin{array}{|c|} \hline \text{[diagonal elements]} \\ \hline \end{array} = \begin{array}{|c|} \hline \mathbf{u}_1 \\ \hline \end{array} \begin{array}{|c|} \hline \mathbf{v}_1^T \\ \hline \end{array} + \begin{array}{|c|} \hline \mathbf{u}_2 \\ \hline \end{array} \begin{array}{|c|} \hline \mathbf{v}_2^T \\ \hline \end{array} + \dots + \begin{array}{|c|} \hline \mathbf{u}_s \\ \hline \end{array} \begin{array}{|c|} \hline \mathbf{v}_s^T \\ \hline \end{array}.$$

The diagram illustrates the decomposition of matrix A into a sum of rank-1 matrices (diads). Each term in the sum is represented by a vertical vector u_k and a horizontal vector v_k^T. The vectors u_k have an increasing number of leading zeros as k increases, while the vectors v_k^T have an increasing number of trailing zeros. For example, u_1 has no leading zeros, while u_s has s-1 leading zeros. Similarly, v_1^T has no trailing zeros, while v_s^T has s-1 trailing zeros. The diagram shows the first three terms explicitly and then uses an ellipsis and a brace to indicate the remaining terms up to s.

Mint az (1.2.16) képletben láttuk, s számú diád összege mindig felírható a diádok oszlopvektoraiból, ill. sorvektoraiból alkotott s oszlopos, ill. s soros mátrix szorzataként, ezzel tehát az \mathbf{A} mátrixot két tényező szorzatára bontottuk, azaz *faktorizáltuk*:

$$(1.3.4) \quad \mathbf{A} = \sum_{k=1}^s \mathbf{u}_k \mathbf{v}_k^T = \mathbf{U} \mathbf{V}^T,$$

vázlatosan:



(A mátrixok sematikus vázolásával a zérus elemek elhelyezkedése jól szemléltethető.)

Most megmutatjuk, hogy az (1.3.3) diádösszegben a diádok száma nem csökkenthető, vagyis érvényes a következő tétel.

1.3.1 tétel. *Bármely \mathbf{A} mátrix egy minimális diadikus előállítását megkapjuk, ha azt az (1.3.2) rekurzív összefüggés segítségével (1.3.3) diádösszegként írjuk fel.*

Bizonyítás. A bizonyításhoz felhasználjuk az \mathbf{A} mátrix (1.3.4) faktorizált alakját. Először is megállapítjuk, hogy az \mathbf{A} mátrix s -edrendű bal felső sarokminora szükségképpen zérustól különböző. Alkalmazzuk ugyanis a szorzatmátrix minormátrixának előállítására szolgáló (1.2.6) összefüggést:

$$(1.3.5) \quad \left| \mathbf{A}_{1 \ 2 \dots s}^1 \ 2 \dots s \right| = \left| \mathbf{U}_{1 \ 2 \dots s}^1 \ 2 \dots s \mathbf{V}_{1 \ 2 \dots s}^{T^1 \ 2 \dots s} \right| = \left| \mathbf{U}_{1 \ 2 \dots s}^1 \ 2 \dots s \right| \cdot \left| \mathbf{V}_{1 \ 2 \dots s}^{T^1 \ 2 \dots s} \right| = \\ = \prod_{k=1}^s u_{kk} v_{kk} = \prod_{k=1}^s a_{kk}^{(k)} \neq 0.$$

Mármost azt, hogy az előállítás minimális, indirekt módszerrel bizonyítjuk be, azaz feltesszük, hogy az \mathbf{A} mátrix kevesebb, pl. $s-1$ diád összegeként is előállítható:

$$\mathbf{A} = \sum_{k=1}^{s-1} \tilde{\mathbf{u}}_k \tilde{\mathbf{v}}_k^T.$$

Itt a jobb oldali összeg nem változik meg, ha egy olyan diádot adunk hozzá, amelynek a sorvektora csupa zérus elemet tartalmaz, hiszen az ilyen diád a zérusmátrixot jelenti:

$$\mathbf{A} = \sum_{k=1}^{s-1} \tilde{\mathbf{u}}_k \tilde{\mathbf{v}}_k^T + \tilde{\mathbf{u}}_s \cdot \mathbf{0}.$$

Ugyanezt faktorizált alakban is felírhatjuk:

$$\mathbf{A} = \begin{bmatrix} \tilde{\mathbf{u}}_1 & \tilde{\mathbf{u}}_2 & \dots & \tilde{\mathbf{u}}_s \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{v}}_1^T \\ \tilde{\mathbf{v}}_2^T \\ \vdots \\ \tilde{\mathbf{v}}_{s-1}^T \\ 0 \quad 0 \quad \dots \quad 0 \end{bmatrix} = \tilde{\mathbf{U}}_s \tilde{\mathbf{V}}_s^T.$$

Ha most kiszámítjuk az \mathbf{A} mátrix bármelyik s -edrendű minorát, akkor azt kapjuk hogy az zérussal egyenlő:

$$|\mathbf{A}_{j_1 j_2 \dots j_s}^{i_1 i_2 \dots i_s}| = |\tilde{\mathbf{U}}_{1 \ 2 \dots s}^{i_1 i_2 \dots i_s} \tilde{\mathbf{V}}_{j_1 j_2 \dots j_s}^{1 \ 2 \dots s}| = |\tilde{\mathbf{U}}_{1 \ 2 \dots s}^{i_1 i_2 \dots i_s}| \cdot |\tilde{\mathbf{V}}_{j_1 j_2 \dots j_s}^{1 \ 2 \dots s}| = 0,$$

mivel a \mathbf{V}^T mátrix utolsó sorának minden eleme zérus. Ezzel ellentmondásra jutottunk, hiszen láttuk, hogy az \mathbf{A} mátrixnak van s -edrendű, zérustól különböző aldeterminánsa. Tehát hamis az a kiinduló feltevés, mely szerint a mátrix s -nél kevesebb diád összegeként is előállítható. Ezzel bebizonyítottuk, hogy az (1.3.3) diadikus előállítás valóban *minimális*. ■

A következőkben azt bizonyítjuk be, hogy a mátrix 1.3.1 definíció szerinti r rangja éppen s , azaz $\rho(\mathbf{A}) = r = s$.

1.3.2 tétel. *Bármely mátrix egy minimális diadikus előállításában a diádok száma egyenlő a mátrix elemeiből kiválasztható legmagasabbrendű, zérustól különböző értékű aldeterminánsának a rendjével.*

Bizonyítás. Mivel azt már beláttuk [(1.3.5) alapján], hogy a mátrix s -edrendű bal felső sarokminora zérustól különböző, elegendő azt megmutatni, hogy a mátrix bármelyik $s + 1$ -edrendű (és magasabb rendű) minora zérus. Az 1.3.1 tétel bizonyításában alkalmazott gondolatmenetet követve, a mátrix (1.3.4) alatti felbontásának a jobb oldalához adjunk hozzá egy olyan diádot, amelynek sorvektora csupa zérus elemet tartalmaz:

$$\mathbf{A} = \sum_{k=1}^s \mathbf{u}_k \mathbf{v}_k^T + \hat{\mathbf{u}}_{s+1} \cdot \mathbf{0},$$

vagy faktorizált alakban:

$$\mathbf{A} = \begin{bmatrix} | & | & & | & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_s & \hat{\mathbf{u}}_{s+1} \\ | & | & & | & | \end{bmatrix} \begin{bmatrix} \text{---} \mathbf{v}_1^\top \text{---} \\ \text{---} \mathbf{v}_2^\top \text{---} \\ \vdots \\ \text{---} \mathbf{v}_s^\top \text{---} \\ 0 \ 0 \ 0 \ 0 \ 0 \end{bmatrix} = \hat{\mathbf{U}} \hat{\mathbf{V}}^\top.$$

Most számítsuk ki az \mathbf{A} mátrix *bármelyik* $s + 1$ -edrendű minorának értékét a kapott faktorizáció felhasználásával:

$$\left| \mathbf{A}_{j_1 j_2 \dots j_{s+1}}^{i_1 i_2 \dots i_{s+1}} \right| = \left| \hat{\mathbf{U}}_{1 \ 2 \dots s+1}^{i_1 i_2 \dots i_{s+1}} \hat{\mathbf{V}}_{j_1 j_2 \dots j_{s+1}}^{1 \ 2 \dots s+1} \right| = \left| \hat{\mathbf{U}}_{1 \ 2 \dots s+1}^{i_1 i_2 \dots i_{s+1}} \right| \cdot \left| \hat{\mathbf{V}}_{j_1 j_2 \dots j_{s+1}}^{1 \ 2 \dots s+1} \right| = 0,$$

hiszen a $\hat{\mathbf{V}}^\top$ mátrix utolsó sorának minden eleme zérus. Innen következik, hogy a mátrix valamennyi ennél magasabb rendű minora is zérus, tehát a rangja valóban megegyezik a minimális diadikus előállításban szereplő diádok számával. ■

Ezzel egyben megmutattuk azt is, hogy a rang 1.3.1 definíciója ekvivalens a következő definícióval.

1.3.3 definíció. *A mátrix rangja egy minimális diadikus előállításában szereplő diádok száma.*

Megjegyzés. Ennek a definíciónak az a jelentősége, hogy ezzel egyrészt könnyen kezelhető (és könnyen gépesíthető) eljárást találtunk mátrixoknak a rang értékét is automatikusan szolgáltató faktorizációjára, másrészt a mátrixok rangjára vonatkozó számos tétel könnyen bizonyítható a segítségével.

A mátrixok minimális diadikus előállítását az alábbi számpéldán mutatjuk be.

12. Példa. Határozzuk meg az

$$\mathbf{A} = \begin{bmatrix} 3 & 3 & 6 & 5 & 5 \\ 7 & 4 & 7 & 2 & 0 \\ -1 & -2 & -3 & -4 & -5 \\ -1 & -3 & -8 & -9 & -10 \end{bmatrix}$$

mátrix egy minimális diadikus előállítását, és ennek alapján írjuk fel a mátrixot két tényező szorzataként!

Megoldás. Az (1.3.2) rekurziós képlet felhasználásával

$$\begin{aligned} \mathbf{A} - \frac{1}{3} \begin{bmatrix} 3 \\ 7 \\ -1 \\ -1 \end{bmatrix} [3 \ 3 \ 6 \ 5 \ 5] &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & -3 & -7 & -\frac{29}{3} & -\frac{35}{3} \\ 0 & -1 & -1 & -\frac{7}{3} & -\frac{10}{3} \\ 0 & -2 & -6 & -\frac{22}{3} & -\frac{25}{3} \end{bmatrix} = {}^{(2)}\mathbf{A}; \\ {}^{(2)}\mathbf{A} - \left(-\frac{1}{3}\right) \begin{bmatrix} 0 \\ -3 \\ -1 \\ -2 \end{bmatrix} [0 \ -3 \ -7 \ -\frac{29}{3} \ -\frac{35}{3}] &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{4}{3} & \frac{8}{9} & \frac{5}{9} \\ 0 & 0 & -\frac{4}{3} & -\frac{8}{9} & -\frac{5}{9} \end{bmatrix} = {}^{(3)}\mathbf{A}; \\ {}^{(3)}\mathbf{A} - \frac{3}{4} \begin{bmatrix} 0 \\ 0 \\ \frac{4}{3} \\ -\frac{4}{3} \end{bmatrix} [0 \ 0 \ \frac{4}{3} \ \frac{8}{9} \ \frac{5}{9}] &= \mathbf{0}. \end{aligned}$$

Az adott \mathbf{A} mátrix egy minimális diadikus előállítása tehát a következő:

$$\begin{aligned} \mathbf{A} &= \frac{1}{3} \begin{bmatrix} 3 \\ 7 \\ -1 \\ -1 \end{bmatrix} [3 \ 3 \ 6 \ 5 \ 5] + \left(-\frac{1}{3}\right) \begin{bmatrix} 0 \\ -3 \\ -1 \\ -2 \end{bmatrix} [0 \ -3 \ -7 \ -\frac{29}{3} \ -\frac{35}{3}] + \\ &+ \frac{3}{4} \begin{bmatrix} 0 \\ 0 \\ \frac{4}{3} \\ -\frac{4}{3} \end{bmatrix} [0 \ 0 \ \frac{4}{3} \ \frac{8}{9} \ \frac{5}{9}]. \end{aligned}$$

Ha a diádok sorvektorát beszorozzuk az egyes diádok együtthatójával (a generáló elemek reciprokával) és a diádösszeget az oszlopvektorokból, ill. a sorvektorokból alkotott mátrixok szorzataként írjuk fel, akkor az adott mátrix alábbi faktorizációját nyerjük:

$$\mathbf{A} = \begin{bmatrix} 3 & 0 & 0 \\ 7 & -3 & 0 \\ -1 & -1 & \frac{4}{3} \\ -1 & -2 & -\frac{4}{3} \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 & \frac{5}{3} & \frac{5}{3} \\ 0 & 1 & \frac{7}{3} & \frac{29}{9} & \frac{35}{9} \\ 0 & 0 & 1 & \frac{2}{3} & \frac{5}{12} \end{bmatrix}.$$

Mint látható, ezzel nemcsak azt értük el, hogy az egyes tényezők – a zérus elemek elhelyezkedése folytán – „derékszögű trapéz” alakúak, hanem azt is,

hogy a második tényező kk indexű elemei az egységgel egyenlők.

* * *

A mátrixok minimális diadikus előállítására vonatkozó 1.3.1 tétel következménye, hogy a nemszinguláris mátrixok – esetleg a sorok és oszlopok megfelelő átrendezése után – egy alsó és egy felső háromszögmátrix szorzataként írhatók fel (lásd az 1.4.6 definíciót). Ezt a mátrix *trianguláris faktorizációjának* nevezzük. Ha az alsó háromszögmátrix főátlójában lévő elemeket jobbra, a felső háromszögmátrix főátlójában lévő elemeket balra kiemeljük és ezek $u_k \neq 0$ szorzatát egy diagonálmátrix elemének tekintjük, akkor a mátrix trianguláris faktorizációjára

$$(1.3.6) \quad \mathbf{A} = \begin{bmatrix} 1 & & & & \\ * & 1 & & & \\ * & * & 1 & & \\ & & & \ddots & \\ * & * & * & \dots & 1 \end{bmatrix} \begin{bmatrix} u_1 & & & & \\ & u_2 & & & \\ & & u_3 & & \\ & & & \ddots & \\ & & & & u_n \end{bmatrix} \begin{bmatrix} 1 & * & * & & * \\ & 1 & * & & * \\ & & 1 & & * \\ & & & \ddots & \vdots \\ & & & & 1 \end{bmatrix}$$

adódik. Bebonyítjuk a következő tételt:

1.3.3 tétel. *Ha a nemszinguláris \mathbf{A} mátrix trianguláris faktorizációja (1.3.6) alakú, akkor*

$$(1.3.7) \quad u_k = \frac{D_k}{D_{k-1}},$$

ahol D_k az \mathbf{A} mátrix k -adrendű bal felső sarokminora.

Bizonyítás. Az 1.3.1 tétel bizonyítása során követett gondolatmenet alapján, az (1.3.6) összefüggés mindkét oldalán képezve a k -adrendű bal felső sarokminorokat:

$$(1.3.8) \quad D_k = u_1 u_2 \dots u_k$$

és innen azonnal következik (1.3.7), valamint az is, hogy valamennyi bal felső sarokminor zérustól különböző. ■

1.3.2 Vektorok lineáris függetlensége

A következőkben bevezetjük a vektorok lineáris függetlenségének, ill. lineáris összefüggésének a fogalmát.

1.3.4 definíció. *Az $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ (zérustól különböző) vektorokat akkor mondjuk lineárisan függetlennek, ha az*

$$(1.3.9) \quad x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \dots + x_r \mathbf{a}_r = \mathbf{0}$$

összefüggés csak

$$(1.3.10) \quad x_1 = x_2 = \dots = x_r = 0$$

esetén áll fenn.

Ha viszont található olyan x_1, x_2, \dots, x_r értékrendszer, hogy az x_i értékek közül *legalább* egy zérustól különböző, és az (1.3.9) összefüggés fennáll, akkor az $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ vektorok *lineárisan összefüggőek*.

13. Példa. Állapítsuk meg, hogy az

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}; \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

vektorok lineárisan függetlenek-e!

Megoldás. Képezzük az adott vektorok lineáris kombinációját valamilyen c_1, c_2, c_3 együtthatókkal:

$$c_1 \mathbf{e}_1 + c_2 \mathbf{e}_2 + c_3 \mathbf{e}_3 = \begin{bmatrix} c_1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ c_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ c_3 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}.$$

Mint látható, ez csakis akkor lehet zérus, ha $c_1 = c_2 = c_3 = 0$, azaz ha a lineáris kombináció valamennyi együtthatója zérus. Az $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ vektorok tehát lineárisan függetlenek.

* * *

14. Példa. Tekintsük az

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}; \quad \mathbf{a}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}; \quad \mathbf{a}_3 = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$$

vektorokat és győződjünk meg arról, hogy ezek lineárisan összefüggőek!

Megoldás. Nem nehéz észrevenni, hogy az első és harmadik vektor összege a második vektor kétszeresét adja, tehát $\mathbf{a}_1 - 2\mathbf{a}_2 + \mathbf{a}_3 = \mathbf{0}$, és ezzel találtunk olyan

$$c_1 = 1, \quad c_2 = -2, \quad c_3 = 1$$

együtthatókat, amelyek közül *legalább egy* nem zérus (esetünkben egyik sem zérus) és az adott vektorok velük képzett lineáris kombinációja zérus, tehát az adott vektorok valóban lineárisan összefüggőek.

* * *

15. Példa. Bizonyítsuk be, hogy egy lineárisan független vektorrendszer a zérusvektorral kiegészítve mindig lineárisan összefüggő vektorrendszert ad.

Megoldás. Tekintsük az $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ lineárisan független vektorrendszert, amelyre definíció szerint az (1.3.9) összefüggés csak akkor állhat fenn, ha (1.3.10) is teljesül. Egészítsük ki az adott rendszert a zérusvektorral, és képezzük az így nyert $r + 1$ vektor lineáris kombinációját:

$$x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \dots + x_r \mathbf{a}_r + x_{r+1} \cdot \mathbf{0} = \mathbf{0}.$$

Nyilvánvaló, hogy az $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r, \mathbf{0}$ vektoroknak az

$$x_1 = x_2 = \dots = x_r = 0, \quad x_{r+1} \neq 0$$

együtthatókkal képezett lineáris kombinációja zérus, tehát a zérusvektorral kiegészített rendszer valóban lineárisan összefüggő.

* * *

A vektorok lineáris függetlenségének, ill. lineáris összefüggésének definícióját az alábbi tömörebb alakban is megfogalmazhatjuk, ha bevezetjük az $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ oszlopvektorokból álló \mathbf{A} mátrixot és az x_1, x_2, \dots, x_r elemű \mathbf{x} oszlopvektort.

1.3.5 definíció. Az \mathbf{A} mátrix oszlopvektorai lineárisan függetlenek, ha az

$$(1.3.11) \quad \mathbf{Ax} = \mathbf{0}$$

egyenletnek csak az

$$(1.3.12) \quad \mathbf{x} = \mathbf{0}$$

triviális megoldása van.

Az (1.3.11) egyenletnek akkor és csak akkor van triviálistól különböző megoldása, ha \mathbf{A} oszlopvektorai lineárisan összefüggőek.

Ezek alapján bebizonyítjuk a következő fontos tételt.

1.3.4 tétel. Mátrixok minimális diadikus előállításában szereplő diádok \mathbf{u}_k oszlopvektorai, illetve \mathbf{v}_k^T sorvektorai lineárisan független vektorrendszert alkotnak.

Bizonyítás. A tételt indirekt módszerrel bizonyítjuk. Feltesszük pl., hogy a minimális diadikus előállítás oszlopvektorai lineárisan összefüggőek, tehát hogy található olyan x_1, x_2, \dots, x_r értékrendszer, amelynek elemei közül legalább egy zérustól különböző és amellyel

$$x_1 \mathbf{u}_1 + x_2 \mathbf{u}_2 + \dots + x_r \mathbf{u}_r = \mathbf{0}.$$

Az általánosság megszorítása nélkül feltehető, hogy $x_r \neq 0$. A felírt összefüggésből tehát kifejezhető az \mathbf{u}_r vektor:

$$(1.3.13) \quad \mathbf{u}_r = -\frac{x_1}{x_r}\mathbf{u}_1 - \frac{x_2}{x_r}\mathbf{u}_2 - \cdots - \frac{x_{r-1}}{x_r}\mathbf{u}_{r-1}.$$

Behelyettesítünk az \mathbf{A} mátrix (1.3.4) alatti diádösszegre bontott alakjába:

$$\mathbf{A} = \mathbf{u}_1 \mathbf{v}_1^\top + \cdots + \mathbf{u}_{r-1} \mathbf{v}_{r-1}^\top + \left(-\frac{x_1}{x_r}\mathbf{u}_1 - \cdots - \frac{x_{r-1}}{x_r}\mathbf{u}_{r-1} \right) \mathbf{v}_r^\top;$$

összevonás után az alábbi adódik:

$$(1.3.14) \quad \mathbf{A} = \mathbf{u}_1 \left(\mathbf{v}_1^\top - \frac{x_1}{x_r}\mathbf{v}_r^\top \right) + \cdots + \mathbf{u}_{r-1} \left(\mathbf{v}_{r-1}^\top - \frac{x_{r-1}}{x_r}\mathbf{v}_r^\top \right).$$

Ezzel az \mathbf{A} mátrixot $r-1$ diád összegeként állítottuk elő, ami azonban ellentmond annak, hogy az (1.3.4) alatti felbontás minimális. Az $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ vektorok tehát nem lehetnek lineárisan összefüggőek. Tehát a minimális diadikus felbontás oszlopai lineárisan független vektorok. Hasonló módon belátható ugyanez a diádok sorvektoraira is. ■

Ennek a tételnek a következményeképpen megmutatjuk, hogy az

$$(1.3.15) \quad \mathbf{A}_{j_1 j_2 \dots j_{r+1}}^{1 2 \dots n} \tilde{\mathbf{x}} = \mathbf{U} \mathbf{V}_{j_1 j_2 \dots j_{r+1}}^{\top 1 2 \dots r} \tilde{\mathbf{x}} = \mathbf{0}$$

egyenletnek biztosan van triviálistól különböző megoldása. (Itt most $\tilde{\mathbf{x}}$ egy $r+1$ elemű vektor.) Az 1.3.4 tétel alapján \mathbf{U} oszlopai lineárisan független vektorok, tehát (1.3.15) csak úgy állhat fenn, ha

$$(1.3.16) \quad \mathbf{V}_{j_1 j_2 \dots j_{r+1}}^{\top 1 2 \dots r} \tilde{\mathbf{x}} = \mathbf{0}.$$

Megfelelő particionálással $\tilde{\mathbf{x}} = \begin{bmatrix} \mathbf{x} \\ x_{r+1} \end{bmatrix}$, ahol \mathbf{x} ismét r elemű vektor, így

$$(1.3.16) \quad \mathbf{V}^\top \text{ megfelelő oszlopai segítségével}$$

$$(1.3.17) \quad \mathbf{V}_{j_1 j_2 \dots j_r}^{\top 1 2 \dots r} \mathbf{x} + \mathbf{V}_{j_{r+1}}^{\top 1 2 \dots r} x_{r+1} = \mathbf{0}$$

alakban írható fel. Most két esetet különböztethetünk meg:

(a) ha $|\mathbf{V}_{j_1 j_2 \dots j_r}^{\top 1 2 \dots r}| = 0$, akkor $x_{r+1} = 0$ választással (1.3.17)-ből

$$\mathbf{V}_{j_1 j_2 \dots j_r}^{\top 1 2 \dots r} \mathbf{x} = \mathbf{0}$$

adódik és mivel $\mathbf{V}_{j_1 j_2 \dots j_r}^{\top 1 2 \dots r}$ szinguláris, ennek az egyenletnek van triviálistól különböző megoldása, tehát $\mathbf{V}_{j_1 j_2 \dots j_r}^{\top 1 2 \dots r}$ oszlopai és így $\mathbf{V}_{j_1 j_2 \dots j_{r+1}}^{\top 1 2 \dots r}$ oszlopai is lineárisan összefüggőek;

(b) ha $|\mathbf{V}_{j_1 j_2 \dots j_r}^{\mathbf{T}_{12\dots r}}| \neq 0$, azaz $\mathbf{V}_{j_1 j_2 \dots j_r}^{\mathbf{T}_{12\dots r}}$ nonszinguláris, akkor az inverzével balról megszorozva az (1.3.17) egyenletet és a második tagot a jobb oldalra átvive

$$\mathbf{x} = -(\mathbf{V}_{j_1 j_2 \dots j_r}^{\mathbf{T}_{12\dots r}})^{-1} \mathbf{V}_{j_{r+1}}^{\mathbf{T}_{12\dots r}} x_{r+1}.$$

Ez azt jelenti, hogy x_{r+1} értékét tetszőlegesen megválasztva, az (1.3.15) egyenletnek van triviálistól különböző megoldása, vagyis az \mathbf{A} mátrix bármely $r + 1$ oszlopa lineárisan összefüggő.

1.3.3 Rangra vonatkozó tételek

1.3.5 tétel. *Mátrixok összegének a rangja nem lehet nagyobb, mint az egyes mátrixok rangjának összege:*

$$(1.3.18) \quad \varrho(\mathbf{A} + \mathbf{B}) \leq \varrho(\mathbf{A}) + \varrho(\mathbf{B}).$$

Bizonyítás. Legyen \mathbf{A} , ill. \mathbf{B} egy-egy minimális diadikus előállítása

$$\mathbf{A} = \sum_{k=1}^r \mathbf{u}_k \mathbf{v}_k^{\mathbf{T}}, \quad \text{ill.} \quad \mathbf{B} = \sum_{l=1}^s \mathbf{w}_l \mathbf{z}_l^{\mathbf{T}};$$

ekkor $\mathbf{A} + \mathbf{B}$ egy diadikus felbontása

$$\mathbf{A} + \mathbf{B} = \sum_{k=1}^r \mathbf{u}_k \mathbf{v}_k^{\mathbf{T}} + \sum_{l=1}^s \mathbf{w}_l \mathbf{z}_l^{\mathbf{T}},$$

amely $r + s$ diádból áll. E felbontás azonban nem szükségképpen minimális, így az (1.3.18) egyenlőtlenség bizonyítva van. ■

1.3.6 tétel. *Mátrix rangját a szorzás nem növelheti:*

$$(1.3.19) \quad \varrho(\mathbf{AB}) \leq \varrho(\mathbf{A}) \quad \text{és} \quad \varrho(\mathbf{AB}) \leq \varrho(\mathbf{B}).$$

Bizonyítás. Legyen \mathbf{A} egy minimális diadikus előállítása

$$\mathbf{A} = \sum_{k=1}^{r_{\mathbf{A}}} \mathbf{u}_k \mathbf{v}_k^{\mathbf{T}} \quad \text{ahol} \quad \varrho(\mathbf{A}) = r_{\mathbf{A}};$$

ekkor \mathbf{AB} egy diadikus előállítása

$$\mathbf{AB} = \sum_{k=1}^{r_{\mathbf{A}}} \mathbf{u}_k (\mathbf{v}_k^{\mathbf{T}} \mathbf{B}) = \sum_{k=1}^{r_{\mathbf{A}}} \mathbf{u}_k \mathbf{w}_k^{\mathbf{T}},$$

ami $r_{\mathbf{A}}$ diádból áll. Minthogy e felbontás nem szükségképpen minimális, így az (1.3.19) első egyenlőtlensége igazolva van.

Teljesen hasonlóan látható be az (1.3.19) második egyenlőtlenségének érvényessége. ■

1.3.7 tétel. *Nemszinguláris mátrixszal végzett szorzás a mátrix rangját változtatlanul hagyja:*

$$(1.3.20) \quad \varrho(\mathbf{AM}) = \varrho(\mathbf{MA}) = \varrho(\mathbf{A}), \quad \text{ha} \quad |\mathbf{M}| \neq 0.$$

Bizonyítás. Alkalmazzuk az 1.3.6 tételt az \mathbf{AM} , majd az $(\mathbf{AM})\mathbf{M}^{-1}$ szorzatra:

$$\varrho(\mathbf{AM}) \leq \varrho(\mathbf{A}), \quad \text{és} \quad \varrho(\mathbf{A}) = \varrho[(\mathbf{AM})\mathbf{M}^{-1}] \leq \varrho(\mathbf{AM}).$$

(Az utóbbinál felhasználtuk, hogy $(\mathbf{AM})\mathbf{M}^{-1} = \mathbf{A}(\mathbf{MM}^{-1}) = \mathbf{A}$.) A kapott két egyenlőtlenség egyidejűleg csak az egyenlőség esetén állhat fenn. Teljesen hasonlóan belátható a tétel érvényessége a nemszinguláris \mathbf{M} mátrixszal balról végzett szorzásra is. ■

Az alábbi tételekben következtetéseket vonunk le abból, ha két vagy három mátrix szorzata zérus.

1.3.8 tétel. *Ha az $m \times r$ típusú \mathbf{U} mátrix oszlopai lineárisan függetlenek, akkor az*

$$(1.3.21) \quad \underset{(m,r)(r,n)}{\mathbf{U}} \cdot \mathbf{C} = \mathbf{0},$$

azaz

$$\left\{ \begin{array}{c} \overbrace{\hspace{1cm}}^r \\ \mathbf{U} \\ \underbrace{\hspace{1cm}}_m \end{array} \right\} \cdot \left\{ \begin{array}{c} \mathbf{C} \\ \underbrace{\hspace{1cm}}_n \\ \overbrace{\hspace{1cm}}^r \end{array} \right\} = \mathbf{0}$$

egyenletből

$$(1.3.22) \quad \underset{(r,n)}{\mathbf{C}} = \mathbf{0}$$

következik.

Bizonyítás. Az (1.3.21) egyenletet \mathbf{U} oszlopvektorainak felhasználásával az alábbi alakban írhatjuk:

$$(1.3.23) \quad \sum_{k=1}^r \mathbf{u}_k c_{kj} = \mathbf{0} \quad (j = 1, 2, \dots, n),$$

az \mathbf{u}_k vektorok lineáris függetlenségéből pedig már

$$(1.3.24) \quad c_{kj} = 0 \quad (k = 1, 2, \dots, r; j = 1, 2, \dots, n)$$

következik. ■

Hasonlóképpen belátható a következő tétel.

1.3.9 tétel. *Ha az $r \times n$ típusú \mathbf{V}^T mátrix sorai lineárisan függetlenek, akkor*

$$(1.3.25) \quad \underset{(m,r)}{\mathbf{B}} \underset{(r,n)}{\mathbf{V}^T} = \mathbf{0},$$

azaz

$$\begin{array}{c} \overbrace{\hspace{2cm}}^r \\ \left\{ \begin{array}{c} \mathbf{B} \\ (m, r) \end{array} \right\} \end{array} \begin{array}{c} \left\{ \begin{array}{c} \mathbf{V}^T \\ (r, n) \end{array} \right\} \\ \underbrace{\hspace{2cm}}_n \end{array} = \mathbf{0}$$

egyenletből

$$\underset{(m,r)}{\mathbf{B}} = \mathbf{0}$$

következik.

Az 1.3.8 és 1.3.9 tétel egyesítésének következménye a következő tétel.

1.3.10 tétel. *Ha az $m \times p$ típusú \mathbf{U} mátrix oszlopai és a $q \times n$ típusú \mathbf{V}^T mátrix sorai lineárisan függetlenek, akkor az*

$$\underset{(m,p)}{\mathbf{U}} \underset{(p,q)}{\mathbf{C}} \underset{(q,n)}{\mathbf{V}^T} = \mathbf{0}$$

egyenletből $\underset{(p,q)}{\mathbf{C}} = \mathbf{0}$ következik.

1.3.4 Elemi transzformációk, ekvivalens transzformációk, mátrix normálalakja

Az 1.3.7 tételben láttuk, hogy ha adott \mathbf{A} mátrixot megszorozunk valamely nemszinguláris \mathbf{M} mátrixszal, akkor ez a mátrix rangját változtatlanul hagyja.

Azt az utasítást, amely az \mathbf{A} mátrixhoz az \mathbf{MA} vagy az \mathbf{AM} mátrixot rendeli hozzá, az \mathbf{A} mátrix *transzformációjának* nevezzük.

Először tekintsük azt az elemi transzformációt, amelyik az adott mátrix *sorainak*, ill. *oszlopainak* az átrendezését hajtja végre. Ez az ún. *permutáló mátrixokkal* (lásd az 1.4.1 pontot) való szorzással érhető el. Mint a determinánselméletből ismeretes, ha egy kvadratikus mátrix sorait (ill. oszlopait) átrendezzük, akkor a mátrix determinánsa legfeljebb előjelet válthat (és ez akkor következik be, ha a sor-, ill. oszlopindexek permutációjában az inverziók száma páratlan).

A második elemi transzformáció a mátrix k -adik sorát (ill. oszlopát) *szorozza egy d_k számmal*. Ezt egy $\mathbf{D} = \langle 1 \ 1 \dots d_k \dots 1 \rangle$ diagonálmátrixszal balról (ill. jobbról) végrehajtott szorzással lehet elérni. Kvadratikus mátrix determinánsa ekkor nyilván a d_k számmal szorozódik.

A harmadik elemi transzformáció a mátrix i -edik sorának c -szeresét hozzáadja j -edik sorához. Könnyen belátható, hogy ezt a transzformációt a

$$\mathbf{Q}^{(ij)} = \mathbf{E} + c\mathbf{e}_j\mathbf{e}_i^T = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & \vdots & \ddots & & \\ & & & & c & \dots 1 \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}$$

mátrixszal balról végzett szorzás hajtja végre. Hasonlóképpen konstruálható meg az az

$$\mathbf{R}^{(ij)} = \mathbf{E} + c\mathbf{e}_i\mathbf{e}_j^T = \left(\mathbf{Q}^{(ij)}\right)^T$$

mátrix, amellyel jobbról szorozva az adott mátrixot, az i -edik oszlop c -szeresét adja hozzá a j -edik oszlophoz. A $\mathbf{Q}^{(ij)}$ és az $\mathbf{R}^{(ij)}$ mátrixok determinánsának értéke 1 – az ilyen mátrixokat *unimoduláris* mátrixoknak nevezzük – a velük való szorzás tehát az adott kvadratikus mátrix determinánsát nem változtatja meg.

Mindhárom típusú elemi transzformációnak nemszinguláris mátrixszal való szorzás felel meg, ezek tehát az 1.3.7 tétel értelmében az adott mátrix rangját nem változtatják meg.

A következőkben bevezetjük az *ekvivalens transzformáció* fogalmát, majd elemi transzformációk felhasználásával megmutatjuk, miként hozható bármely mátrix ekvivalens transzformációval „legegyszerűbb” (vagyis maximális számú zérus elemet tartalmazó) normálalakra.

1.3.6 definíció. *Ha egy mátrixot balról és jobbról egy-egy nemszinguláris mátrixszal szorzunk, akkor azt mondjuk, hogy a mátrixon **ekvivalens transzformációt** végeztünk.*

Az ekvivalens transzformáció e definíciójából és az 1.3.7 tételből következik, hogy a mátrix rangja ekvivalens transzformációval szemben invariáns.

Ezek után bebizonyítjuk a következő tételt.

1.3.11 tétel. *Bármely $m \times n$ típusú r -edrangú \mathbf{A} mátrixhoz található olyan m -edrendű nonszinguláris \mathbf{Q} és n -edrendű nonszinguláris \mathbf{R} mátrix, hogy a velük végzett ekvivalens transzformációval az \mathbf{A} mátrix*

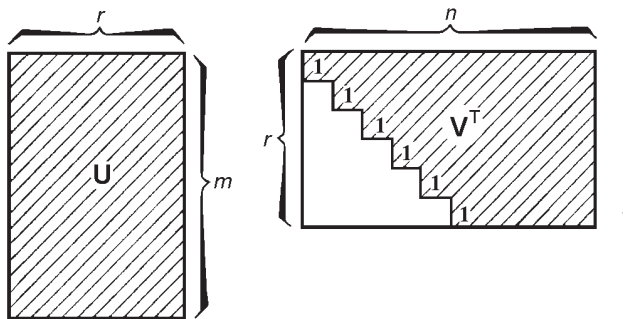
$$(1.3.26) \quad \mathbf{QAR} = \left[\begin{array}{cc} \mathbf{E}_r & 0 \\ 0 & \underbrace{0}_{n-r} \end{array} \right] \begin{array}{l} \underbrace{}_r \\ \underbrace{}_{m-r} \end{array} = \mathbf{P}_r$$

normálalakra hozható.

Bizonyítás. A tétel bizonyításához az \mathbf{A} mátrix egy minimális diadikus előállításából indulunk ki. A diádokat generáló elemet az \mathbf{A} mátrix első, második, \dots , r -edik oszlopából fogjuk választani. (Ezt azonban csak akkor tehetjük meg, ha \mathbf{A} első r oszlopa lineárisan független. Ha ez nem teljesül, akkor első lépésként egy megfelelő \mathbf{P} permutáló mátrixszal* jobbról való szorzással elérhető, hogy az átrendezett mátrix első r oszlopa már lineárisan független legyen.) Ezzel ugyanis biztosítható, hogy a minimális diadikus felbontással nyert faktorizáció második tényezője, a \mathbf{V}^T mátrix „derékszögű trapéz” alakot vegyen fel. Ha pedig a generáló elemmel az egyes diádok sorvektorát osztjuk, akkor ezzel elérjük azt, hogy \mathbf{V}^T minden v_{kk} eleme 1 legyen. Vagyis

$$\mathbf{AP} = \mathbf{UV}^T,$$

ahol \mathbf{U} és \mathbf{V}^T vázlatos szerkezete



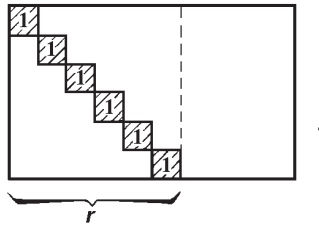
Ha most a \mathbf{V}^T mátrix első oszlopának alkalmas többszöröseit hozzáadjuk a második, harmadik, \dots , n -edik oszlopához, azután a második oszlop alkalmas

*A \mathbf{P} permutáló mátrix oszlopait úgy kapjuk meg, hogy az egységmátrix oszlopait permutáljuk; a vele való szorzás (jobbról) a mátrix oszlopait megfelelően permutálja (lásd részletesen az 1.4.7 definíciót).

többszöröseit a harmadik, \dots , n -edik oszlophoz, és így tovább, végül az r -edik oszlop alkalmas többszöröseit az $r+1$ -edik, $r+2$ -edik, \dots , n -edik oszlophoz – vagyis alkalmasan választott $\mathbf{R}^{(ij)}$ típusú mátrixokkal jobbról szorozva, elemi transzformációk egy sorozatát hajtjuk végre –, akkor ennek a mátrixnak a $v_{kk} = 1$ ($k = 1, 2, \dots, r$) elemein kívül valamennyi elemét zérussá tesszük. Jelölje \mathbf{T} ezeknek az $\mathbf{R}^{(ij)}$ mátrixoknak a szorzatát; akkor

$$\mathbf{V}^T \mathbf{T} = [\mathbf{E}_r \quad 0]$$

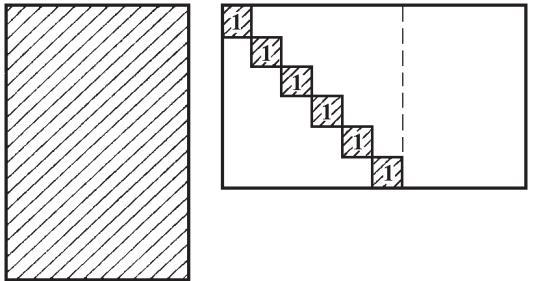
adódik, amelynek vázlatos szerkezete a következő:



Visszatérve az eredetileg adott \mathbf{A} mátrixra, az alábbi összefüggést nyerjük:

$$\mathbf{A} \mathbf{P} \mathbf{T} = \mathbf{U} \mathbf{V}^T \mathbf{T},$$

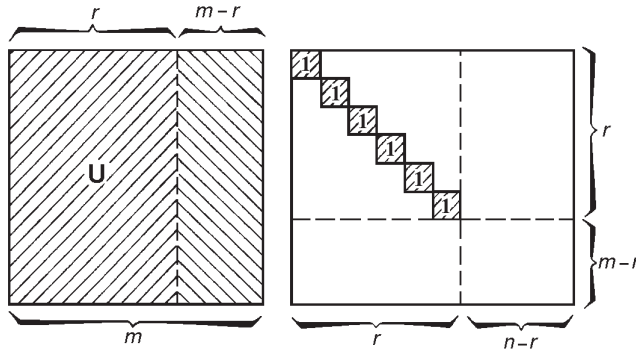
vázlatosan:



A kapott egyenlőség nyilván nem változik, ha a jobb oldalon az \mathbf{U} tényezőt $m - r$ oszloppal egy nemszinguláris $\tilde{\mathbf{U}}$ mátrixszá egészítjük ki, a $\mathbf{V}^T \mathbf{T}$ mátrixot pedig egyidejűleg $m - r$ csupa zérust tartalmazó sorral a \mathbf{P}_r mátrixszá egészítjük ki:

$$\mathbf{A} \mathbf{P} \mathbf{T} = \tilde{\mathbf{U}} \mathbf{P}_r;$$

az $\tilde{\mathbf{U}} \mathbf{P}_r$ mátrixszorzat szerkezete



Bevezetve a

$$\mathbf{Q} = \tilde{\mathbf{U}}^{-1}, \quad \mathbf{P}\mathbf{T} = \mathbf{R}$$

jelöléseket, valóban (1.3.26) adódik, amit bizonyítani kellett. ■

16. Példa. Határozzunk meg olyan \mathbf{Q} és \mathbf{R} nonsinguláris mátrixot, amelyek segítségével az

$$\mathbf{A} = \begin{bmatrix} \boxed{1} & 5 & 9 & 2 & -4 \\ 2 & 6 & 10 & 1 & -7 \\ 3 & 7 & 11 & 3 & -7 \\ 4 & 8 & 12 & 6 & -6 \end{bmatrix}$$

mátrix ekvivalens transzformációval az (1.3.26) normálalakra hozható. (A generáló elemet minden lépésben bekeretezzük.)

Megoldás. 1. Első lépésként meghatározzuk az \mathbf{A} mátrix egy minimális diagonális felbontását és faktorizált alakját.

$$\mathbf{A} - \frac{1}{1} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 5 & 9 & 2 & -4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \boxed{-4} & -8 & -3 & 1 \\ 0 & -8 & -16 & -3 & 5 \\ 0 & -12 & -24 & -2 & 10 \end{bmatrix} = {}^{(2)}\mathbf{A};$$

$${}^{(2)}\mathbf{A} - \left(-\frac{1}{4}\right) \begin{bmatrix} 0 \\ -4 \\ -8 \\ -12 \end{bmatrix} \begin{bmatrix} 0 & -4 & -8 & -3 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \boxed{3} & 3 \\ 0 & 0 & 0 & 7 & 7 \end{bmatrix} = {}^{(3)}\mathbf{A};$$

$${}^{(3)}\mathbf{A} - \frac{1}{3} \begin{bmatrix} 0 \\ 0 \\ 3 \\ 7 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 3 & 3 \end{bmatrix} = \mathbf{0}.$$

A diádok együtthatóival beszorozva a megfelelő sorvektorokat, az adott mátrix faktorizált alakja adódik:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & -4 & 0 \\ 3 & -8 & 3 \\ 4 & -12 & 7 \end{bmatrix} \begin{bmatrix} 1 & 5 & 9 & 2 & -4 \\ 0 & 1 & 2 & \frac{3}{4} & -\frac{1}{4} \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

Ebből már következik, hogy \mathbf{A} rangja 3.

2. A \mathbf{P} permutáló mátrix meghatározása. Tekintettel arra, hogy az \mathbf{A} mátrixnak nemcsak a második, hanem a harmadik oszlopa is csupa zérus elemet tartalmaz (ami annak a következménye, hogy az \mathbf{A} mátrix első három oszlopa lineárisan összefüggő), ezért most jobbról egy olyan \mathbf{P} permutáló mátrixszal kell szorozni, amely a faktorizált alak második tényezőjének – és ezzel egyúttal az \mathbf{A} mátrixnak – oszlopait úgy rendezi át, hogy \mathbf{AP} első három oszlopa lineárisan független legyen. Ezzel biztosítható ugyanis a faktorizáció jobb oldali tényezőjének trapéz alakja. Mivel \mathbf{A} negyedik oszlopa már nem csupa zérus elemet tartalmaz, ezért elegendő a harmadik és negyedik oszlopot felcserélni. Tehát jobbról a

$$\mathbf{P} = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_4 \quad \mathbf{e}_3 \quad \mathbf{e}_5]$$

permutáló mátrixszal kell szoroznunk:

$$\mathbf{AP} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & -4 & 0 \\ 3 & -8 & 3 \\ 4 & -12 & 7 \end{bmatrix} \begin{bmatrix} 1 & 5 & 2 & 9 & -4 \\ 0 & 1 & \frac{3}{4} & 2 & -\frac{1}{4} \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} = \mathbf{UV}^T.$$

3. A \mathbf{T} mátrix meghatározása. Következő lépésben a \mathbf{V}^T mátrix ii indexű elemein kívüli összes elemét megfelelő elemi transzformációk sorozatával – mégpedig $\mathbf{R}^{(ij)}$ típusú nonsinguláris mátrixokkal jobbról való szorzással – zérussá kell tenni. Nem nehéz belátni, hogy ha \mathbf{V}^T első oszlopának (-5) -szörösét a második oszlophoz, (-2) -szeresét a harmadik oszlophoz, (-9) -szeresét a negyedik oszlophoz és 4-szeresét az ötödik oszlophoz hozzáadjuk,

azaz jobbról szorzunk a

$$\mathbf{T}_1 = \begin{bmatrix} 1 & -5 & -2 & -9 & 4 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

mátrixszal, akkor \mathbf{V}^T első sorának elemei (természetesen az első kivételével) zérussá válnak. Ha az így kapott $\mathbf{V}^T \mathbf{T}_1$ mátrix második oszlopának $(-\frac{3}{4})$ -szeresét a harmadik oszlophoz, (-2) -szeresét a negyedik oszlophoz és $(\frac{1}{4})$ -szeresét az ötödik oszlophoz adjuk hozzá, azaz jobbról szorzunk a

$$\mathbf{T}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -\frac{3}{4} & -2 & \frac{1}{4} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

mátrixszal, akkor – anélkül, hogy az első sor zérus elemeit „elrontanánk” – a második sor elemei is zérussá válnak (természetesen ennek második eleme kivételével). Végül, ha a kapott $\mathbf{V}^T \mathbf{T}_1 \mathbf{T}_2$ mátrix harmadik oszlopát levonjuk az ötödik oszlopból, vagyis jobbról szorzunk a

$$\mathbf{T}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

mátrixszal, akkor az így adódó $\mathbf{V}^T \mathbf{T}_1 \mathbf{T}_2 \mathbf{T}_3$ mátrixnak az azonos sor- és oszlopindexű helyeken álló elemei egyenlők 1-gyel, az összes többi helyen pedig 0 áll. A \mathbf{V}^T mátrix elemeit „kipusztító” \mathbf{T} mátrix a \mathbf{T}_1 , \mathbf{T}_2 , és \mathbf{T}_3 unimoduláris mátrixok szorzataként adódik:

$$\mathbf{T} = \mathbf{T}_1 \mathbf{T}_2 \mathbf{T}_3 = \begin{bmatrix} 1 & -5 & \frac{7}{4} & 1 & 1 \\ 0 & 1 & \frac{3}{4} & -2 & 1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

4. Az **R** mátrix meghatározása. A keresett ekvivalens transzformáció **R** mátrixa végül az így kapott **T** mátrixnak és az előző lépésben meghatározott **P** permutáló mátrixnak **PT** szorzata (esetünkben ez annyit jelent, hogy a **T** mátrix harmadik és negyedik sorát fel kell cserélni):

$$\begin{aligned} \mathbf{R} = \mathbf{PT} &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -5 & \frac{7}{4} & 1 & 1 \\ 0 & 1 & -\frac{3}{4} & -2 & 1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \\ &= \begin{bmatrix} 1 & -5 & \frac{7}{4} & 1 & 1 \\ 0 & 1 & -\frac{3}{4} & -2 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

5. A **Q** mátrix meghatározása. Az ekvivalens transzformáció **Q** mátrixát úgy nyerjük, hogy a 2. lépésben kapott

$$\mathbf{U} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & -4 & 0 \\ 3 & -8 & 3 \\ 4 & -12 & 7 \end{bmatrix}$$

mátrixot egy negyedik oszloppal tetszőleges nonszinguláris $\tilde{\mathbf{U}}$ mátrixra egészítjük ki, és ennek vesszük az inverzét. Például

$$\tilde{\mathbf{U}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & -4 & 0 & 0 \\ 3 & -8 & 3 & 0 \\ 4 & -12 & 7 & 1 \end{bmatrix}, \text{ és innen } \mathbf{Q} = \tilde{\mathbf{U}}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & -\frac{1}{4} & 0 & 0 \\ \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} & 0 \\ -\frac{1}{3} & \frac{5}{3} & -\frac{7}{3} & 1 \end{bmatrix}.$$

6. Az ekvivalens transzformáció felírása. Az így nyert **R** és **Q** mátrix behelyettesítésével végül valóban a keresett normálalak, vagyis

$$\mathbf{QAR} = \begin{bmatrix} \mathbf{E}_3 & 0 \\ 0 & 0 \end{bmatrix}$$

adódik, ugyanis

$$\begin{aligned}
 & \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & -\frac{1}{4} & 0 & 0 \\ \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} & 0 \\ -\frac{1}{3} & \frac{5}{3} & -\frac{7}{3} & 1 \end{bmatrix} \begin{bmatrix} 1 & 5 & 9 & 2 & -4 \\ 2 & 6 & 10 & 1 & -7 \\ 3 & 7 & 11 & 3 & -7 \\ 4 & 8 & 12 & 6 & -6 \end{bmatrix} \begin{bmatrix} 1 & -5 & \frac{7}{4} & 1 & 1 \\ 0 & 1 & -\frac{3}{4} & -2 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \\
 & = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.
 \end{aligned}$$

* * *

1.3.5 A nullitás fogalma; a Sylvester-féle nullitási tétel

Az 1.3.11 tétel felhasználásával bebizonyítható a következő tétel, amely két mátrix szorzatának a rangjára fontos alsó becslést ad. Mielőtt erre rátérnénk, bevezetjük kvadratikus mátrixokra a *nullitás* vagy *defektus* fogalmát.

1.3.7 definíció. Egy kvadratikus mátrix rendszámának és rangjának a különbségét a mátrix **nullitásának** vagy **defektusának** nevezzük.

Ha a nullitást d jelöli, a mátrix rendszáma n és a rangja r , akkor tehát

$$(1.3.27) \quad d = n - r.$$

Az 1.3.6 tételből következik, hogy két mátrix szorzatának a rangja nem lehet nagyobb, mint bármelyik tényező rangja, ezzel tehát egy felső becslést kaptunk a szorzat rangjára vonatkozóan:

$$(1.3.28) \quad \varrho(\mathbf{AB}) \leq \min(r_{\mathbf{A}}, r_{\mathbf{B}}),$$

ahol

$$\varrho(\mathbf{A}) = r_{\mathbf{A}}, \quad \varrho(\mathbf{B}) = r_{\mathbf{B}}.$$

Ezzel analóg alsó becslést ad a következő tétel.

1.3.12 tétel (Sylvester-féle nullitási tétel).* Ha az \mathbf{A} , ill. \mathbf{B} mátrix rangja $r_{\mathbf{A}}$, ill. $r_{\mathbf{B}}$, továbbá nullitása $d_{\mathbf{A}}$, ill. $d_{\mathbf{B}}$, akkor az \mathbf{AB} szorzat rangjára érvényes a

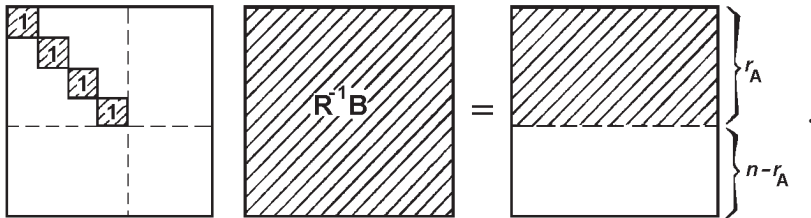
$$(1.3.29) \quad \varrho(\mathbf{AB}) \geq \min(r_{\mathbf{A}}, r_{\mathbf{B}}) - \min(d_{\mathbf{A}}, d_{\mathbf{B}})$$

becslés (lásd pl. [20]).

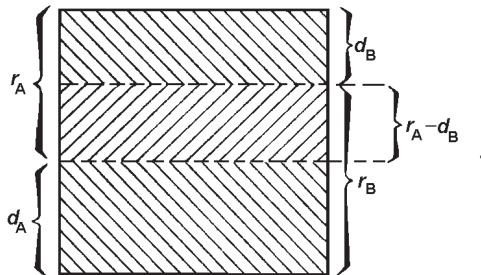
Bizonyítás. Tegyük fel, hogy $r_{\mathbf{A}} \leq r_{\mathbf{B}}$ (ha ez nem teljesül, akkor a $\mathbf{B}^T \mathbf{A}^T$ szorzat rangját vizsgáljuk). Hozzuk az \mathbf{A} mátrixot az (1.3.26) szerinti normálalakra, és írjuk fel az \mathbf{AB} szorzat rangját:

$$\varrho(\mathbf{AB}) = \varrho(\mathbf{QAB}) = \varrho((\mathbf{QAR})(\mathbf{R}^{-1}\mathbf{B})) = \varrho(\mathbf{P}_{r_{\mathbf{A}}}\mathbf{R}^{-1}\mathbf{B}).$$

A nemszinguláris \mathbf{Q} , ill. \mathbf{R} és \mathbf{R}^{-1} mátrixszal való szorzás a rangot nem változtatja meg, a normálalakú $\mathbf{P}_{r_{\mathbf{A}}}$ mátrixszal való szorzás pedig az $r_{\mathbf{B}}$ rangú $\mathbf{R}^{-1}\mathbf{B}$ mátrixnak csak az első $r_{\mathbf{A}}$ sorát tartja meg. Vázlatosan:



Ha feltesszük, hogy $\mathbf{R}^{-1}\mathbf{B}$ utolsó $n - r_{\mathbf{A}}$ sora lineárisan független, akkor a megmaradó $r_{\mathbf{A}}$ sor között $r_{\mathbf{B}} - (n - r_{\mathbf{A}}) = (r_{\mathbf{B}} - n) + r_{\mathbf{A}} = r_{\mathbf{A}} - d_{\mathbf{B}}$ lineárisan független sornak kell lennie (feltéve, hogy $r_{\mathbf{B}} > d_{\mathbf{A}}$):



Ebből következik, hogy az \mathbf{AB} szorzat rangja $r_{\mathbf{A}} \leq r_{\mathbf{B}}$ esetén nem kisebb, mint $r_{\mathbf{A}} - d_{\mathbf{B}}$. Ezzel a tételt bebizonyítottuk. ■

* J. J. Sylvester (1814–1897) angol matematikus.

Egyesítve a szorzatmátrix rangjára vonatkozó (1.3.29) alsó és (1.3.28) felső becslést, a következő egyenlőtlenséget kapjuk:

$$(1.3.30) \quad \min(r_{\mathbf{A}}, r_{\mathbf{B}}) - \min(d_{\mathbf{A}}, d_{\mathbf{B}}) \leq \varrho(\mathbf{AB}) \leq \min(r_{\mathbf{A}}, r_{\mathbf{B}}).$$

Speciális esetként adódik annak szükséges feltétele, hogy két kvadrati-kus mátrix szorzata zérus legyen: mivel a szorzat rangja ekkor zérus, ezért (1.3.29) alapján

$$(1.3.31) \quad r_{\mathbf{A}} + r_{\mathbf{B}} - n \leq 0,$$

azaz $r_{\mathbf{A}} + r_{\mathbf{B}} \leq n$.

Ezzel bebizonyítottuk a következő tételt:

1.3.13 tétel. *Két kvadrati-kus mátrix szorzata csak akkor lehet zérus, ha a rangjuk összege nem nagyobb, mint a rendszámuk.*

Megjegyzés. A tétel fordítva nem igaz, vagyis abból, hogy két mátrix rangjának az összege kisebb, mint rendszámuk, még nem következik, hogy a szorzatuk zérus!

1.4 SPECIÁLIS TULAJDONSÁGÚ MÁTRIXOK

Ebben a szakaszban néhány olyan mátrixtípussal ismerkedünk meg, amelyeket különleges tulajdonságaik tüntetnek ki a kvadrati-kus mátrixok körében. Először olyan *speciális szerkezetű mátrixokkal* foglalkozunk, amelyekben szabályosan elhelyezkedő, sok zérus elem található. Az irodalomban ezeket *strukturált mátrixoknak* is szokták nevezni. Megjegyezzük, hogy ilyen „sok” zérus elemet tartalmazó ún. *sparse* vagy *ritka* mátrixoknak a numerikus kezelés szempontjából is számos előnyük van. Mi elsősorban azokra az előnyökre kívánunk rámutatni, amelyek elvileg könnyítik meg a velük való műveletek elvégzését.

Ezek sorában részletesen foglalkozunk ilyen speciális tulajdonságú strukturált mátrixok invertálásával, az inverz szerkezetének feltárásával és azokkal az „ad hoc” módszerekkel, amelyek lehetővé teszik az inverz elemeinek a rendszám függvényében való előállítását. Ezek segítségével egyrészt előkészítjük a későbbiek során előforduló bonyolultabb feladatok megoldását, másrészt pedig lehetővé kívánjuk tenni az Olvasó számára annak a rutin-nak a megszerzését, amely képessé teszi majd az itt megismert módszerek alkalmazására.

1.4.1 Speciális mátrixok

1.4.1 definíció. *Az olyan $[a_{ij}]$ mátrixot, amelyre*

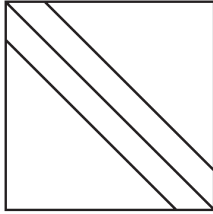
$$a_{ij} = \begin{cases} \neq 0, & \text{ha } |j - i| \leq 1 \\ = 0, & \text{ha } |j - i| > 1, \end{cases}$$

kontinuáns (tridiagonális) mátrixnak, másképpen **Jacobi-féle*** mátrixnak *nevezzük* (lásd [1], [6], [13]). Általános alakja:

$$(1.4.1) \quad \begin{bmatrix} a_1 & b_1 & 0 & . & . & . & 0 \\ c_1 & a_2 & b_2 & 0 & . & . & 0 \\ 0 & c_2 & a_3 & b_3 & 0 & . & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & c_{n-1} & a_n \end{bmatrix},$$

ahol a főátlóval „szomszédos”; b_i , c_i elemek a mátrix ún. *kodiagonálisában* helyezkednek el.

Ha csupán az ilyen szerkezetre kívánunk utalni, akkor a következő vázlatos jelölést használjuk:



Ha egy kontinuáns mátrix szimmetrikus is, azaz $b_i = c_i$ ($i = 1, 2, \dots, n-1$), akkor *szimmetrikus kontinuáns mátrixnak* nevezzük.

Megjegyzés. Egy későbbi fejezetben (4.4.3 pont) azt is megmutatjuk, hogy az általános nonszimmetrikus kontinuáns mátrixok mindig „szimmetrizálhatók”, ami azt jelenti, hogy elegendő, ha a szimmetrikus eset vizsgálatára szorítkozunk.

Ha a kontinuáns mátrixban $a_1 = a_2 = \dots = a_n = a$, $b_1 = b_2 = \dots = b_{n-1} = b$ és $c_1 = c_2 = \dots = c_{n-1} = c$, akkor *egyenletes kontinuáns mátrixról*, ha még $b = c$, akkor *szimmetrikus egyenletes kontinuáns mátrixról* van szó. Az alkalmazásokban ezeknek fontos szerepük van, ezért érdemes velük részletesebben foglalkozni.

Tekintsük tehát az alábbi, n -edrendű szimmetrikus egyenletes kontinuáns mátrixot:

$$(1.4.2) \quad \begin{bmatrix} a & b & 0 & . & . \\ b & a & b & . & . \\ 0 & b & a & b & . \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & . & b & a \end{bmatrix} \begin{matrix} 1) \\ 2) \\ . \\ \vdots \\ n) \end{matrix}.$$

*C. G. Jacobi (1804–1851) német matematikus.

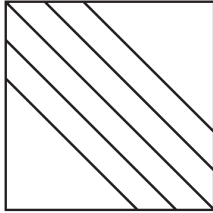
Nem nehéz belátni, hogy minden szimmetrikus egyenletes kontinuáns mátrix felírható a

$$(1.4.3) \quad \mathbf{K} = \begin{bmatrix} 0 & 1 & 0 & . & . \\ 1 & 0 & 0 & . & . \\ 0 & 1 & 0 & 1 & . \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}$$

szimmetrikus egyenletes kontinuáns mátrix polinomjaként, $a\mathbf{E} + b\mathbf{K}$ alakban.

1.4.2 definíció. Ha egy mátrixban nem csupán a főátlóval szomszédos első elemek, hanem még a második, harmadik stb. elemek is különbözhetnek zérustól, de a főátlótól bizonyos távolságra már valamennyi elem zérus, akkor azt **sávmátrixnak** (vagy szalagmátrixnak) nevezzük.

Ilyenkor azt is meg szoktuk jelölni, hogy hány „ferde” sorban található zérustól különböző elemek. Például egy ötsoros sávmátrix vázlatos jelölése a következő:



Megjegyezzük, hogy ezt *pentadiagonális mátrixnak* is szokták nevezni.

1.4.3 definíció. Ha egy mátrixban minden „ferde” sor csupa megegyező elem-ből áll, azaz a mátrix elemei csupán oszlop- és sorindexük különbségétől függenek, akkor azt **Toeplitz-típusú*** mátrixnak nevezzük.

Tehát az \mathbf{A} mátrix Toeplitz-típusú, ha $a_{ij} = a_{j-i}$.

Részletesen kiírva:

$$\mathbf{A} = \begin{bmatrix} a_0 & a_1 & a_2 & \dots & a_{n-1} \\ a_{-1} & a_0 & a_1 & \dots & a_{n-2} \\ a_{-2} & a_{-1} & a_0 & \dots & a_{n-3} \\ \dots & \dots & \dots & \dots & \dots \\ a_{-n+1} & a_{-n+2} & \dots & a_0 \end{bmatrix}.$$

Amennyiben egy Toeplitz-típusú mátrix még szimmetrikus is, akkor elemeire az $a_{ij} = a_{|j-i|}$ összefüggés érvényes.

*O. Toeplitz (1881–1940) német matematikus.

Ezek szerint tehát az egyenletes kontinuáns mátrix speciális Toeplitz-típusú mátrix. Egy másik nevezetes speciális Toeplitz-típusú mátrix a következő:

$$(1.4.4) \quad \mathbf{H} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \dots & 1 \\ 0 & 0 & 0 & \dots & \dots & 0 \end{bmatrix} \begin{matrix} (1) \\ (2) \\ \vdots \\ (n) \end{matrix}$$

Ismételt szorzással közvetlenül belátható, hogy

$$\mathbf{H}^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \dots & 1 \\ 0 & 0 & 0 & \dots & \dots & 0 \\ 0 & 0 & 0 & \dots & \dots & 0 \end{bmatrix}, \dots, \mathbf{H}^{n-1} = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix},$$

és végül $\mathbf{H}^n = \mathbf{0}$.

1.4.4 definíció. Azokat a mátrixokat, amelyekre létezik olyan p természetes szám, hogy

$$(1.4.5) \quad \mathbf{A}^{p-1} \neq \mathbf{0}, \quad \text{de} \quad \mathbf{A}^p = \mathbf{0},$$

p **indexű nilpotens mátrixoknak** nevezzük.

Az itt bemutatott \mathbf{H} mátrix a „legegyszerűbb szerkezetű”, vagyis a legtöbb zérus elemet tartalmazó, n indexű nilpotens mátrix.

Tekintsük most a következő Toeplitz-típusú mátrixot:

$$(1.4.6) \quad \mathbf{A}(a_0, a_1, \dots, a_{n-1}) = \begin{bmatrix} a_0 & a_1 & a_2 & \dots & a_{n-1} \\ 0 & a_0 & a_1 & \dots & a_{n-2} \\ 0 & 0 & a_0 & \dots & a_{n-3} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_0 \end{bmatrix}.$$

Az (1.4.4) alatti \mathbf{H} nilpotens mátrix hatványainak szerkezetéből közvetlenül adódik, hogy az (1.4.6) mátrix a \mathbf{H} mátrix polinomjaként írható fel:

$$(1.4.7) \quad \mathbf{A}(a_0, a_1, \dots, a_{n-1}) = a_0 \mathbf{E} + a_1 \mathbf{H} + a_2 \mathbf{H}^2 + \dots + a_{n-1} \mathbf{H}^{n-1}.$$

A nilpotens mátrixok indexének fogalma kiterjeszthető tetszőleges kvadratus mátrixokra. Ha meggondoljuk, hogy p indexű nilpotens mátrixok esetén p az a legkisebb pozitív egész, amelyre \mathbf{A}^p rangja 0 és ennél magasabb

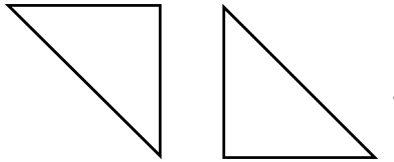
hatványra emelve a rangja már nem változik, akkor az indexnek ezt a tulajdonságát tetszőleges mátrixra általánosítva, megadható a következő definíció.

1.4.5 definíció. Az \mathbf{A} mátrix **indexe** az a legkisebb pozitív egész p , amelynél magasabb hatványra emelve a mátrixot, a hatvány rangja nem változik, azaz

$$\varrho(\mathbf{A}^{p+1}) = \varrho(\mathbf{A}^p).$$

1.4.6 definíció. Az olyan mátrixokat, amelyeknek a főátló **alatti** elemei mind zérussal egyenlők, **felső háromszögmátrixoknak**, azokat pedig, amelyekben a főátló **feletti** elemek egyenlők zérussal, **alsó háromszögmátrixoknak** (trianguláris mátrixoknak) nevezzük.

Ezeket szematikusan az alábbi ábrákkal vázoljuk:



(Ilyen mátrixokkal már találkoztunk, lásd az (1.3.6) faktorizációt.)

1.4.7 definíció. Ha egy n -edrendű kvadrátikus mátrix minden sorában és oszlopában pontosan egy 1-es áll, a többi elem pedig 0, akkor ez a mátrix fel fogható úgy, hogy az egységmátrixból oszlopainak (vagy sorainak) valamilyen permutációjával adódott; ezért az ilyen mátrixot **permutáló mátrixnak** nevezzük. Ez felírható az alábbi alakban:

$$\mathbf{P} = [\mathbf{e}_{j_1} \quad \mathbf{e}_{j_2} \quad \dots \quad \mathbf{e}_{j_n}],$$

ill.

$$(1.4.8) \quad \mathbf{P}^T = \begin{bmatrix} \mathbf{e}_{j_1}^T \\ \mathbf{e}_{j_2}^T \\ \vdots \\ \mathbf{e}_{j_n}^T \end{bmatrix},$$

ahol a $j_1 j_2 \dots j_n$ szám- n -es az $1, 2, \dots, n$ számok valamilyen permutációját jelenti.

Bármely mátrixot a fenti \mathbf{P} permutáló mátrixszal jobbról szorozva, annak oszlopait a szorzás éppen a $j_1 j_2 \dots j_n$ sorrendbe rendezi át:

$$(1.4.9) \quad \mathbf{AP} = [\mathbf{a}_1 \quad \mathbf{a}_2 \dots \mathbf{a}_n] [\mathbf{e}_{j_1} \quad \mathbf{e}_{j_2} \dots \mathbf{e}_{j_n}] = [\mathbf{a}_{j_1} \quad \mathbf{a}_{j_2} \dots \mathbf{a}_{j_n}].$$

Ilyen permutáló mátrixszal már találkoztunk az 1.3.11 tétel bizonyítása során.

Hasonlóképpen, bármely mátrixot \mathbf{P}^\top -vel *balról szorozva*, annak *sorait* a szorzás a $j_1 j_2 \dots j_n$ sorrendbe rendezi át:

$$(1.4.10) \quad \mathbf{P}^T \mathbf{A} = \begin{bmatrix} \mathbf{e}_{j_1}^T \\ \mathbf{e}_{j_2}^T \\ \vdots \\ \mathbf{e}_{j_n}^T \end{bmatrix} \begin{bmatrix} \mathbf{a}^1 \\ \mathbf{a}^2 \\ \vdots \\ \mathbf{a}^n \end{bmatrix} = \begin{bmatrix} \mathbf{a}^{j_1} \\ \mathbf{a}^{j_2} \\ \vdots \\ \mathbf{a}^{j_n} \end{bmatrix}.$$

Speciális permutáló mátrixot nyerünk, ha az egység mátrix oszlopait (sorait) *ciklikusan* permutáljuk: az így adódó mátrixokat *ciklikus permutáló mátrixoknak* nevezzük.

1.4.8 definíció. Az $\Omega_k = [\mathbf{e}_{n-k+1} \ \mathbf{e}_{n-k+2} \ \dots \ \mathbf{e}_n \ \mathbf{e}_1 \ \dots \ \mathbf{e}_{n-k}]$ mátrixot (k -adik) ciklikus permutáló mátrixnak nevezzük.

Az „első” ciklikus permutáló mátrix, vagyis

$$(1.4.11) \quad \Omega = \Omega_1 = [\mathbf{e}_n \quad \mathbf{e}_1 \quad \mathbf{e}_2 \quad \dots \quad \mathbf{e}_{n-1}] = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix}$$

az *elemi ciklikus mátrix*. Közvetlenül belátható, hogy a többi ciklikus permutáló mátrix az elemi ciklikus mátrix hatványozásával adódik:

$$\begin{aligned}
 \Omega_2 &= \Omega^2 = [\mathbf{e}_{n-1} & \mathbf{e}_n & \mathbf{e}_1 \dots \mathbf{e}_{n-2}]; \\
 \Omega_3 &= \Omega^3 = [\mathbf{e}_{n-2} & \mathbf{e}_{n-1} & \mathbf{e}_n \dots \mathbf{e}_{n-3}]; \\
 &\dots\dots\dots \\
 \Omega_n &= \Omega^n = [\mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \dots \mathbf{e}_n] = \mathbf{E}.
 \end{aligned}
 \tag{1.4.12}$$

Az is látható, hogy a ciklikus mátrixok valamennyien Toeplitz-típusú mátrixok.

1.4.9 definíció. Az első sor elemei által egyértelműen meghatározott

$$(1.4.13) \quad \mathbf{C}(c_0, c_1, \dots, c_{n-1}) = \begin{bmatrix} c_0 & c_1 & c_2 \dots c_{n-1} \\ c_{n-1} & c_0 & c_1 \dots c_{n-2} \\ c_{n-2} & c_{n-1} & c_0 \dots c_{n-3} \\ \dots & \dots & \dots \\ c_1 & c_2 & c_3 \dots c_0 \end{bmatrix}$$

mátrixot ciklikus mátrixnak nevezzük.

Az (1.4.11) alatti elemi ciklikus mátrix hatványainak (1.4.12) szerkezetéből közvetlenül adódik, hogy az (1.4.13) mátrix felírható az elemi ciklikus mátrix polinomjaként:

$$(1.4.14) \quad \mathbf{C}(c_0, c_1, \dots, c_{n-1}) = c_0 \mathbf{E} + c_1 \mathbf{\Omega} + c_2 \mathbf{\Omega}^2 + \dots + c_{n-1} \mathbf{\Omega}^{n-1}.$$

Végül meg kell jegyezni, hogy egy mátrix hatványainak – és polinomjainak – kommutativitásából következik az alábbi nevezetes tétel.

1.4.1 tétel. *Bármely két – azonos rendszámú – ciklikus mátrix a szorzásra kommutatív.*

Az inverz fogalma lehetővé teszi olyan speciális mátrixok bevezetését, amelyekre nem a „szerkezet” (struktúra) jellemző, hanem más tulajdonságaik, amelyeknek igen nagy szerepük lesz az elmélet további kiépítése során.

1.4.10 definíció. *Azokat a (komplex elemű) mátrixokat, amelyekre*

$$(1.4.15) \quad \mathbf{A}^{-1} = \mathbf{A}^H, \quad \text{vagyis} \quad \mathbf{A}\mathbf{A}^H = \mathbf{E}$$

fennáll, unitér mátrixoknak nevezzük. Valós elemek esetén az (1.4.15)-nek megfelelő összefüggés

$$(1.4.16) \quad \mathbf{A}^{-1} = \mathbf{A}^T, \quad \text{ill.} \quad \mathbf{A}\mathbf{A}^T = \mathbf{E},$$

és az ezt kielégítő mátrixokat valós ortogonális mátrixoknak nevezzük.

Az unitér, ill. ortogonális mátrixok 1.4.10 definíciója alapján könnyen bebizonyítható a következő tétel.

1.4.2 tétel. *Unitér mátrixok szorzata is unitér mátrix.*

Bizonyítás. Legyen \mathbf{A} és \mathbf{B} egy-egy n -edrendű unitér mátrix; képezzük az \mathbf{AB} szorzat inverzét, és helyettesítsük be az (1.4.15) definiáló összefüggéseket:

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} = \mathbf{B}^H\mathbf{A}^H = (\mathbf{AB})^H.$$

Azt kaptuk tehát, hogy az \mathbf{AB} szorzat is kielégíti az (1.4.15) összefüggést, vagyis \mathbf{AB} is unitér mátrix. ■

A tételből speciális esetként adódik, hogy ortogonális mátrixok szorzata is ortogonális mátrix.

1.4.11 definíció. *Azokat a mátrixokat, amelyekre*

$$(1.4.17) \quad \mathbf{A}^{-1} = \mathbf{A}, \quad \text{vagyis} \quad \mathbf{A}^2 = \mathbf{E},$$

involutórius mátrixoknak, amelyekre pedig

$$(1.4.18) \quad \mathbf{A}^{-1} = -\mathbf{A}, \quad \text{vagyis} \quad \mathbf{A}^2 = -\mathbf{E},$$

félig involutórius mátrixoknak nevezzük.

17. Példa. Közvetlen számolással meggyőződhetünk arról, hogy az

$$\mathbf{A} = \begin{bmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

mátrix ortogonális, mert $\mathbf{A}\mathbf{A}^T = \mathbf{E}$, továbbá, hogy

$$\mathbf{B} = \begin{bmatrix} 11 & 8 & 20 \\ 30 & 19 & 50 \\ -18 & -12 & -31 \end{bmatrix}$$

involutórius mátrix, mert $\mathbf{B}^2 = \mathbf{E}$ és végül, hogy

$$\mathbf{C} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{3}} + \frac{1}{2}i & -\frac{1}{2\sqrt{3}} - \frac{1}{2}i \\ \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{3}} - \frac{1}{2}i & -\frac{1}{2\sqrt{3}} + \frac{1}{2}i \end{bmatrix}$$

unitér mátrix, mert $\mathbf{C}\mathbf{C}^H = \mathbf{E}$.

* * *

1.4.12 definíció. Az

$$(1.4.19) \quad \mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$$

összefüggést kielégítő mátrixokat **normális mátrixoknak** vagy röviden, **normálmátrixoknak** nevezzük.

A normálmátrixokat tehát az jellemzi, hogy felcserélhetők a saját *transzponált konjugáltjukkal*.

A normálmátrixok a mátrixok nagy és fontos osztályát alkotják. Például az 1.4.10 definícióban megismert unitér és ortogonális mátrixok, valamint a már korábban megismert hermitikus (szimmetrikus), valamint a ferdén hermitikus (ferdén szimmetrikus) mátrixok valamennyien a normálmátrixok speciális esetei.

Az 1.1 szakaszban megismert hermitikus, ferdén hermitikus, valós szimmetrikus és ferdén szimmetrikus mátrixok definíciójából, valamint az unitér (valós ortogonális), involutórius és félig involutórius mátrixok 1.4.10 és 1.4.11 definíciójából közvetlenül következik, hogy

(a) ha egy mátrix hermitikus és unitér (ill. valós szimmetrikus és ortogonális), akkor egyúttal involutórius is,

(b) ha egy mátrix ferdén hermitikus és unitér (ill. valós ferdén szimmetrikus és ortogonális), akkor együttal félig involutórius is.

Már korábban hangsúlyoztuk, hogy a mátrixok körében különös jelentősége van annak, ha két mátrix a szorzásra kommutatív. Tekintettel a legújabban megismert mátrixtípusokra, összefoglaljuk, hogy milyen esetekkel találkoztunk az eddigiekben, ahol a kommutativitás fennáll.

Bármely kvadratikus mátrix	fel- cserélhető	a zérusmátrixszal az egységmátrixszal saját polinomjával az adjungáltjával az inverzével (ha az létezik) inverze polinomjával
Bármely diagonálmátrix		bármely diagonálmátrixszal
Bármely ciklikus mátrix		bármely ciklikus mátrixszal
Bármely Toeplitz-típusú alsó, ill. felső háromszögmátrix		bármely ugyanolyan rendű Toeplitz-típusú alsó, ill. felső háromszög-mátrixszal
Bármely normálmátrix, azaz hermitikus ferdén hermitikus szimmetrikus valós ferdén szimmetrikus valós unitér ortogonális valós		a saját transzponált konjugáltjával

1.4.2 Szimmetrikus egyenletes kontinuáns mátrix invertálása

Láttuk, hogy a szimmetrikus egyenletes kontinuáns mátrixok (1.4.2) szerinti általános alakja felírható az (1.4.3) szerinti \mathbf{K} mátrix polinomjaként, $a\mathbf{E} + b\mathbf{K}$ alakban. Ennek inverze ($-b$ kiemelésével és az $x = -\frac{a}{b}$ jelölés bevezetésével) az alábbi alakban írható:

$$(a\mathbf{E} + b\mathbf{K})^{-1} = -\frac{1}{b}(x\mathbf{E} - \mathbf{K})^{-1}.$$

Tehát az általános szimmetrikus egyenletes kontinuáns mátrix inverzének meghatározásához elegendő az egyetlen paramétert tartalmazó

$$(1.4.20) \quad x\mathbf{E} - \mathbf{K} = \begin{bmatrix} x-1 & 0 & \dots & 0 \\ -1 & x-1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & x-1 \\ 0 & 0 & \dots & -1 & x \end{bmatrix}$$

speciális szimmetrikus kontinuáns mátrix invertálása.

A feladatot az inverz mátrixra kapott (1.2.34) képlet segítségével oldhatjuk meg. Ehhez ismernünk kell egyrészt a mátrix determinánsát, másrészt adjungáltjának elemeit. Nyilvánvaló, hogy n -edrendű mátrix esetén a determináns az x paraméter n -edfokú polinomja, az adjungált mátrix elemei pedig legfeljebb $(n-1)$ -edfokú polinomjai. Felmerül a gondolat, nem lehet-e ezeket a polinomokat valamilyen egységes és viszonylag egyszerű alakban kifejezni? Látni fogjuk, hogy ez lehetséges, mégpedig a Csebisev-polinomok és a trigonometrikus, ill. hiperbolikus függvények közötti szoros kapcsolat kihasználásával. Már most megemlítjük, hogy ennek igazi előnye a későbbiek során, az ún. sajátérték-feladat tárgyalásakor – a 3.4 szakaszban – fog megmutatkozni.

Annak érdekében, hogy végig valós mennyiségekkel számoljunk, a paramétertartomány három szakaszát külön vizsgáljuk. Ezek: $|x| < 2$, $x > 2$, ill. $x < -2$. (Az $x = \pm 2$ eset ezekből egyszerű határátmenettel adódik.)

(a) Vizsgáljuk először az

$$(1.4.21) \quad |x| < 2$$

esetet. Vezessük be ekkor az

$$(1.4.22) \quad x = 2 \cos \theta$$

transzformációval definiált θ új változót (lásd pl. [13]).

A mátrix determinánsának meghatározása: Jelölje D_n a keresett n -edrendű determinánst:

$$D_n = \begin{vmatrix} 2 \cos \theta & -1 & 0 & \dots & 0 \\ -1 & 2 \cos \theta & -1 & \dots & 0 \\ 0 & -1 & 2 \cos \theta & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & -1 \\ 0 & \dots & \dots & -1 & 2 \cos \theta \end{vmatrix}.$$

Teljes indukcióval bebizonyítjuk, hogy

$$(1.4.23) \quad D_n = \frac{\sin(n+1)\theta}{\sin \theta}.$$

Közvetlenül belátható, hogy az állítás $n = 1$ és $n = 2$ esetén igaz, ugyanis $\frac{\sin 2\theta}{\sin \theta} = 2 \cos \theta$ és $\frac{\sin 3\theta}{\sin \theta} = 4 \cos^2 \theta - 1$. Tegyük most fel, hogy az állítás $(n - 1)$ -re és n -re igaz, vagyis

$$D_{n-1} = \frac{\sin n\theta}{\sin \theta} \quad \text{és} \quad D_n = \frac{\sin(n+1)\theta}{\sin \theta},$$

és számítsuk ki a D_{n+1} determinánst. Az utolsó sor szerint kifejtve, az alábbi rekurziós képletet nyerjük: $D_{n+1} = 2 \cos \theta \cdot D_n - D_{n-1}$.

Behelyettesítve D_n és D_{n-1} fenti kifejezését,

$$D_{n+1} = \frac{2 \cos \theta \sin(n+1)\theta - \sin n\theta}{\sin \theta} = \frac{\sin(n+2)\theta}{\sin \theta}$$

adódik, és ezzel az állításunkat bebizonyítottuk.

Megjegyezzük, hogy a $\frac{\sin(n+1)\theta}{\sin \theta}$ kifejezés mint $\cos \theta$ polinomja, az ún. n -edfokú másodfajú Csebisev polinom*.

Az adjungált mátrix meghatározása: Az (1.2.24) összefüggés alapján az invertálandó $x \mathbf{E} - \mathbf{K}$ mátrix adjungáltjának elemeit az alábbi determinánsból határozzuk meg:

$$\begin{array}{c} \left. \begin{array}{c} i-1 \\ i \\ j-i \\ j \\ n-j \end{array} \right\} \begin{array}{c} \overbrace{\begin{array}{ccc} x-1 & & \\ -1 & x & -1 \\ & \ddots & \\ & -1 & x-1 \\ & & -1 & x \end{array}}^{i-1} & \overbrace{\begin{array}{ccc} & & \\ & & \\ & & \\ & & \\ & & \end{array}}^{j-i} & \overbrace{\begin{array}{ccc} & & \\ & & \\ & & \\ & & \\ & & \end{array}}^{j} & \overbrace{\begin{array}{ccc} & & \\ & & \\ & & \\ & & \\ & & \end{array}}^{n-j} \\ \hline & -1 & & & \\ \hline & & -1 & & \\ & & x-1 & & \\ & & -1 & x-1 & \\ & & & \ddots & \\ & & & -1 & x-1 \\ \hline & & & -1 & \\ & & & x-1 & \\ & & & -1 & x-1 \\ & & & & \ddots \\ & & & & -1 & x-1 \\ & & & & & -1 & x \end{array} \end{array}.$$

*P. L. Csebisev (1821–1894) orosz matematikus. A Csebisev-polinomokra vonatkozóan lásd pl. [18], 2. kötet 30. oldal.

Ezt a determinánst úgy kaptuk, hogy az eredeti mátrix determinánsából elhagytuk a j -edik sort és i -edik oszlopot. Az áttekinthetőség kedvéért a zérus elemeket nem írjuk ki. (Itt azt az esetet tekintjük, amikor $i \leq j$; amennyiben $i \geq j$, akkor az indexek szerepet cserélnek.) Tekintettel arra, hogy a determináns a szaggatott vonalak segítségével háromszor három olyan blokkra particionálható, ahol a főátló blokkjai felett csupa zérust tartalmazó blokkok állnak, ezért – a determinánselmélet alapján – a determináns értéke a főátlóban levő blokkok determinánsának a szorzata. Tehát

$$\{\text{adj } (x \mathbf{E} - \mathbf{K})\}_{ij} = (-1)^{i+j} (-1)^{j-i} D_{i-1} D_{n-j}, \quad i \leq j.$$

(A bal felső, ill. a jobb alsó blokk szerkezete ugyanis megegyezik az eredeti determináns szerkezetével, a középső blokk determinánsa pedig a főátló elemeinek a szorzata – ezeknek az elemeknek mindegyike -1 .) Behelyettesítve a determinánssra bizonyított (1.4.23) kifejezést,

$$(1.4.24) \quad \{\text{adj } (x \mathbf{E} - \mathbf{K})\}_{ij} = \begin{cases} \frac{\sin i\theta}{\sin \theta} \frac{\sin(n+1-j)\theta}{\sin \theta}, & \text{ha } i \leq j, \\ \frac{\sin j\theta}{\sin \theta} \frac{\sin(n+1-i)\theta}{\sin \theta}, & \text{ha } i \geq j. \end{cases}$$

Az inverz mátrix felírása: Az eredeti mátrix determinánsának és adjungáltjának ismeretében most már (1.2.34) alapján a keresett inverz elemeire a következő adódik:

$$(1.4.25) \quad \{(x \mathbf{E} - \mathbf{K})^{-1}\}_{ij} = \begin{cases} \frac{\sin i\theta}{\sin \theta} \frac{\sin(n+1-j)\theta}{\sin(n+1)\theta}, & \text{ha } i \leq j, \\ \frac{\sin j\theta}{\sin \theta} \frac{\sin(n+1-i)\theta}{\sin(n+1)\theta}, & \text{ha } i \geq j. \end{cases}$$

(b) Abban az esetben, ha $x > 2$, az

$$(1.4.26) \quad x = 2 \cosh \theta$$

transzformációval azt kapjuk, hogy

$$(1.4.27) \quad D_n = \frac{\sinh(n+1)\theta}{\sinh \theta}.$$

(Ez a kifejezés ugyanolyan polinomja a $\cosh \theta$ függvénynek, mint $\frac{\sin(n+1)\theta}{\sin \theta}$ a $\cos \theta$ függvénynek.) Az inverz elemei az (1.4.25) összefüggéssel analóg módon írhatók fel, csupán a trigonometrikus függvényeket kell hiperbolikus függvényekkel helyettesíteni.

(c) Ha $x < -2$, akkor az

$$(1.4.28) \quad x = -2 \cosh \theta$$

helyettesítéssel – előjeltől eltekintve – ugyanarra az eredményre jutunk, mint az előbb. Ekkor ugyanis a (-1) -et kiemelve, az adjungált elemeit a sakktáblaszabálynak megfelelő előjellel kell ellátni (a középső blokk determinánsa 1, ezért marad meg a sakktábla-előjel). Vagyis ebben az esetben

$$(1.4.29) \quad \begin{bmatrix} -2\operatorname{ch} \theta & -1 & 0 & \dots & 0 \\ -1 & -2\operatorname{ch} \theta & -1 & \dots & 0 \\ 0 & -1 & -2\operatorname{ch} \theta & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -1 & -2\operatorname{ch} \theta \end{bmatrix}_{ij}^{-1} = \begin{cases} (-1)^{i+j-1} \frac{\operatorname{sh} i\theta}{\operatorname{sh} \theta} \frac{\operatorname{sh} (n+1-j)\theta}{\operatorname{sh} (n+1)\theta}, & \text{ha } i \leq j, \\ (-1)^{i+j-1} \frac{\operatorname{sh} j\theta}{\operatorname{sh} \theta} \frac{\operatorname{sh} (n+1-i)\theta}{\operatorname{sh} (n+1)\theta}, & \text{ha } i \geq j. \end{cases}$$

(d) Végül abban a határesetben, amikor $x = \pm 2$, vagyis az (1.4.22), ill. az (1.4.28) transzformációban $\theta = 0$, tekintettel arra, hogy a determináns értéke elemeinek folytonos függvénye, az inverz elemeinek a kiszámítására az (1.4.23), ill. az (1.4.27) kifejezésben elvégezzük a $\theta \rightarrow 0$ határátmenetet.

A l'Hospital-szabály alkalmazásával (1.4.24), ill. (1.4.29) alapján $x = 2$ esetén

$$(1.4.30) \quad \begin{bmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ 0 & -1 & 2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -1 & 2 \end{bmatrix}_{ij}^{-1} = \begin{cases} \frac{i(n+1-j)}{n+1}, & \text{ha } i \leq j, \\ \frac{j(n+1-i)}{n+1}, & \text{ha } i \geq j; \end{cases}$$

$x = -2$ esetén pedig

$$\begin{bmatrix} -2 & -1 & 0 & \dots & 0 \\ -1 & -2 & -1 & \dots & 0 \\ 0 & -1 & -2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -1 & -2 \end{bmatrix}_{ij}^{-1} = \begin{cases} (-1)^{i+j-1} \frac{i(n+1-j)}{n+1}, & \text{ha } i \leq j, \\ (-1)^{i+j-1} \frac{j(n+1-i)}{n+1}, & \text{ha } i \geq j. \end{cases}$$

Megjegyzés. Bevezetve az (1.4.20) mátrix determinánsára a

$$(1.4.31) \quad D_n = U_n(x)$$

jelölést, az inverz mátrix elemeit a teljes paramétertartományra egységesen írhatjuk fel a következő alakban:

$$(1.4.32) \quad \{(x\mathbf{E} - \mathbf{K})^{-1}\}_{ij} = \begin{cases} \frac{U_{i-1}(x)U_{n-j}(x)}{U_n(x)}, & \text{ha } i \leq j, \\ \frac{U_{j-1}(x)U_{n-i}(x)}{U_n(x)}, & \text{ha } i \geq j. \end{cases}$$

A szimmetrikus egyenletes kontinuáns mátrix inverzére kapott eredményekből látható, hogy az – ugyancsak szimmetrikus – inverz mátrix elemei mindig két olyan tényező szorzataként írhatók fel, amelyek egyike csak a sorindextől függ, a másik pedig csak az oszlopindextől. Általában, az ilyen tulajdonságú mátrixokat *egypárú mátrixoknak* nevezzük (lásd pl. [7]).

1.4.13 definíció. *Egypárú az $\mathbf{A} = [a_{ij}]$ mátrix, ha elemei a következő alakúak:*

$$(1.4.33) \quad a_{ij} = \begin{cases} \varphi_i \psi_j, & \text{ha } i \leq j, \\ \varphi_j \psi_i, & \text{ha } i \geq j. \end{cases}$$

Azt is mondhatjuk, hogy az egypárú mátrix származtatható úgy, hogy a diád főátló feletti elemeit a főátlóra tükrözzük.

A szimmetrikus kontinuáns mátrix és az egypárú mátrix kapcsolatára az 1.5.3 pontban visszatérünk.

1.4.3 Nilpotens és ciklikus mátrix polinomjának invertálása

Az előzőekben megismerkedtünk a speciális nilpotens \mathbf{H} mátrixszal és annak $\mathbf{A}(a_0, a_1, \dots, a_{n-1})$ polinomjával. Most a \mathbf{H} mátrix speciális tulajdonságainak kihasználásával meghatározzuk az

$$(1.4.34) \quad \mathbf{A}(1, -x, 0, \dots, 0) = \mathbf{E} - x\mathbf{H} = \begin{bmatrix} 1 & -x & 0 & \dots & 0 \\ 0 & 1 & -x & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & -x \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix}$$

mátrixpolinom inverzét. A feladat megoldásához nyilvánvalóan azt kell kihasználnunk, hogy a \mathbf{H} mátrix n indexű nilpotens mátrix, azaz

$$(1.4.35) \quad \mathbf{H}^n = 0.$$

Az $\mathbf{E} - x\mathbf{H}$ mátrixpolinomnak megfelelő közösleges (skalár) polinom $(1 - x)$; ez fellép az alábbi közismert azonosságban:

$$(1.4.36) \quad (1 - x)(1 + x + x^2 + \dots + x^{n-1}) \equiv 1 - x^n.$$

Vegyük észre, hogy az azonosságban x helyébe az $x\mathbf{H}$ mátrixot helyettesítve, (1.4.35) miatt a jobb oldalon az egységmátrix adódik. Ezért a bal oldal második tényezője adja a keresett inverz mátrixot:

$$(\mathbf{E} - x\mathbf{H})(\mathbf{E} + x\mathbf{H} + x^2\mathbf{H}^2 + \dots + x^{n-1}\mathbf{H}^{n-1}) = \mathbf{E} - x^n\mathbf{H}^n = \mathbf{E},$$

tehát

$$(1.4.37) \quad (\mathbf{E} - x\mathbf{H})^{-1} = \mathbf{E} + x\mathbf{H} + x^2\mathbf{H}^2 + \dots + x^{n-1}\mathbf{H}^{n-1} =$$

$$= \begin{bmatrix} 1 & x & x^2 & x^3 & \dots & x^{n-1} \\ 0 & 1 & x & x^2 & \dots & x^{n-2} \\ 0 & 0 & 1 & x & \dots & x^{n-3} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & x \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}.$$

Lényegében a most alkalmazott módszer segítségével invertálható az alábbi speciális ciklikus mátrix:

$$(1.4.38) \quad \mathbf{C}(1, x, x^2, \dots, x^{n-1}) = \begin{bmatrix} 1 & x & x^2 & x^3 & \dots & x^{n-1} \\ x^{n-1} & 1 & x & x^2 & \dots & x^{n-2} \\ x^{n-2} & x^{n-1} & 1 & x & \dots & x^{n-3} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x^2 & x^3 & x^4 & \dots & 1 & x \\ x & x^2 & x^3 & \dots & x^{n-1} & 1 \end{bmatrix}.$$

A feladat megoldása során a ciklikus mátrix (1.4.14) tulajdonságát használjuk ki, amely szerint \mathbf{C} felírható az $\mathbf{\Omega}$ elemi ciklikus mátrix polinomjaként:

$$\mathbf{C}(1, x, x^2, \dots, x^{n-1}) = \mathbf{E} + x\mathbf{\Omega} + x^2\mathbf{\Omega}^2 + \dots + x^{n-1}\mathbf{\Omega}^{n-1}.$$

Ha most az (1.4.36) skalár azonosságban x helyére az $x\mathbf{\Omega}$ mátrixot írjuk:

$$(\mathbf{E} - x\mathbf{\Omega})(\mathbf{E} + x\mathbf{\Omega} + x^2\mathbf{\Omega}^2 + \dots + x^{n-1}\mathbf{\Omega}^{n-1}) = \mathbf{E} - x^n\mathbf{\Omega}^n,$$

és figyelembe vesszük, hogy (1.4.12) utolsó egyenlete szerint $\mathbf{\Omega}^n = \mathbf{E}$, akkor rögtön adódik az eredmény, és egyúttal az invertálhatóság feltétele is. Ugyanis, ha $1 - x^n \neq 0$, azaz x nem egyenlő egyik n -edik egységgyökkel sem, akkor oszthatunk az $1 - x^n$ mennyiséggel és így

$$(1.4.39) \quad \{\mathbf{C}(1, x, x^2, \dots, x^{n-1})\}^{-1} = \frac{1}{1 - x^n}(\mathbf{E} - x\mathbf{\Omega}) =$$

$$= \frac{1}{1 - x^n} \mathbf{C}(1, -x, 0, 0, \dots, 0) =$$

$$= \frac{1}{1 - x^n} \begin{bmatrix} 1 & -x & 0 & 0 & \dots & 0 \\ 0 & 1 & -x & 0 & \dots & 0 \\ 0 & 0 & 1 & -x & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & -x \\ -x & 0 & 0 & \dots & 0 & 1 \end{bmatrix}.$$

1.5 HIPERMÁTRIXOK

Már az 1.1 szakaszban megismertük a particionálás fogalmát és a particionált mátrixok összeadására és szorzására vonatkozó néhány egyszerű szabályt. Láttuk, hogy particionálással a mátrix téglalap alakú blokkokra bontható. A blokkokból álló mátrixot olyan „mátrixnak” tekinthetjük, amelynek elemei maguk is mátrixok. Az ilyen „mátrix elemű mátrixot” hipermátrixnak nevezzük. Hasonló értelemben használjuk a hipervektor fogalmát is, amelynek elemei maguk is vektorok.

A hipermátrix fogalmának bevezetése számos gyakorlati előnnyel jár. Például, ha olyan nagy rendszámú mátrixokkal kell numerikus feladatokat megoldani, hogy a rendelkezésre álló számítógép memóriakapacitása a szokásos módszerekhez nem elegendő, akkor a particionálás lehetővé teszi, hogy a hipermátrixok blokkjaival – vagyis kisebb rendszámú mátrixokkal – dolgozzunk.

Ebben a fejezetben a legegyszerűbb – négy blokkból álló – hipermátrixokra vonatkozó fontosabb összefüggéseket ismertetjük. Ezeket az eredményeket alkalmazzuk a módosított mátrixok inverzének és adjungáltjának kiszámítására, majd megmutatjuk egy adott mátrix inverze és egy minor-mátrixának inverze közötti kapcsolatot. Ezt alkalmazzuk a szimmetrikus kontinuáns mátrix és az egypárú mátrixok között fennálló érdekes kapcsolat megmutatására. Ez azt mondja ki, hogy a szimmetrikus kontinuáns mátrixok (és csak azok) olyan tulajdonságúak, hogy az inverzük egypárú mátrix.

1.5.1 Hipermátrixok szorzása és faktorizációja

Először néhány új fogalmat vezetünk be, amelyek többnyire a közönséges mátrixokra már ismert fogalmak általánosításai hipermátrixokra.

1.5.1 definíció. Az

$$(1.5.1) \quad [\mathbf{A}_{ij}] = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \dots \mathbf{A}_{1n} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \dots \mathbf{A}_{2n} \\ \dots & \dots \dots \dots \dots \dots \dots \\ \mathbf{A}_{m1} & \mathbf{A}_{m2} \dots \mathbf{A}_{mn} \end{bmatrix}$$

alakú hipermátrixot $m \times n$ típusú hipermátrixnak, sorait **blokk soroknak**, oszlopait pedig **blokk oszlopoknak** nevezzük; ha $m = n$, akkor **n -edrendű hipermátrixról** beszélünk.

1.5.2 definíció. Ha az $[\mathbf{A}_{ij}]$ hipermátrix főátlójában álló \mathbf{A}_{ii} blokkok kvadratus mátrixok, akkor a hipermátrixot **szimmetrikusan particionált**nak nevezzük.

Speciális szimmetrikusan particionált mátrixszal már az eddigiekben is találkoztunk: ilyenek a kvázidiagonál-mátrixok, amelyeknek csak a főátlójá-

ban állhatnak zérustól különböző kvadratikus blokkok. Ezért a hipermátrixok terminológiájának megfelelően, a kvázidiagonál-mátrixokat *hiperdiagonál-mátrixoknak* is nevezhetjük.

A következőkben a hipermátrixokra vonatkozó szabályokat és összefüggéseket az egyszerűség kedvéért általában n -edrendű hipermátrixokra mondjuk ki, de megjegyezzük, hogy azok – értelemszerűen – általános típusú hipermátrixokra is vonatkozhatnak.

A közönséges mátrixok összeadásának és szorzásának definíciójából következik a hipermátrixok összeadási és szorzási szabálya.

(a) Ha az \mathbf{A}_{ij} és \mathbf{B}_{ij} blokkok minden ij indexpárra megegyező típusúak, akkor a hipermátrixok blokkonként összegezhetőek:

$$(1.5.2) \quad [\mathbf{A}_{ij}] + [\mathbf{B}_{ij}] = [\mathbf{A}_{ij} + \mathbf{B}_{ij}].$$

(b) Ha az $[\mathbf{A}_{ik}]$ hipermátrix oszlopok szerinti particionálása megegyezik a $[\mathbf{B}_{kj}]$ hipermátrix sorok szerinti particionálásával, azaz \mathbf{A}_{ik} oszlopainak a száma minden k -ra megegyezik \mathbf{B}_{kj} sorainak a számával, akkor a hipermátrixok szorzása blokkjaik segítségével ugyanúgy végezhető el, mint a közönséges mátrixok szorzása elemeik segítségével:

$$(1.5.3) \quad [\mathbf{A}_{ik}][\mathbf{B}_{kj}] = \left[\sum_{k=1}^n \mathbf{A}_{ik} \mathbf{B}_{kj} \right].$$

A közönséges mátrixok körében definiált diádoknak megfelelően, bevezetjük a hiperdiád fogalmát.

1.5.3 definíció. Az egyetlen blokkoszlopból és egyetlen blokkosorból álló hipermátrixok szorzatát – feltéve, hogy a szorzás értelmezve van –, **hiperdiádnak** nevezzük:

$$(1.5.4) \quad \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_n \end{bmatrix} \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 & \dots & \mathbf{B}_n \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 \mathbf{B}_1 & \mathbf{A}_1 \mathbf{B}_2 & \dots & \mathbf{A}_1 \mathbf{B}_n \\ \mathbf{A}_2 \mathbf{B}_1 & \mathbf{A}_2 \mathbf{B}_2 & \dots & \mathbf{A}_2 \mathbf{B}_n \\ \dots & \dots & \dots & \dots \\ \mathbf{A}_n \mathbf{B}_1 & \mathbf{A}_n \mathbf{B}_2 & \dots & \mathbf{A}_n \mathbf{B}_n \end{bmatrix}.$$

A hipermátrixok szorzására vonatkozó (1.5.3) összefüggésből kapjuk, hogy

$$(1.5.5) \quad \left[\sum_{k=1}^n \mathbf{A}_{ik} \mathbf{B}_{kj} \right] = \sum_{k=1}^n [\mathbf{A}_{ik} \mathbf{B}_{kj}].$$

Ez azt jelenti, hogy a szorzatösszeg minden tagját egy-egy hipermátrixba foglaljuk össze; formailag a mátrixjel és a szummajel felcserélődik. A jobb oldali kifejezés rögzített k mellett egy-egy hiperdiád. Két hipermátrix szorzata tehát kifejezhető hiperdiádok összegeként. Fordított irányban olvasva

az (1.5.5) összefüggést, megállapítható az a közönséges mátrixoknál megismert összefüggés, mely szerint hiperdiádok összege megegyezik a hiperdiádok blokkoszlopaiból, ill. blokkSORaiból alkotott hipermátrixok szorzatával:

$$(1.5.6) \quad \sum_{k=1}^n [\mathbf{A}_{ik} \mathbf{B}_{kj}] = \left[\sum_{k=1}^n \mathbf{A}_{ik} \mathbf{B}_{kj} \right] = [\mathbf{A}_{ik}] [\mathbf{B}_{kj}].$$

Hipermátrixok szorzási szabályának ismételt alkalmazásával jutunk három hipermátrix szorzatához:

$$(1.5.7) \quad [\mathbf{A}_{ik}] [\mathbf{B}_{kl}] [\mathbf{C}_{lj}] = \left[\sum_{l=1}^n \sum_{k=1}^n \mathbf{A}_{ik} \mathbf{B}_{kl} \mathbf{C}_{lj} \right].$$

Speciális esetként, amikor a $[\mathbf{B}_{kl}]$ hipermátrixnak csak a főátlójában állnak zérustól különböző blokkok, vagyis amikor a középső tényező hiperdiagonálmátrix, akkor az (1.5.7) jobb oldalán álló kettős szumma helyett egyszeres szummát kapunk, majd az (1.5.5) összefüggéshez hasonlóan, felcserélve a mátrixjelet a szummajellel, a következő adódik:

$$(1.5.8) \quad [\mathbf{A}_{ik}] \langle \mathbf{B}_k \rangle [\mathbf{C}_{kj}] = \left[\sum_{k=1}^n \mathbf{A}_{ik} \mathbf{B}_k \mathbf{C}_{kj} \right] = \sum_{k=1}^n [\mathbf{A}_{ik} \mathbf{B}_k \mathbf{C}_{kj}].$$

Ez a hiperdiádok „lineáris kombinációját” jelenti, úgy értve ezt, hogy a lineáris kombináció „együtthatói” – amelyek maguk is mátrixok – a hiperdiádok egy-egy blokkoszlopa és blokkSORa közé ékelődnek:

$$(1.5.9) \quad \begin{bmatrix} \mathbf{A}_{11} \dots \mathbf{A}_{1n} \\ \dots \dots \dots \\ \mathbf{A}_{n1} \dots \mathbf{A}_{nn} \end{bmatrix} \begin{bmatrix} \mathbf{B}_1 & & \\ & \ddots & \\ & & \mathbf{B}_n \end{bmatrix} \begin{bmatrix} \mathbf{C}_{11} \dots \mathbf{C}_{1n} \\ \dots \dots \dots \\ \mathbf{C}_{n1} \dots \mathbf{C}_{nn} \end{bmatrix} = \\ = \begin{bmatrix} \mathbf{A}_{11} \\ \vdots \\ \mathbf{A}_{n1} \end{bmatrix} \mathbf{B}_1 [\mathbf{C}_{11} \dots \mathbf{C}_{1n}] + \begin{bmatrix} \mathbf{A}_{12} \\ \vdots \\ \mathbf{A}_{n2} \end{bmatrix} \mathbf{B}_2 [\mathbf{C}_{21} \dots \mathbf{C}_{2n}] + \dots \\ \dots + \begin{bmatrix} \mathbf{A}_{1n} \\ \vdots \\ \mathbf{A}_{nn} \end{bmatrix} \mathbf{B}_n [\mathbf{C}_{n1} \dots \mathbf{C}_{nn}].$$

Ezt az összefüggést fordított irányban olvasva kimondható, hogy hiperdiádok fenti értelemben vett lineáris kombinációja három mátrix szorzataként írható fel: a szorzat első tényezője a hiperdiádok blokkoszlopaiból, harmadik tényezője a hiperdiádok blokkSORaiból alkotott hipermátrix, második tényezője pedig a kombináció „együtthatóiból” álló hiperdiagonálmátrix.

1.5.2 Szimmetrikusan particionált másodrendű hipermátrix faktorizálása, determinánsa, inverze

Az alábbiakban szimmetrikusan particionált másodrendű hipermátrixokkal foglalkozunk. Először előállítjuk ezek kétféle faktorizációját, majd ennek segítségével meghatározzuk a determinásukat és inverzüket.

1.5.1 tétel. Ha

$$(1.5.10) \quad \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$$

szimmetrikusan particionált másodrendű hipermátrix és \mathbf{A} nonszinguláris, akkor az (1.5.10) hipermátrix egy faktorizációja

$$(1.5.11) \quad \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{CA}^{-1} & \mathbf{E} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{E} & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{E} \end{bmatrix}$$

alakban nyerhető.

Bizonyítás. Ha az (1.5.10) hipermátrixból levonjuk a nonszinguláris \mathbf{A} blokkal generált \mathbf{A}^{-1} „együtthatójú”

$$\begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} \mathbf{A}^{-1} [\mathbf{A} \quad \mathbf{B}]$$

hiperdiádot, akkor azt kapjuk, hogy

$$(1.5.12) \quad \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} - \begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} \mathbf{A}^{-1} [\mathbf{A} \quad \mathbf{B}] = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B} \end{bmatrix};$$

ez egyetlen hiperdiád „lineáris kombinációja”:

$$(1.5.13) \quad \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{E} \end{bmatrix} (\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}) [\mathbf{0} \quad \mathbf{E}].$$

Behelyettesítve az (1.5.13) összefüggést (1.5.12)-be, az adott (1.5.10) hipermátrix hiperdiádok lineáris kombinációjaként adódik:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} \mathbf{A}^{-1} [\mathbf{A} \quad \mathbf{B}] + \begin{bmatrix} \mathbf{0} \\ \mathbf{E} \end{bmatrix} (\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}) [\mathbf{0} \quad \mathbf{E}].$$

A jobb oldal első tagjában \mathbf{A} a blokkoszlopból jobbra, a blokkSORból balra „kiemelhető”, és így

$$(1.5.14) \quad \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{E} \\ \mathbf{CA}^{-1} \end{bmatrix} \mathbf{A} [\mathbf{E} \quad \mathbf{A}^{-1}\mathbf{B}] + \begin{bmatrix} \mathbf{0} \\ \mathbf{E} \end{bmatrix} (\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}) [\mathbf{0} \quad \mathbf{E}].$$

Felhasználva az (1.5.9) összefüggést, megkapjuk az (1.5.11) faktorizációt. ■

A most bebizonyított tétellel analóg a következő tétel.

1.5.2 tétel. *Ha az (1.5.10) hipermátrixban \mathbf{D} nemszinguláris blokk, akkor a hipermátrix*

$$(1.5.15) \quad \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{E} & \mathbf{BD}^{-1} \\ \mathbf{0} & \mathbf{E} \end{bmatrix} \begin{bmatrix} \mathbf{A} - \mathbf{BD}^{-1}\mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{D}^{-1}\mathbf{C} & \mathbf{E} \end{bmatrix}$$

alakban írható fel.

A tétel bizonyítása hasonló az 1.5.1 tételéhez, itt azonban nem az \mathbf{A} , hanem a \mathbf{D} mátrixszal generált, \mathbf{D}^{-1} „együtthatójú” hiperdiádot vonjuk le az (1.5.10) hipermátrixból. A bizonyítás további menetét az Olvasóra bízuk.

Megjegyzés. Az itt fellépő $\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}$ és $\mathbf{A} - \mathbf{BD}^{-1}\mathbf{C}$ alakú kifejezéseket a szimmetrikusan particionált hipermátrix \mathbf{A} , ill. \mathbf{D} blokkjához tartozó *Schur-komplementumnak** nevezzük (lásd [38]).

Az (1.5.11) és (1.5.15) alakú faktorizáció lehetővé teszi a szimmetrikusan particionált másodrendű – vagyis négy blokkból álló – hipermátrixok determinánsának és inverzének egyszerű kiszámítását. Ha ugyanis felhasználjuk a determinánsok szorzásának azt a szabályát, mely szerint szorzatmátrix determinánsa egyenlő az egyes tényezők determinánsának a szorzatával, akkor az (1.5.11), ill. (1.5.15) előállításból közvetlenül kiolvasható a következő tétel.

1.5.3 tétel. *Az (1.5.10) hipermátrix determinánsa $|\mathbf{A}| \neq 0$ esetén az*

$$(1.5.16) \quad \begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{vmatrix} = |\mathbf{A}| \cdot |\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}|,$$

$|\mathbf{D}| \neq 0$ esetén pedig az

$$(1.5.17) \quad \begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{vmatrix} = |\mathbf{A} - \mathbf{BD}^{-1}\mathbf{C}| \cdot |\mathbf{D}|$$

képlettel számítható.

A kapott képletekből az is kiolvasható, hogy $|\mathbf{A}| \neq 0$ esetén az adott hipermátrix akkor és csak akkor nemszinguláris, ha

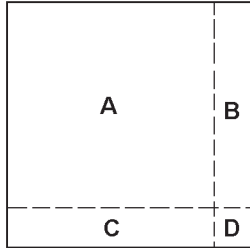
$$(1.5.18) \quad |\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}| \neq 0,$$

$|\mathbf{D}| \neq 0$ esetén pedig ha

$$(1.5.19) \quad |\mathbf{A} - \mathbf{BD}^{-1}\mathbf{C}| \neq 0.$$

*I. Schur (1875–1941) német matematikus.

Tekintsük azt a speciális esetet, amikor \mathbf{B} egyetlen oszlopvektor, \mathbf{C} egyetlen sorvektor, \mathbf{D} pedig egyetlen elem:



Ennek megfelelően, kis betűkkel jelölve a megfelelő blokkokat, vagyis $\mathbf{B} = \mathbf{b}$, $\mathbf{C} = \mathbf{c}^\top$, $\mathbf{D} = d$, az egyetlen sorral és oszloppal szegélyezett \mathbf{A} mátrix determinánsának kiszámítására az (1.5.16) képlet alkalmazásával a következő adódik:

$$(1.5.20) \quad \begin{vmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{c}^\top & d \end{vmatrix} = |\mathbf{A}|(d - \mathbf{c}^\top \mathbf{A}^{-1} \mathbf{b}).$$

Beszorzás után a szegélyezett mátrix determinánsára a következő képletet kapjuk:

$$(1.5.21) \quad \begin{vmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{c}^\top & d \end{vmatrix} = d|\mathbf{A}| - \mathbf{c}^\top (\text{adj } \mathbf{A}) \mathbf{b}.$$

Az (1.5.20) összefüggésből kiolvasható annak szükséges és elégséges feltétele, hogy a szegélyezett mátrix szinguláris legyen:

$$d - \mathbf{c}^\top \mathbf{A}^{-1} \mathbf{b} = 0.$$

Mivel a determináns az elemeinek folytonos függvénye, az (1.5.21) képletből kiolvasható, hogy egy szegélyezett szinguláris \mathbf{A} mátrix determinánsa az alábbi képlettel számítható:

$$\begin{vmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{c}^\top & d \end{vmatrix} = -\mathbf{c}^\top (\text{adj } \mathbf{A}) \mathbf{b} \quad (\text{ha } |\mathbf{A}| = 0).$$

Most bevezetjük a hipermátrix inverzének a fogalmát.

1.5.4 definíció. Hipermátrix inverzén azt a hipermátrixot értjük, amelylyel az adott hipermátrixot megszorozva, a – megfelelően particionált – egység-mátrixot kapjuk.

Az (1.5.11) és (1.5.15) faktorizált előállítás lehetővé teszi az (1.5.10) hipermátrix inverzének egyszerű előállítását. Ugyanis a mátrixszorzat inverzének kiszámítására vonatkozó szabályt alkalmazva, az egyes tényezők inverzét

fordított sorrendben kell összeszorozni. A középső tényező mindkét faktORIZÁCIÓ esetén hiperdiagonál-mátrix, és ennek inverze – amint az közvetlenül belátható – az egyes blokkok invertálásával állítható elő. A bal, ill. jobb oldali tényező pedig mindkét esetben az egység mátrixnak és egy nilpotens mátrixnak az összege, inverzük tehát az (1.4.37) mátrix invertálása szerint végezhető. Tekintsük például az (1.5.11) képlet jobb oldalán álló első tényezőt:

$$(1.5.22) \quad \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{CA}^{-1} & \mathbf{E} \end{bmatrix} = \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{E} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{CA}^{-1} & \mathbf{0} \end{bmatrix}.$$

Mivel

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{CA}^{-1} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{CA}^{-1} & \mathbf{0} \end{bmatrix} = \mathbf{0},$$

tehát (1.5.22) az egység mátrixnak és egy 2 indexű nilpotens mátrixnak az összege. Mint tudjuk (lásd (1.4.37)), ennek inverze

$$(\mathbf{E} + \mathbf{N})^{-1} = \mathbf{E} - \mathbf{N}$$

alakban írható fel, tehát

$$\begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{CA}^{-1} & \mathbf{E} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ -\mathbf{CA}^{-1} & \mathbf{E} \end{bmatrix}.$$

Ugyanígy számítható a többi tényező inverze is, tehát az (1.5.10) hiper-mátrix inverzének faktorizált alakja (1.5.11) alapján és feltéve, hogy (1.5.18) teljesül

$$(1.5.23) \quad \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{E} & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{E} \end{bmatrix} \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B})^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ -\mathbf{CA}^{-1} & \mathbf{E} \end{bmatrix}.$$

Hasonlóképpen kaphatjuk meg az (1.5.15) képlet segítségével ugyanennek egy másik alakját (feltéve, hogy (1.5.19) teljesül):

$$(1.5.24) \quad \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{E} \end{bmatrix} \begin{bmatrix} (\mathbf{A} - \mathbf{BD}^{-1}\mathbf{C})^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{E} & -\mathbf{BD}^{-1} \\ \mathbf{0} & \mathbf{E} \end{bmatrix}.$$

Ha az (1.5.23), ill. (1.5.24) képletben szereplő beszorzásokat elvégezzük, akkor megkapjuk a szimmetrikusan négy blokkra particionált hiper-mátrixok inverzének particionált alakját. A kapott eredményt a következő tételben foglaljuk össze.

1.5.4 tétel. Az (1.5.10) hiper-mátrix inverzének particionált alakú előállítás

$|\mathbf{A}| \neq 0$ esetén

$$(1.5.25) \quad \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \\ -(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{bmatrix},$$

$|\mathbf{D}| \neq 0$ esetén pedig

$$(1.5.26) \quad \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & -(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1} \end{bmatrix}.$$

Speciális esetként tekintünk az (1.5.20) alatti szegélyezett mátrixot. Ennek inverzére (1.5.25) szerint az alábbi összefüggést kapjuk:

$$(1.5.25a) \quad \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{c}^\top & d \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} + \frac{\mathbf{A}^{-1}\mathbf{b}\mathbf{c}^\top\mathbf{A}^{-1}}{d - \mathbf{c}^\top\mathbf{A}^{-1}\mathbf{b}} & -\frac{\mathbf{A}^{-1}\mathbf{b}}{d - \mathbf{c}^\top\mathbf{A}^{-1}\mathbf{b}} \\ -\frac{\mathbf{c}^\top\mathbf{A}^{-1}}{d - \mathbf{c}^\top\mathbf{A}^{-1}\mathbf{b}} & \frac{1}{d - \mathbf{c}^\top\mathbf{A}^{-1}\mathbf{b}} \end{bmatrix}.$$

18. Példa. Határozzuk meg az

$$\mathbf{M} = \begin{bmatrix} 2 & -1 & & & & & \\ -1 & 2 & -1 & & & & \\ & & \ddots & & & & \\ & & & -1 & & & \\ & & & & -1 & 2 & -1 \\ \cdots & & & & & -1 & 1 \end{bmatrix} \begin{matrix} \underbrace{1} \\ \underbrace{2} \\ \vdots \\ \underbrace{n} \end{matrix} = \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{c}^\top & d \end{bmatrix}$$

n -edrendű szegélyezett mátrix inverzét, felhasználva az (1.4.30) inverz mátrixot.

Megoldás. Az \mathbf{A} itt $n - 1$ -edrendű mátrix, inverzének elemei

$$[\mathbf{A}^{-1}]_{ij} = \begin{cases} \frac{i(n-j)}{n}, & i \leq j, \\ \frac{j(n-i)}{n}, & i \geq j \end{cases}$$

alakban írhatók. Mivel $\mathbf{c}^\top = [0 \cdots -1]$, $\mathbf{b} = \mathbf{c}$ és $d = 1$, az (1.5.25a) képlet alapján

$$d - \mathbf{c}^\top\mathbf{A}^{-1}\mathbf{b} = 1 - \frac{n-1}{n} = \frac{1}{n},$$

$\mathbf{A}^{-1}\mathbf{b} = \begin{bmatrix} i \\ n \end{bmatrix}$, $\mathbf{c}^T\mathbf{A}^{-1} = \begin{bmatrix} j \\ n \end{bmatrix}$, és így $\left[\frac{\mathbf{A}^{-1}\mathbf{b}\mathbf{c}^T\mathbf{A}^{-1}}{d - \mathbf{c}^T\mathbf{A}^{-1}\mathbf{b}} \right]_{ij} = \frac{i \cdot j}{n}$ adódik. Azonban

$$\left[\mathbf{A}^{-1} + \frac{\mathbf{A}^{-1}\mathbf{b}\mathbf{c}^T\mathbf{A}^{-1}}{d - \mathbf{c}^T\mathbf{A}^{-1}\mathbf{b}} \right]_{ij} = \begin{cases} \frac{i(n-j)}{n} + \frac{ij}{n} = i, & i \leq j \\ \frac{j(n-i)}{n} + \frac{ij}{n} = j, & i \geq j, \end{cases}$$

$i, j = 1, 2, \dots, n-1$. Ezt behelyettesítve az (1.5.25a) képletbe, a szegélyezett mátrix elemeire kapjuk, hogy

$$\{\mathbf{M}^{-1}\}_{ij} = \begin{cases} i, & i \leq j, \\ j, & i \geq j, \end{cases} \quad i, j = 1, 2, \dots, n.$$

Például $n = 5$ esetén az inverz mátrix:

$$\mathbf{M}^{-1} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}.$$

* * *

1.5.3 Módosított mátrix és minormátrix inverze

Az 1.5.4 tételből igen egyszerűen adódik az alkalmazások során gyakran használt további két fontos eredmény. Az első a módosított mátrixok inverzének a kiszámítására, a második pedig egy minormátrix inverzének előállítására vonatkozik.

Az első feladat a következőképpen fogalmazható meg: tegyük fel, hogy adott n -edrendű mátrixnak ismerjük az inverzét, és változtassuk meg az adott mátrix bizonyos elemeit; kérdés, hogy meghatározható-e a módosított mátrix inverze az adott mátrix inverzének a segítségével anélkül, hogy ismét n -edrendű mátrixot kellene invertálni? Bebizonyítjuk, hogy ha a módosított mátrix és az eredeti mátrix különbsége r -edrangú mátrix ($r < n$), akkor a módosított mátrix inverzének kiszámítása egy r -edrendű mátrix invertálására redukálható. Az eredményt tétel formájában fogalmazzuk meg (lásd [45]).

1.5.5 tétel. (Woodbury* tétele.) Legyen \mathbf{A} nemszinguláris n -edrendű mátrix, és \mathbf{M} olyan r -edrangú, n -edrendű mátrix, amelynek egy minimális diadikus felbontása $\mathbf{M} = \mathbf{BC}$. Ha

$$(1.5.27) \quad |\mathbf{E} - \mathbf{CA}^{-1}\mathbf{B}| \neq 0,$$

*M. A. Woodbury USA-beli biomatematikus.

akkor az $\mathbf{A} - \mathbf{M}$ módosított mátrix nonszinguláris, és inverze előállítható

$$(1.5.28) \quad (\mathbf{A} - \mathbf{M})^{-1} = \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{E} - \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1}$$

alakban.

Bizonyítás. Az állítás első része az 1.5.3 tételből következik $\mathbf{D} = \mathbf{E}$ választás mellett. Az (1.5.16) és (1.5.17) képletek összevetéséből ugyanis

$$|\mathbf{A}| \cdot |\mathbf{E} - \mathbf{CA}^{-1}\mathbf{B}| = |\mathbf{A} - \mathbf{BC}|$$

adódik, tehát (1.5.27) fennállásából következik, hogy $\mathbf{A} - \mathbf{M} = \mathbf{A} - \mathbf{BC}$ nonszinguláris. Az állítás második része az 1.5.4 tételből következik. Ha felteszünk, hogy \mathbf{A} és \mathbf{D} nonszinguláris, akkor az inverz mátrix egyértelműsége miatt az (1.5.25) és (1.5.26) hipermátrixok bal felső blokkjai azonosak:

$$(\mathbf{A} - \mathbf{BD}^{-1}\mathbf{C})^{-1} = \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1},$$

és innen $\mathbf{D} = \mathbf{E}$ választással azonnal adódik az (1.5.28) összefüggés. ■

Megjegyzés. $\mathbf{D}^{-1} = \mathbf{K}$ jelölés bevezetésével a módosított mátrix inverzének a meghatározására a következő alternatív formula is alkalmazható:

$$(\mathbf{A} - \mathbf{BKC})^{-1} = \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{K}^{-1} - \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1}.$$

Speciális esetként – amikor az adott mátrixot csupán egyetlen diáddal módosítjuk – az 1.5.5 tételből kapjuk, hogy

$$(1.5.29) \quad (\mathbf{A} - \mathbf{uv}^T)^{-1} = \mathbf{A}^{-1} + \frac{(\mathbf{A}^{-1}\mathbf{u})(\mathbf{v}^T\mathbf{A}^{-1})}{1 - \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}.$$

Ezt az eredményt *Sherman–Morrison-formulának** nevezzük (lásd [39]). Az egyetlen diáddal módosított mátrix invertálhatóságának szükséges és elégséges feltétele – az (1.5.27) feltétel alapján –

$$1 - \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u} \neq 0.$$

Egy másik speciális eset, amikor az adott mátrixnak csak egyetlen elemét, pl. az ij indexű elemét módosítjuk. Ekkor a módosított mátrix inverze

$$(1.5.30) \quad (\mathbf{A} - k\mathbf{e}_i\mathbf{e}_j^T)^{-1} = \mathbf{A}^{-1} + \frac{k(\mathbf{A}^{-1}\mathbf{e}_i)(\mathbf{e}_j^T\mathbf{A}^{-1})}{1 - k\mathbf{e}_j^T\mathbf{A}^{-1}\mathbf{e}_i}.$$

Mint látható, a módosított mátrix inverzének meghatározásában a számítás egyszerűsödése attól függ, hogy az adott mátrixot milyen rangú mátrixszal – más szóval hány diáddal – módosítjuk, nem pedig attól, hogy a mátrix hány elemét változtatjuk meg.

*J. Sherman USA-beli matematikus.

19. Példa. Határozzuk meg az

$$\mathbf{M} = \begin{bmatrix} 2+x & -1 & & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & \ddots & \\ & & & & -1 & 2 \end{bmatrix}$$

n -edrendű mátrix inverzét!

Megoldás. Mivel \mathbf{M} a 18. Példában szereplő, de n -edrendű \mathbf{A} mátrix 11 indexű elemének módosításával írható fel, alkalmazzuk az (1.5.30) alatti képletet:

$$(\mathbf{A} + x\mathbf{e}_1\mathbf{e}_1^\top)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{e}_1\mathbf{e}_1^\top\mathbf{A}^{-1}}{\frac{1}{x} + \mathbf{e}_1^\top\mathbf{A}^{-1}\mathbf{e}_1}.$$

Felhasználva az (1.4.30) alatti inverz mátrixot:

$$[\mathbf{A}^{-1}]_{ij} = \begin{cases} \frac{i(n+1-j)}{n+1}, & i \leq j, \\ \frac{j(n+1-i)}{n+1}, & i \geq j, \end{cases}$$

$$\mathbf{A}^{-1}\mathbf{e}_1 = \begin{bmatrix} \frac{n+1-i}{n+1} \end{bmatrix}, \quad \mathbf{e}_1^\top\mathbf{A}^{-1} = \begin{bmatrix} \frac{n+1-j}{n+1} \end{bmatrix}, \quad \mathbf{e}_1^\top\mathbf{A}^{-1}\mathbf{e}_1 = \frac{n}{n+1},$$

a keresett inverz elemeire $i \leq j$ esetén

$$\begin{aligned} \frac{i(n+1-j)}{n+1} - \frac{\frac{n+1-i}{n+1} \frac{n+1-j}{n+1}}{\frac{1}{x} - \frac{n}{n+1}} &= \frac{n+1-j}{n+1} \frac{i[n+1+nx] - x(n+1-i)}{n+1+nx} = \\ &= \frac{n+1-j}{n+1} \frac{(n+1)i + (n+1)ix - (n+1)x}{n+1+nx}, \end{aligned}$$

azaz

$$[\mathbf{M}^{-1}]_{ij} = \begin{cases} \frac{i(x+1)-x}{n(x+1)+1}(n+1-j), & i \leq j, \\ \frac{j(x+1)-x}{n(x+1)+1}(n+1-i), & i \geq j, \end{cases}$$

adódik. Az is látható innen, hogy a módosított mátrix

$$\frac{1}{x} + \frac{n}{n+1} = 0, \quad \text{azaz} \quad x = -\frac{n+1}{n}$$

esetén válik szingulárisá.

* * *

20. Példa. Határozzuk meg az

$$\mathbf{N} = \begin{bmatrix} 2+x & -1 & & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & \\ & & & \ddots & \\ & & & & 2 & -1 \\ & & & & -1 & 2+y \end{bmatrix}$$

n -edrendű mátrix inverzét.

Megoldás. A feladatot tekinthetjük mint a 19. Példabeli, 11 indexű és az nn indexű elemének módosításával felírt \mathbf{A} mátrix inverzét, azaz

$$\left\{ \mathbf{A} + \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_n \end{bmatrix} \begin{bmatrix} x & \\ & y \end{bmatrix} \begin{bmatrix} \mathbf{e}_1^\top \\ \mathbf{e}_n^\top \end{bmatrix} \right\}^{-1}.$$

Alkalmazva az (1.5.28) Woodbury-féle képletet, a keresett inverzet (*)

$$\mathbf{A}^{-1} - \mathbf{A}^{-1} \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_n \end{bmatrix} \left\{ \begin{bmatrix} \frac{1}{x} & \\ & \frac{1}{y} \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1^\top \\ \mathbf{e}_n^\top \end{bmatrix} \mathbf{A}^{-1} \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_n \end{bmatrix} \right\}^{-1} \begin{bmatrix} \mathbf{e}_1^\top \\ \mathbf{e}_n^\top \end{bmatrix} \mathbf{A}^{-1}$$

alakban kapjuk. Az invertálhatóság feltétele, hogy

$$\begin{vmatrix} \frac{1}{x} + \mathbf{e}_1^\top \mathbf{A}^{-1} \mathbf{e}_1 & \mathbf{e}_1^\top \mathbf{A}^{-1} \mathbf{e}_n \\ \mathbf{e}_n^\top \mathbf{A}^{-1} \mathbf{e}_1 & \frac{1}{y} + \mathbf{e}_n^\top \mathbf{A}^{-1} \mathbf{e}_n \end{vmatrix} \neq 0.$$

Behelyettesítve az \mathbf{A}^{-1} inverz megfelelő elemeit, innen

$$\begin{vmatrix} \frac{1}{x} + \frac{n}{n+1} & \frac{1}{n+1} \\ \frac{1}{n+1} & \frac{1}{y} + \frac{n}{n+1} \end{vmatrix} \equiv \left(\frac{1}{x} + \frac{n}{n+1} \right) \left(\frac{1}{y} + \frac{n}{n+1} \right) - \frac{1}{(n+1)^2} \neq 0$$

adódik, ahonnan némi algebrai átalakítás után azt kapjuk, hogy

$$(**) \quad (n-1)(x+1)(y+1) + (x+1) + (y+1) = 0$$

esetén a módosított mátrix szinguláris, tehát nem invertálható. Ha x és y nem elégíti ki ezt az egyenletet, a módosított mátrix inverze a (*) képlet szerint számítható. Ez azonban nagyon körülményes, túl sok algebrai átalakítás után vezetne csak áttekinthető eredményre, ezért helyette azt javasoljuk, hogy

a 19. Példában szereplő \mathbf{M} mátrix ismert inverzét felhasználva, annak nn indexű elemét módosítsuk y -nal és így ismét a Sherman–Morrison-képletet alkalmazzuk. Ezek szerint

$$\mathbf{N} = \mathbf{M} + y\mathbf{e}_n\mathbf{e}_n^\top,$$

és így

$$\mathbf{N}^{-1} = \mathbf{M}^{-1} - \frac{\mathbf{M}^{-1}\mathbf{e}_n\mathbf{e}_n^\top\mathbf{M}^{-1}}{\frac{1}{y} + \mathbf{e}_n^\top\mathbf{M}^{-1}\mathbf{e}_n}.$$

A 19. Példa alapján

$$\begin{aligned}\mathbf{M}^{-1}\mathbf{e}_n &= \left[\frac{i(x+1)-x}{n(x+1)+1} \right]; \quad \mathbf{e}_n^\top\mathbf{M}^{-1} = \left[\frac{j(x+1)-x}{n(x+1)+1} \right]; \\ \mathbf{e}_n^\top\mathbf{M}^{-1}\mathbf{e}_n &= \frac{n(x+1)-x}{n(x+1)+1}.\end{aligned}$$

Behelyettesítés után $i \leq j$ esetén

$$\frac{i(x+1)-x}{n(x+1)+1}(n+1-j) - \frac{\frac{i(x+1)-x}{n(x+1)+1} \frac{j(x+1)-x}{n(x+1)+1}}{\frac{1}{y} + \frac{n(x+1)-x}{n(x+1)+1}}$$

adódik, majd összevonás után a keresett inverz elemeire kapjuk, hogy

$$[\mathbf{N}^{-1}]_{ij} = \begin{cases} [i(x+1)-x] \frac{(n+1-j)(y+1)-y}{n(x+1)+1+y[n(x+1)-x]}, & i \leq j, \\ [j(x+1)-x] \frac{(n+1-i)(y+1)-y}{n(x+1)+y[n(x+1)-x]}, & i \geq j. \end{cases}$$

A kapott eredmény jól mutatja, hogy az inverz egypárú mátrix (lásd az 1.4.13 definíciót), továbbá az invertálhatóság feltétele

$$n(x+1)+1+y[n(x+1)-x] \neq 0,$$

ami megegyezik a (**) alatti feltétellel.

* * *

A másik feladat, ami az 1.5.4 tétel segítségével oldható meg, a következőképpen fogalmazható. Adott nonszinguláris n -edrendű mátrixnak ismerjük az inverzét, és meg akarjuk határozni egy r -edrendű minormátrixának inverzét. Kérdés, hogy amennyiben r csak kevéssel különbözik n -től, elérhető-e, hogy a keresett inverz meghatározásának feladatát egy r -nél alacsonyabb rendű mátrix invertálására redukáljuk? A válasz a következő tételben fogalmazható meg.

1.5.6 tétel. *Tegyük fel, hogy ismerjük egy szimmetrikusan négy blokkra particionált hipermátrix inverzét:*

$$(1.5.31) \quad \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{S} & \mathbf{T} \end{bmatrix}.$$

Az \mathbf{A} minormátrix inverzét megkapjuk, ha az adott mátrix inverzéből a komplementer minormátrixszal – vagyis a \mathbf{T} blokkal – generált hiperdiádot levonjuk:

$$(1.5.32) \quad \begin{bmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{S} & \mathbf{T} \end{bmatrix} - \begin{bmatrix} \mathbf{Q} \\ \mathbf{T} \end{bmatrix} \mathbf{T}^{-1} [\mathbf{S} \quad \mathbf{T}] = \begin{bmatrix} \mathbf{A}^{-1} & 0 \\ 0 & 0 \end{bmatrix}.$$

Bizonyítás. A négy blokkra particionált hipermátrixok inverzének (1.5.25) alakú előállításából közvetlenül adódik az állítás, ugyanis

$$\begin{aligned} & \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \\ -(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{bmatrix} - \\ & - \begin{bmatrix} -\mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \\ (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{bmatrix} (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}) [-(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} \quad (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}] = \\ & = \begin{bmatrix} \mathbf{A}^{-1} & 0 \\ 0 & 0 \end{bmatrix}. \quad \blacksquare \end{aligned}$$

Ezt az eredményt még a következőképpen is megfogalmazhatjuk: egy szimmetrikusan particionált nonszinguláris hipermátrix bal felső, ill. jobb alsó blokkjának inverzét (feltéve, hogy az létezik) megkapjuk, ha vesszük az adott hipermátrix inverzében a komplementer minormátrixhoz tartozó

$$(1.5.33) \quad \mathbf{A}^{-1} = \mathbf{P} - \mathbf{Q}\mathbf{T}^{-1}\mathbf{S},$$

illetve

$$(1.5.34) \quad \mathbf{D}^{-1} = \mathbf{T} - \mathbf{S}\mathbf{P}^{-1}\mathbf{Q}$$

Schur-komplementumot.

Ezek felhasználásával – Jacobitól származó – érdekes összefüggéseket kaphatunk egy determináns és bizonyos minorai között. Ha ugyanis az 1.5.3 tételt az (1.5.31) szerint particionált inverz mátrixra alkalmazzuk:

$$\begin{vmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{S} & \mathbf{T} \end{vmatrix} = |\mathbf{P}| \cdot |\mathbf{T} - \mathbf{S}\mathbf{P}^{-1}\mathbf{Q}| = |\mathbf{T}| \cdot |\mathbf{P} - \mathbf{Q}\mathbf{T}^{-1}\mathbf{S}|,$$

és figyelembe vesszük az (1.5.33) és (1.5.34) összefüggéseket, akkor

$$(1.5.35) \quad |\mathbf{A}| = \begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{vmatrix} \cdot |\mathbf{T}|,$$

illetve

$$(1.5.36) \quad |\mathbf{D}| = \begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{vmatrix} \cdot |\mathbf{P}|$$

adódik, azaz a mátrix egy főminorá egyenlő a mátrix determinánsának és az inverz mátrix komplementer főminorának a szorzatával (lásd [10]).

Az alábbiakban a fentiek általánosítását adjuk az ún. *perszimmetrikusan particionált hipermátrixokra*.

1.5.5 definíció. Ha az $[\mathbf{A}_{ij}]$ hipermátrix mellékátlójában álló $\mathbf{A}_{i,n+1-i}$ blokkok kvadratikus mátrixok, akkor a hipermátrixot **perszimmetrikusan particionált hipermátrixnak** nevezzük.

Tekintsünk most egy perszimmetrikusan particionált, (1.5.10) alakú másodrendű hipermátrixot, ahol \mathbf{B} p -edrendű, \mathbf{C} pedig q -adrendű kvadratikus mátrix. Tegyük fel, hogy ismerjük az inverzét, amelyet az (1.5.31) szerint ugyancsak perszimmetrikusan particionált alakban írhatunk fel. A blokkok összeszorozhatóságának feltétele azonban csak akkor teljesül, ha az inverzben \mathbf{S} p -edrendű és \mathbf{Q} q -adrendű. Ilyenkor *komplementer perszimmetrikus* particionálásról beszélünk. Sematikusan vázolva:

$$\left\{ \begin{array}{c} p \\ q \end{array} \right\} \begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{vmatrix}^{-1} = \begin{vmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{S} & \mathbf{T} \end{vmatrix} \left\{ \begin{array}{c} q \\ p \end{array} \right\}.$$

Ha most pl. a \mathbf{C} blokk inverzét kívánjuk meghatározni az inverz blokkjai segítségével, akkor az adott mátrixot az oszlopok permutációjával először egy szimmetrikusan particionált hipermátrixra transzformáljuk, és ekkor már alkalmazható az 1.5.6 tétel. Vagyis

$$\begin{aligned} (1.5.37) \quad \begin{array}{c} p \\ q \end{array} \begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{vmatrix}^{-1} &= \left\{ \begin{array}{c} \begin{vmatrix} \mathbf{B} & \mathbf{A} \\ \mathbf{D} & \mathbf{C} \end{vmatrix} \\ \begin{array}{c} \widehat{p} \\ \widehat{q} \end{array} \end{array} \cdot \begin{array}{c} \mathbf{E}_p \\ \mathbf{E}_q \end{array} \right\}^{-1} = \\ &= \begin{bmatrix} \mathbf{E}_q & \mathbf{E}_p \end{bmatrix} \cdot \begin{bmatrix} \mathbf{B}^{-1} + \mathbf{B}^{-1}\mathbf{A}(\mathbf{C} - \mathbf{D}\mathbf{B}^{-1}\mathbf{A})^{-1}\mathbf{D}\mathbf{B}^{-1} & -\mathbf{B}^{-1}\mathbf{A}(\mathbf{C} - \mathbf{D}\mathbf{B}^{-1}\mathbf{A})^{-1} \\ -(\mathbf{C} - \mathbf{D}\mathbf{B}^{-1}\mathbf{A})^{-1}\mathbf{D}\mathbf{B}^{-1} & (\mathbf{C} - \mathbf{D}\mathbf{B}^{-1}\mathbf{A})^{-1} \end{bmatrix} = \\ &= \begin{bmatrix} -(\mathbf{C} - \mathbf{D}\mathbf{B}^{-1}\mathbf{A})^{-1}\mathbf{D}\mathbf{B}^{-1} & (\mathbf{C} - \mathbf{D}\mathbf{B}^{-1}\mathbf{A})^{-1} \\ \mathbf{B}^{-1} + \mathbf{B}^{-1}\mathbf{A}(\mathbf{C} - \mathbf{D}\mathbf{B}^{-1}\mathbf{A})^{-1}\mathbf{D}\mathbf{B}^{-1} & -\mathbf{B}^{-1}\mathbf{A}(\mathbf{C} - \mathbf{D}\mathbf{B}^{-1}\mathbf{A})^{-1} \end{bmatrix} = \begin{bmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{S} & \mathbf{T} \end{bmatrix}. \end{aligned}$$

Innen a \mathbf{B} minormátrix keresett inverzére

$$(1.5.38) \quad \mathbf{B}^{-1} = \mathbf{S} - \mathbf{T}\mathbf{Q}^{-1}\mathbf{P}$$

adódik, illetve hasonló megfontolással, \mathbf{C} inverzére

$$(1.5.39) \quad \mathbf{C}^{-1} = \mathbf{Q} - \mathbf{P}\mathbf{S}^{-1}\mathbf{T}.$$

Ezt az eredményt a következő tételben fogalmazzuk meg:

1.5.7 tétel. *Egy perszimmetrikusan particionált nemszinguláris másodrendű hipermátrix bal alsó, ill. jobb felső blokkjának inverze (feltéve, hogy létezik) egyenlő az adott hipermátrix inverzének komplementer particionálásával nyert hipermátrixban a komplementer minormátrixhoz tartozó Schur-komplementummal.*

Alkalmazzuk most az 1.5.3 tételt a perszimmetrikusan particionált mátrixokra. Mivel

$$\begin{matrix} p) \\ q) \end{matrix} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{B} & \mathbf{A} \\ \mathbf{D} & \mathbf{C} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{E}_p & \\ & \mathbf{E}_q \end{bmatrix},$$

$\underbrace{\hspace{1.5cm}}_{\mathbf{P}} \quad \underbrace{\hspace{1.5cm}}_{\mathbf{Q}}$

a determinánsa

$$(1.5.40) \quad \begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{vmatrix} = (-1)^{pq} |\mathbf{C}| \cdot |\mathbf{B} - \mathbf{A}\mathbf{C}^{-1}\mathbf{D}| = (-1)^{pq} |\mathbf{B}| \cdot |\mathbf{C} - \mathbf{A}\mathbf{B}^{-1}\mathbf{D}|$$

alakban, inverzének determinánsa pedig

$$\begin{aligned} \left| \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} \right| &= \left| \begin{bmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{S} & \mathbf{T} \end{bmatrix} \right| = \left| \begin{bmatrix} \mathbf{Q} & \mathbf{P} \\ \mathbf{T} & \mathbf{S} \end{bmatrix} \right| \cdot \left| \mathbf{E}_p \quad \mathbf{E}_q \right| = \\ &= (-1)^{pq} |\mathbf{Q}| \cdot |\mathbf{S} - \mathbf{T}\mathbf{Q}^{-1}\mathbf{P}| = (-1)^{pq} |\mathbf{S}| \cdot |\mathbf{Q} - \mathbf{P}\mathbf{S}^{-1}\mathbf{T}| \end{aligned}$$

alakban írható. Figyelembe véve az (1.5.38) és (1.5.39) összefüggéseket, innen

$$(1.5.41) \quad |\mathbf{C}| = (-1)^{pq} \begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{vmatrix} \cdot |\mathbf{S}|,$$

illetve

$$(1.5.42) \quad |\mathbf{B}| = (-1)^{pq} \begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{vmatrix} \cdot |\mathbf{Q}|,$$

adódik, azaz a perszimmetrikusan particionált mátrix jobb felső (bal alsó) minora – előjeltől eltekintve – egyenlő a mátrix determinánsának és az inverz komplementer particionálásával nyert mátrix jobb felső (bal alsó) minorának a szorzatával.

Ennek a tételnek egy érdekes alkalmazásaként most bebizonyítjuk az 1.4.2 pontban megismert szimmetrikus kontinuáns mátrixok és az egypárú mátrixok között fennálló alábbi fontos kapcsolatot.

Megjegyzés. Az egypárú mátrix fogalmát Gantmacher és Krejn vezették be, és ők bizonyították be először az 1.5.8 tételt (lásd [7]). Az itt közölt bizonyítás gondolata megtalálható a [21] és [36] dolgozatokban.

1.5.8 tétel. *A nemszinguláris szimmetrikus kontinuáns mátrixok – és csakis ezen mátrixok – inverze egypárú mátrix.*

Bizonyítás. Tekintsük a

$$\mathbf{K} = \begin{bmatrix} a_1 & b_1 & 0 & \dots & 0 \\ b_1 & a_2 & b_2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_{n-1} & 0 \\ 0 & 0 & \dots & b_{n-1} & a_n \end{bmatrix}$$

szimmetrikus kontinuáns mátrixot és tegyük fel, hogy $b_i \neq 0$ (ez azt biztosítja, hogy a mátrix irreducibilis; lásd majd a 4.0.1 definíciót).

Ha ezt a mátrixot egy sorral (fent) és egy oszloppal (jobb oldalt) szegélyezzük, akkor alsó háromszögmátrixot kapunk, amelynek inverze ugyancsak alsó háromszögmátrix. Másrészt a szegélyezett mátrix egy perszimmetrikusan particionált hipermátrixnak tekinthető, amelynek bal alsó (n -edrendű) minormátrixa az invertálandó \mathbf{K} mátrix, alkalmazható tehát az (1.5.39) képlet. A kiegészített \mathbf{K} mátrix inverzének elemeire vezessük be az alábbi jelöléseket:

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ a_1 & b_1 & 0 & \dots & 0 \\ b_1 & a_2 & b_2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_{n-1} & 0 \\ 0 & 0 & \dots & b_{n-1} & a_n \end{bmatrix}^{-1} = \begin{bmatrix} u_1 & 0 & 0 & \dots & 0 \\ u_2 & l_{21} & 0 & \dots & 0 \\ u_3 & l_{31} & l_{32} & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ u_n & l_{n1} & l_{n2} & \dots & 0 \\ -u_0 & u_0 v_1 & u_0 v_2 & \dots & u_0 v_n \end{bmatrix},$$

azaz

$$(1.5.43) \quad \begin{bmatrix} \mathbf{e}_1^T & 0 \\ \mathbf{K} & \mathbf{e}_n \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{u} & \mathbf{L} \\ -u_0 & u_0 \mathbf{v}^T \end{bmatrix}.$$

Az (1.5.39) képlet alkalmazásával

$$(1.5.44) \quad \mathbf{K}^{-1} = \mathbf{L} + \mathbf{u} \mathbf{v}^T,$$

ahol $l_{ij} = 0$, ha $i \leq j$. Tehát a $\mathbf{K}^{-1} = \mathbf{R} = [r_{ij}]$ inverz mátrix elemeire $i \leq j$ esetén $r_{ij} = u_i v_j$ adódik, s mivel \mathbf{K} szimmetrikus és így \mathbf{R} is az, végül is

$$(1.5.45) \quad r_{ij} = \begin{cases} u_i v_j, & \text{ha } i \leq j, \\ v_i u_j, & \text{ha } i \geq j \end{cases}$$

írható, ami az 1.4.13 definíció szerint éppen azt mondja ki, hogy \mathbf{R} egypárú mátrix.

Az (1.5.43)–(1.5.45) összefüggésekből következik, hogy

$$(1.5.46) \quad b_i^{-1} = l_{i+1,i} = r_{i+1,i} - u_{i+1} v_i = v_{i+1} u_i - u_{i+1} v_i \neq 0 \quad (i = 1, 2, \dots, n-1). \blacksquare$$

Most bebizonyítjuk a fordított állítást: tekintsünk egy nemszinguláris, (1.5.45) alakú egypárú mátrixot, és lássuk be, hogy az inverze szimmetrikus kontinuáns mátrix. Alakítsuk át az (1.5.45) alakban megadott mátrix elemeit a következőképpen:

$$(1.5.47) \quad r_{ij} = \begin{cases} u_i v_j, & \text{ha } i \leq j, \\ u_i v_j + (v_i u_j - u_i v_j), & \text{ha } i \geq j. \end{cases}$$

Ez azt jelenti, hogy bevezetve az

$$(1.5.48) \quad l_{ij} = \begin{cases} 0, & \text{ha } i \leq j, \\ v_i u_j - u_i v_j, & \text{ha } i > j \end{cases}$$

elemekből alkotott alsó háromszögmátrixot, az \mathbf{R} egypárú mátrix felírható

$$(1.5.49) \quad \mathbf{R} = \mathbf{L} + \mathbf{u} \mathbf{v}^T$$

alakban. Ennek inverzét a következő megfontolásból nyerjük. Szegélyezzük az \mathbf{L} mátrixot balról az \mathbf{u} vektorral, alulról a \mathbf{v}^T vektorral, valamint a (-1) -gyel a bal alsó sarokban. Az így nyert mátrix az

$$(1.5.50) \quad l_{i+1,i} = v_{i+1} u_i - u_{i+1} v_i \neq 0 \quad (i = 1, 2, \dots, n-1), \quad u_1 \neq 0, \quad v_n \neq 0$$

feltételek teljesülése esetén perszimmetrikusan particionált *nemszinguláris* alsó háromszögmátrix:

$$(1.5.51) \quad \left[\begin{array}{c|ccc} u_1 & 0 & \dots & 0 & 0 \\ u_2 & l_{21} & & & \\ \vdots & \vdots & & & \\ u_n & l_{n1} & l_{n,n-1} & 0 & \\ \hline -1 & v_1 & v_{n-1} & v_n & \end{array} \right].$$

Ennek inverze ugyancsak alsó háromszögmátrix; ha ezt komplementer particionálással az

$$(1.5.52) \quad \begin{bmatrix} u_1^{-1} & 0 & \dots & 0 & \vdots & 0 \\ \hline * & l_{21}^{-1} & & & \vdots & \\ \vdots & & \ddots & & & \\ * & & & l_{n,n-1}^{-1} & \vdots & \\ * & \dots & \dots & * & \vdots & v_n^{-1} \end{bmatrix} = \begin{bmatrix} \frac{1}{u_1} \mathbf{e}_1^T & \vdots & 0 \\ \hline & \mathbf{H} & \frac{1}{v_n} \mathbf{E}_n \end{bmatrix}$$

alakban írjuk fel, akkor az (1.5.39) összefüggés felhasználásával azt kapjuk, hogy

$$(1.5.53) \quad (\mathbf{L} + \mathbf{u}\mathbf{v}^T)^{-1} = \mathbf{H},$$

ahol \mathbf{H} ún. felső Hessenberg-féle mátrix,^{*} amelyet elemeinek a következő tulajdonsága jellemez:

$$h_{ij} = 0, \quad \text{ha } j - i > 1.$$

Mivel az invertálandó egypárú mátrix szimmetrikus és így az inverze is szimmetrikus, ebből következik, hogy az inverz egyúttal alsó Hessenberg-féle mátrix, ami csakis úgy lehetséges, hogy szimmetrikus kontinuáns mátrix. ■

Érdemes megjegyezni, hogy a kontinuáns mátrix inverzének $u_i v_j$ elemei az (1.5.43) összefüggés alapján egyszerű rekurzió segítségével számíthatók. Ugyanis figyelembe véve, hogy

$$\begin{bmatrix} \mathbf{e}_1^T & 0 \\ \mathbf{K} & \mathbf{e}_n \end{bmatrix} \cdot \begin{bmatrix} \mathbf{u} & \mathbf{L} \\ -\mathbf{u}_0 & \mathbf{u}_0 \mathbf{v}^T \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{E} \end{bmatrix},$$

innen

$$(1.5.54) \quad \begin{bmatrix} \mathbf{e}_1^T & 0 \\ \mathbf{K} & \mathbf{e}_n \end{bmatrix} \cdot \begin{bmatrix} \mathbf{u} \\ -u_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

és ebből közvetlenül adódik az

$$(1.5.55) \quad \begin{aligned} u_1 &= 1, \quad u_2 = -\frac{1}{b_1} a_1, \\ u_{i+1} &= -\frac{1}{b_i} (a_i u_i + b_{i-1} u_{i-1}) \quad (i = 2, 3, \dots, n-1), \\ u_0 &= a_n u_n + b_{n-1} u_{n-1} \end{aligned}$$

^{*}G. Hessenberg (1874–1925) német matematikus.

alakú rekurziós összefüggés. Hasonlóképpen, $\begin{bmatrix} \mathbf{u} & \mathbf{L} \\ -u_0 & u_0 \mathbf{v}^\top \end{bmatrix} \cdot \begin{bmatrix} \mathbf{e}_1^\top & 0 \\ \mathbf{K} & \mathbf{e}_n \end{bmatrix} =$
 $= \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix}$, innen az

$$(1.5.56) \quad \begin{bmatrix} -u_0 & u_0 \mathbf{v}^\top \end{bmatrix} \begin{bmatrix} \mathbf{e}_1^\top & 0 \\ \mathbf{K} & \mathbf{e}_n \end{bmatrix} = \begin{bmatrix} \mathbf{0} & 1 \end{bmatrix}$$

összefüggésből a

$$(1.5.57) \quad \begin{aligned} v_n &= \frac{1}{u_0}, & v_{n-1} &= -\frac{1}{b_{n-1}} a_n v_n, \\ v_{n-i} &= \frac{1}{b_{n-i}} (a_{n+1-i} v_{n+1-i} + b_{n+1-i} v_{n+2-i}) \quad (i = 2, 3, \dots, n-1) \end{aligned}$$

rekurziós összefüggés adódik.

Meg kell jegyezni még, hogy a \mathbf{K} kontinuáns mátrix determinánsa az (1.5.43) összefüggésből az (1.5.41) képlet szerint a következőképpen számítható:

$$(1.5.58) \quad |\mathbf{K}| = (-1)^{n+1} u_0 b_1 b_2 \dots b_{n-1}.$$

Innen kiolvasható, hogy a \mathbf{K} kontinuáns mátrix akkor és csak akkor nonszinguláris, ha az (1.5.55) rekurzióval számított $u_0 \neq 0$ (azt már kezdetben feltettük, hogy valamennyi $b_i \neq 0$).

Végül, az adott \mathbf{R} egypárú mátrix determinánsát az (1.5.41) képlet alkalmazásával kapjuk, figyelembe véve, hogy az (1.5.51) mátrix inverze az (1.5.52) háromszögmátrix:

$$-1 = (-1)^n u_1 \prod_{i=2}^n l_{i,i-1} \cdot v_n \cdot |\mathbf{H}|.$$

Innen (1.5.53) és (1.5.48) behelyettesítésével

$$(1.5.59) \quad |\mathbf{R}| = (-1)^{n-1} u_1 \prod_{i=1}^{n-1} \begin{vmatrix} u_i & v_i \\ u_{i+1} & v_{i+1} \end{vmatrix} \cdot v_n$$

adódik, ahonnan az is kiolvasható, hogy ha az (1.5.50) feltételek teljesülnek, akkor az egypárú mátrix nonszinguláris. Az inverz mátrix elemeit az (1.5.46) összefüggések alapján

$$(1.5.60) \quad b_i = \frac{1}{\begin{vmatrix} u_i & v_i \\ u_{i+1} & v_{i+1} \end{vmatrix}} \quad (i = 1, 2, \dots, n-1)$$

adja (ezek a kontinuáns mátrix kodiagonális elemei), illetve a főátló elemeit

az (1.5.54) egyenletrendszerből az alábbi rekurzióval kaphatjuk:

$$(1.5.61) \quad a_1 = -\frac{1}{u_1}b_1u_2,$$

$$(1.5.62) \quad a_i = -\frac{1}{u_i}(b_{i-1}u_{i-1} + b_iu_{i+1}) \quad (i = 2, 3, \dots, n-1),$$

és végül (1.5.56) n -edik egyenletéből

$$(1.5.63) \quad a_n = -\frac{1}{v_n}b_{n-1}v_{n-1}.$$

21. Példa. Határozzuk meg a

$$\mathbf{K} = \begin{bmatrix} 2 & 2 & & & \\ 2 & 3 & 1 & & \\ & 1 & 3 & 2 & \\ & & 2 & 6 & 3 \\ & & & 3 & 4 \end{bmatrix}$$

szimmetrikus kontinuáns mátrix inverzét és determinánsát.

Megoldás. Az inverz mátrix (1.5.45) alakú egypárú mátrix, amelynek az elemeit (1.5.55) és (1.5.57) rekurziós összefüggések segítségével számíthatjuk:

$$\begin{aligned} u_1 &= 1, \\ u_2 &= -\frac{a_1}{b_1} = -1, \\ u_3 &= -\frac{1}{b_2}(a_2u_2 + b_1u_1) = 1, \\ u_4 &= -\frac{1}{b_3}(a_3u_3 + b_2u_2) = -1, \\ u_5 &= -\frac{1}{b_4}(a_4u_4 + b_3u_3) = \frac{4}{3}, \\ u_0 &= a_5u_5 + b_4u_4 = \frac{7}{3}; \\ v_5 &= \frac{1}{u_0} = \frac{3}{7}, \\ v_4 &= -\frac{1}{b_4}a_5v_5 = -\frac{4}{7}, \\ v_3 &= -\frac{1}{b_3}(a_4v_4 + b_4v_5) = \frac{15}{14}, \\ v_2 &= -\frac{1}{b_2}(a_3v_3 + b_3v_4) = -\frac{29}{14}, \end{aligned}$$

$$v_1 = -\frac{1}{b_1}(a_2v_2 + b_2v_3) = \frac{36}{14}.$$

A keresett inverzet legegyszerűbben úgy írhatjuk fel, hogy az

$$\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \\ \frac{4}{3} \end{bmatrix} \quad \text{és} \quad \mathbf{v}^T = [36 \quad -29 \quad 15 \quad -8 \quad 6] \frac{1}{14}$$

vektorokból alkotott $\mathbf{u}\mathbf{v}^T$ diádot a főátló mentén „elvágjuk“ és a főátló feletti elemeket tükrözzük a főátlóra:

$$\mathbf{K}^{-1} = \frac{1}{14} \begin{bmatrix} 36 & -29 & 15 & -8 & 6 \\ -29 & 29 & -15 & 8 & -6 \\ 15 & -15 & 15 & -8 & 6 \\ -8 & 8 & -8 & 8 & -6 \\ 6 & -6 & 6 & -6 & 8 \end{bmatrix}.$$

A \mathbf{K} mátrix determinánsát az (1.5.58) képlet szerint számítjuk ki:

$$|\mathbf{K}| = (-1)^{5+1}u_0b_1b_2b_3b_4 = 28.$$

* * *

22. Példa. A 21. Példa megoldásának ismeretében határozzuk meg a \mathbf{K} mátrix negyedrendű bal felső minormátrixának az inverzét!

Megoldás. Az 1.5.7 tétel segítségével, az (1.5.33) képlet alkalmazásával oldjuk meg a feladatot. Itt az invertálandó minormátrix

$$\mathbf{A} = \begin{bmatrix} 2 & 2 & & \\ 2 & 3 & 1 & \\ & 1 & 3 & 2 \\ & & 2 & 6 \end{bmatrix}.$$

A \mathbf{K}^{-1} inverz mátrix komplementer minormátrixához tartozó Schur-komplementum ebben az esetben

$$\begin{aligned}
\mathbf{P} - \mathbf{Q}\mathbf{T}^{-1}\mathbf{S} &= \\
&= \frac{1}{14} \begin{bmatrix} 36 & 29 & 15 & -8 \\ -29 & 29 & -15 & 8 \\ 15 & -15 & 15 & -8 \\ -8 & 8 & -8 & 8 \end{bmatrix} - \frac{1}{14} \begin{bmatrix} 6 \\ -6 \\ 6 \\ -6 \end{bmatrix} \frac{14}{8} [6 \quad -6 \quad 6 \quad -6] \frac{1}{14} = \\
&= \frac{1}{14} \left\{ \begin{bmatrix} 36 & -29 & 15 & -8 \\ -29 & 29 & -15 & 8 \\ 15 & -15 & 15 & -8 \\ -8 & 8 & -8 & 8 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 9 & -9 & 9 & -9 \\ -9 & 9 & -9 & 9 \\ 9 & -9 & 9 & -9 \\ -9 & 9 & -9 & 9 \end{bmatrix} \right\} = \\
&= \frac{1}{28} \begin{bmatrix} 63 & -49 & 21 & -7 \\ -49 & 49 & -21 & 7 \\ 21 & -21 & 21 & -7 \\ -7 & 7 & -7 & 7 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 9 & -7 & 3 & -1 \\ -7 & 7 & -3 & 1 \\ 3 & -3 & 3 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix}. \\
&\quad \quad \quad * \quad * \quad *
\end{aligned}$$

23. Példa. Határozzuk meg az alábbi, Toeplitz-típusú mátrix inverzét (lásd pl. [23]):

$$\mathbf{R} = \begin{bmatrix} 1 & x & x^2 & \dots & x^{n-1} \\ x & 1 & x & \dots & x^{n-2} \\ x^2 & x & 1 & \dots & x^{n-3} \\ \dots & \dots & \dots & \dots & \dots \\ x^{n-1} & \dots & \dots & \dots & 1 \end{bmatrix}.$$

Megoldás. Vegyük észre, hogy az adott mátrix egypárú, ugyanis általános eleme (1.5.45) alakban írható:

$$r_{ij} = \begin{cases} x^{1-i}x^{j-1}, & \text{ha } i \leq j, \\ x^{1-j}x^{i-1}, & \text{ha } i \geq j, \end{cases}$$

azaz

$$u_i = \frac{1}{x^{i-1}}, \quad v_j = x^{j-1} \quad (i, j = 1, \dots, n).$$

Az inverze szimmetrikus kontinúans mátrix. Ennek elemeire az (1.5.60), (1.5.61), (1.5.62) és (1.5.63) képletek alapján, figyelembe véve az

$$\frac{u_{i-1}}{u_i} = x \quad (i = 2, \dots, n-1) \quad \text{és} \quad \frac{u_{i-1}}{u_i} = \frac{1}{x} \quad (i = 1, \dots, n)$$

összefüggéseket,

$$b_i = \frac{1}{\begin{vmatrix} x^{1-i} & x^{i-1} \\ x^{1-i-1} & x^i \end{vmatrix}} = \frac{1}{x - \frac{1}{x}} \quad (i = 1, 2, \dots, n-1),$$

$$a_1 = -\frac{\frac{1}{x}}{x - \frac{1}{x}},$$

$$a_i = -\frac{1}{x - \frac{1}{x}} \cdot \left(x + \frac{1}{x}\right) \quad (i = 2, 3, \dots, n-1),$$

$$a_n = -\frac{1}{x - \frac{1}{x}} \cdot \frac{x^{n-2}}{x^{n-1}} = -\frac{1}{x - \frac{1}{x}} \cdot \frac{1}{x}$$

adódik, és ezzel a keresett inverz:

$$\mathbf{R}^{-1} = \frac{1}{\frac{1}{x} - x} \begin{bmatrix} \frac{1}{x} & -1 & 0 & \dots & 0 \\ -1 & x + \frac{1}{x} & -1 & \dots & 0 \\ 0 & -1 & x + \frac{1}{x} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & x + \frac{1}{x} & -1 \\ 0 & 0 & \dots & -1 & \frac{1}{x} \end{bmatrix}.$$

Az adott mátrix determinánsát az (1.5.59) képlet alapján

$$|\mathbf{R}| = (-1)^{n-1} \prod_{i=1}^{n-1} \begin{vmatrix} x^{1-i} & x^{i-1} \\ x^{-i} & x^i \end{vmatrix} \cdot x^{n-1} = (1 - x^2)^{n-1}$$

alakban kapjuk.

* * *

Ez a példa is mutatja, hogy az (1.5.50) feltétel nem szükséges feltétele annak, hogy az egypárú mátrix nonszinguláris legyen. Ugyanis $x = 0$ esetén az \mathbf{A} mátrix nonszinguláris egypárú mátrix, noha $v_n = x^{n-1} = 0$.

1.6 PROJEKTOROK

A mátrixelméletben jelentős szerepe van azoknak a mátrixoknak, amelyek hatványozás során nem változnak. Ebben a szakaszban megismerkedünk ezek főbb tulajdonságaival. Mint a későbbiekben látni fogjuk, algebrai, geometriai és analízisbeli alkalmazásaik egyaránt központi helyet foglalnak el, ezért külön szakaszban foglalkozunk velük. Bevezetjük a biortogonális vektorrendszer fogalmát és néhány fontos tételt ismertetünk, majd megmutatjuk, hogy a projektorok segítségével hogyan általánosítható az inverz mátrix fogalma tetszőleges mátrixokra.

1.6.1 Projektorokra vonatkozó tételek

1.6.1 definíció. A

$$(1.6.1) \quad \mathbf{P}^2 = \mathbf{P}$$

összefüggést kielégítő mátrixokat **idempotens mátrixoknak**, **projektor-mátrixoknak** vagy röviden **projektoroknak** nevezzük.

24. Példa. Figyelembe véve, hogy a számok közül a 0 és az 1 idempotens, konstruáljunk segítségükkel olyan diagonálmátrixokat, amelyek idempotensek (projektorok).

Megoldás. Ilyenek például az

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}; \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}; \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

mátrixok.

* * *

25. Példa. Lássuk be, hogy az alábbi mátrix projektor:

$$\mathbf{P} = \begin{bmatrix} 6 & 4 & 10 \\ 15 & 10 & 25 \\ -9 & -6 & -15 \end{bmatrix}.$$

Megoldás. A négyzetre emelést az alábbi séma szerint elvégezve

$$\begin{bmatrix} \mathbf{P} \\ [\mathbf{P}] [\mathbf{P}^2] \end{bmatrix},$$

valóban

$$\begin{bmatrix} 6 & 4 & 10 \\ 15 & 10 & 25 \\ -9 & -6 & -15 \end{bmatrix} \begin{bmatrix} 6 & 4 & 10 \\ 15 & 10 & 25 \\ -9 & -6 & -15 \end{bmatrix} = \begin{bmatrix} 6 & 4 & 10 \\ 15 & 10 & 25 \\ -9 & -6 & -15 \end{bmatrix},$$

azaz $\mathbf{P}^2 = \mathbf{P}$ adódik.

* * *

Könnyű meggyőződni arról, hogy ha egy \mathbf{v}^T sorvektor és egy \mathbf{u} oszlopvektor skaláris szorzata 1, akkor ezek $\mathbf{u}\mathbf{v}^T$ alakú diadikus szorzata 1 rangú projektor. A szorzás asszociativitása miatt ugyanis

$$(\mathbf{u}\mathbf{v}^T)^2 = \mathbf{u}(\mathbf{v}^T\mathbf{u})\mathbf{v}^T = \mathbf{u}\mathbf{v}^T, \quad \text{hiszen} \quad \mathbf{v}^T\mathbf{u} = 1.$$

Az alábbiakban bebizonyítunk a projektorokra vonatkozó néhány tételt.

1.6.1 tétel. *Az egyetlen nemszinguláris projektor az egységmátrix.*

Bizonyítás. A projektorok definíciója szerint

$$(1.6.2) \quad \mathbf{P}^2 - \mathbf{P} = \mathbf{P}(\mathbf{P} - \mathbf{E}) = \mathbf{0}.$$

Ha feltételezzük, hogy \mathbf{P} nemszinguláris, akkor invertálható; szorozzuk meg az inverzével az (1.6.2) egyenletet balról:

$$\mathbf{P}^{-1}\mathbf{P}(\mathbf{P} - \mathbf{E}) = \mathbf{0},$$

innen már $\mathbf{P} = \mathbf{E}$ adódik. ■

1.6.2 tétel. *Ha \mathbf{P} egy n -edrendű projektor, amelynek rangja r , akkor $\mathbf{E} - \mathbf{P}$ is projektor és rangja $n - r$.*

Bizonyítás. Hatványozással meggyőződhetünk arról, hogy $\mathbf{E} - \mathbf{P}$ is projektor; ugyanis

$$(\mathbf{E} - \mathbf{P})^2 = \mathbf{E} - 2\mathbf{P} + \mathbf{P}^2 = \mathbf{E} - 2\mathbf{P} + \mathbf{P} = \mathbf{E} - \mathbf{P}.$$

Az $\mathbf{E} - \mathbf{P}$ projektor rangját a két mátrix összegének rangjára vonatkozó 1.3.5 tétel segítségével alulról, a két kvadratikus mátrix zérus szorzatára vonatkozó 1.3.13 tétel felhasználásával felülről becsüljük. Ugyanis

$$\mathbf{P} + (\mathbf{E} - \mathbf{P}) = \mathbf{E},$$

ezért az 1.3.5 tétel szerint

$$(1.6.3) \quad \varrho(\mathbf{P}) + \varrho(\mathbf{E} - \mathbf{P}) \geq \varrho(\mathbf{E}) = n.$$

Másrészt (1.6.1) alapján $\mathbf{P}(\mathbf{E} - \mathbf{P}) = \mathbf{0}$, ezért az 1.3.13 tétel szerint

$$(1.6.4) \quad \varrho(\mathbf{P}) + \varrho(\mathbf{E} - \mathbf{P}) \leq n.$$

Az (1.6.3) és (1.6.4) egyenlőtlenségek egyidejűleg csak úgy állhatnak fenn, ha mindkettőben az egyenlőség teljesül, így

$$(1.6.5) \quad \varrho(\mathbf{E} - \mathbf{P}) = n - \varrho(\mathbf{P}) = n - r,$$

és ezzel a tételt bebizonyítottuk. ■

Megjegyzés. Az $\mathbf{E} - \mathbf{P}$ projektort *komplementer projektornak* nevezzük.

Mielőtt rátérnénk a következő tételre, be kell vezetnünk az ún. *biortogonális vektorrendszer* fogalmát.

1.6.2 definíció. Ha adott n elemű $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$, valamint $\mathbf{v}_1^\top, \mathbf{v}_2^\top, \dots, \mathbf{v}_r^\top$ vektorokra fennáll a

$$(1.6.6) \quad \mathbf{v}_k^\top \mathbf{u}_l = \delta_{kl} \quad (k, l = 1, 2, \dots, r)$$

összefüggés, akkor azt mondjuk, hogy az \mathbf{u}_l és \mathbf{v}_k^\top vektorok **biortogonális vektorrendszert** alkotnak. Ha $r = n$, akkor a vektorrendszert **teljes biortogonális vektorrendszernek** nevezzük.

Ha a teljes biortogonális vektorrendszerben az $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ vektorokat egy \mathbf{U} mátrix oszlopainak, a $\mathbf{v}_1^\top, \mathbf{v}_2^\top, \dots, \mathbf{v}_n^\top$ vektorokat pedig egy \mathbf{V}^\top mátrix sorainak tekintjük, akkor az (1.6.6) definiáló összefüggés alapján fennáll, hogy

$$(1.6.7) \quad \mathbf{V}^\top \mathbf{U} = \mathbf{E},$$

ahol \mathbf{E} az n -edrendű egységmátrix. Mivel \mathbf{U} és \mathbf{V}^\top kvadratikus és (1.6.7) következtében determinánsuk zérustól különböző, ezért invertálhatóak, és

$$\mathbf{V}^\top = \mathbf{U}^{-1}.$$

E tulajdonságuk alapján az $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ és $\mathbf{v}_1^\top, \mathbf{v}_2^\top, \dots, \mathbf{v}_n^\top$ vektorrendszereket *reciprok vektorrendszereknek* is nevezzük.

A következő tétel projektorok minimális diadikus előállítására vonatkozik.

1.6.3 tétel (Egerváry tétele).* Ha egy r -edrangú projektort minimális számú diád összegeként írunk fel, akkor e diádok oszlop-, ill. sorvektorai biortogonális vektorrendszert alkotnak.

Bizonyítás. Legyen a \mathbf{P} projektormátrix egy minimális diadikus előállítása

$$\mathbf{P} = \sum_{k=1}^r \mathbf{u}_k \mathbf{v}_k^\top = \mathbf{U} \mathbf{V}^\top.$$

Behelyettesítve ezt az (1.6.2) egyenletbe, $\mathbf{U} \mathbf{V}^\top \mathbf{U} \mathbf{V}^\top - \mathbf{U} \mathbf{V}^\top = \mathbf{0}$, majd kiemeléssel $\mathbf{U}(\mathbf{V}^\top \mathbf{U} - \mathbf{E}_r) \mathbf{V}^\top = \mathbf{0}$ adódik. Mivel \mathbf{U} oszlopai, ill. \mathbf{V}^\top sorai a minimális diadikus előállításra vonatkozó 1.3.4 tétel értelmében lineárisan függetlenek, ezért a rangra vonatkozó 1.3.10 tétel alapján $-\mathbf{C} = (\mathbf{V}^\top \mathbf{U} - \mathbf{E}_r)$ választásával – ebből következik, hogy

$$(1.6.8) \quad \mathbf{V}^\top \mathbf{U} = \mathbf{E}_r,$$

és ezzel a tételt bebizonyítottuk.

Megjegyzés. A tétel jelentősége az, hogy a projektorokat nem is lehet másképpen minimális számú diád összegére bontani, mint úgy, hogy ezek oszlop-, ill. sorvektorai biortogonális vektorrendszert alkossanak (lásd [24], [26]).

*Egerváry Jenő (1891–1958) magyar matematikus.

26. Példa. Tekintsük a

$$\mathbf{P} = \begin{bmatrix} -5 & -4 & -10 \\ -15 & -9 & -25 \\ 9 & 6 & 16 \end{bmatrix}$$

projektort és készítsük el olyan faktorizációját, amelyben az első tényező oszlopvektorai és a második tényező sorvektorai biortogonális vektorrendszert alkotnak.

Megoldás. A $\mathbf{P}\mathbf{P}$ szorzás elvégzésével meggyőződhetünk arról, hogy $\mathbf{P}^2 = \mathbf{P}$, tehát \mathbf{P} valóban projektor. Készítsük el \mathbf{P} egy minimális diadikus felbontását:

$$\begin{aligned} \mathbf{P} - \frac{1}{-5} \begin{bmatrix} -5 \\ -15 \\ 9 \end{bmatrix} [-5 \quad -4 \quad -10] &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 5 \\ 0 & -1,2 & -2 \end{bmatrix} = {}^{(2)}\mathbf{P}, \\ {}^{(2)}\mathbf{P} - \frac{1}{3} \begin{bmatrix} 0 \\ 3 \\ -1,2 \end{bmatrix} [0 \quad 3 \quad 5] &= \mathbf{0}; \end{aligned}$$

tehát a kapott minimális diadikus felbontás:

$$\mathbf{P} = -\frac{1}{5} \begin{bmatrix} -5 \\ -15 \\ 9 \end{bmatrix} [-5 \quad -4 \quad -10] + \frac{1}{3} \begin{bmatrix} 0 \\ 3 \\ -1,2 \end{bmatrix} [0 \quad 3 \quad 5].$$

Az együtthatókkal beszorozva az egyes sorvektorokat, és a diádösszegeket mátrixszorzat alakjában felírva:

$$\mathbf{P} = \begin{bmatrix} -5 & 0 \\ -15 & 3 \\ 9 & -1,2 \end{bmatrix} \begin{bmatrix} 1 & 0,8 & 2 \\ 0 & 1 & \frac{5}{3} \end{bmatrix} = \mathbf{U}\mathbf{V}^T.$$

A diádok oszlop-, ill. sorvektorainak biortogonalitását igazolhatjuk, ha képezzük az $\mathbf{U}\mathbf{V}^T$ alakú faktorizáció tényezőinek fordított sorrendben vett szorzatát:

$$\mathbf{V}^T\mathbf{U} = \begin{bmatrix} 1 & 0,8 & 2 \\ 0 & 1 & \frac{5}{3} \end{bmatrix} \begin{bmatrix} -5 & 0 \\ -15 & 3 \\ 9 & -1,2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

* * *

Mielőtt a következő tételt megfogalmaznánk, bevezetünk néhány új fogalmat.

1.6.3 definíció. Egy \mathbf{P} mátrixot **hermitikus projektornak** nevezünk, ha egyidejűleg teljesül a

$$(1.6.9) \quad \mathbf{P}^2 = \mathbf{P} \quad \text{és} \quad \mathbf{P}^H = \mathbf{P}$$

összefüggés. Egy **diádot hermitikus diádnak** nevezünk, ha $\mathbf{u}\mathbf{u}^H$ alakban állítható elő. Ha az $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ vektorokra teljesül az

$$\mathbf{u}_i^H \mathbf{u}_j = \delta_{ij} \quad (i, j = 1, 2, \dots, r)$$

feltétel, akkor azt mondjuk, hogy **unitér vektorrendszert** alkotnak. Az unitér vektorrendszert $r = n$ esetén **teljesnek**, $r < n$ esetén **nemteljesnek** nevezzük.

1.6.4 definíció. Egy kvadratikus mátrix főátlójában álló elemek összegét a mátrix **spurjának (nyomának)** nevezzük; jele $\text{Sp } \mathbf{A}^*$. Tehát

$$\text{Sp } \mathbf{A} = \sum_{k=1}^n a_{kk}.$$

A mátrixszorzás definíciójából következik, hogy két mátrix szorzatának a nyoma nem függ a tényezők sorrendjétől, ugyanis

$$\text{Sp}(\mathbf{AB}) = \sum_{i=1}^n \sum_{k=1}^n a_{ik} b_{ki} = \sum_{k=1}^n \sum_{i=1}^n b_{ki} a_{ik} = \text{Sp}(\mathbf{BA}).$$

Ebből következik, hogy pl. három tényezős szorzat esetén

$$\text{Sp}(\mathbf{A} \cdot \mathbf{BC}) = \text{Sp}(\mathbf{BC} \cdot \mathbf{A}), \quad \text{illetve} \quad \text{Sp}(\mathbf{AB} \cdot \mathbf{C}) = \text{Sp}(\mathbf{C} \cdot \mathbf{AB}),$$

vagyis szorzatmátrix nyoma a tényezők sorrendjének ciklikus permutációjával szemben invariáns. A $\mathbf{B} = \mathbf{A}^H$ speciális esetben ez azt jelenti, hogy egy mátrixnak és transzponált konjugáltjának a szorzata olyan mátrixot ad, amelynek a spurja az adott mátrix elemei abszolút értékének négyzetösszege:

$$(1.6.10) \quad \text{Sp}(\mathbf{AA}^H) = \sum_{j=1}^n \sum_{i=1}^n |a_{ij}|^2.$$

Innen adódik hermitikus \mathbf{A} mátrix esetén

$$(1.6.11) \quad \text{Sp}(\mathbf{A}^2) = \sum_{j=1}^n \sum_{i=1}^n |a_{ij}|^2, \quad \text{ha} \quad \mathbf{A} = \mathbf{A}^H.$$

* Az angol nyelvű irodalomban a $\text{tr } \mathbf{A}$ jelölést használják, ami a trace (nyom) elnevezésre utal.

Ennek felhasználásával bizonyíthatjuk a következő tételt.

1.6.4 tétel. *Egy r -edrangu hermitikus projektor mindig felbontható r számú hermitikus diád összegére, ahol a diádok oszlopvektorai nemteljes unitér vektorrendszert alkotnak.*

Bizonyítás. Legyen $\mathbf{P} = [a_{ij}]$ ($a_{ij} = \overline{a_{ji}}$) egy n -edrendű és r -edrangu hermitikus projektor. Ennek főátlójában biztosan van legalább egy pozitív elem. Ugyanis $\mathbf{P}^2 = \mathbf{P}$ miatt $\text{Sp}(\mathbf{P}) = \text{Sp}(\mathbf{P}^2)$. Mivel \mathbf{P} projektor, így

$$a_{ii} = \sum_{k=1}^n a_{ik} a_{ki},$$

és mivel \mathbf{P} hermitikus is, ezért $a_{ki} = \overline{a_{ik}}$, vagyis

$$a_{ii} = \sum_{k=1}^n a_{ik} \overline{a_{ik}} = \sum_{k=1}^n |a_{ik}|^2.$$

Tehát a mátrix nyomára az

$$\text{Sp}(\mathbf{P}) = \sum_{i=1}^n a_{ii} = \sum_{i=1}^n \sum_{k=1}^n |a_{ik}|^2 > 0$$

egyenlőtlenség adódik (hacsak \mathbf{P} nem zérusmátrix). Vagyis \mathbf{P} főátlójában van legalább egy pozitív elem. Az általánosság megszorítása nélkül feltehetjük, hogy $a_{11} > 0$. Vonjuk le a \mathbf{P} hermitikus projektorból az a_{11} elemmel generált hermitikus diádot:

$$\mathbf{P} - \frac{\mathbf{P} \mathbf{e}_1 \mathbf{e}_1^T \mathbf{P}}{\mathbf{e}_1^T \mathbf{P} \mathbf{e}_1} = \mathbf{P} - \frac{1}{\sqrt{a_{11}}} \mathbf{a}_1 \mathbf{a}_1^H \frac{1}{\sqrt{a_{11}}} = \stackrel{(2)}{\mathbf{P}}.$$

Mivel két hermitikus mátrix különbsége is hermitikus, csak azt kell megmutatni, hogy $\stackrel{(2)}{\mathbf{P}}$ is projektor. E célból képezzük $\stackrel{(2)}{\mathbf{P}}$ négyzetét:

$$\begin{aligned} \left(\mathbf{P} - \frac{1}{\sqrt{a_{11}}} \mathbf{a}_1 \mathbf{a}_1^H \frac{1}{\sqrt{a_{11}}} \right)^2 &= \\ &= \mathbf{P}^2 - \frac{1}{\sqrt{a_{11}}} \mathbf{P} \mathbf{a}_1 \mathbf{a}_1^H \frac{1}{\sqrt{a_{11}}} - \frac{1}{\sqrt{a_{11}}} \mathbf{a}_1 \mathbf{a}_1^H \mathbf{P} \frac{1}{\sqrt{a_{11}}} + \frac{\mathbf{a}_1}{\sqrt{a_{11}}} \frac{\mathbf{a}_1^H}{\sqrt{a_{11}}} \frac{\mathbf{a}_1}{\sqrt{a_{11}}} \frac{\mathbf{a}_1^H}{\sqrt{a_{11}}}. \end{aligned}$$

A projektorok definíciójából következik, hogy \mathbf{P}^2 és \mathbf{P} első oszlopa, ill. első sora megegyezik, tehát

$$(1.6.12) \quad \mathbf{P} \mathbf{a}_1 = \mathbf{a}_1 \quad \text{és} \quad \mathbf{a}_1^H \mathbf{P} = \mathbf{a}_1^H,$$

továbbá a bal felső sarokelemek egyenlősége miatt

$$\mathbf{a}_1^H \mathbf{a}_1 = a_{11}.$$

A kapott összefüggéseket behelyettesítve és az összevonást elvégezve,

$$\left(\mathbf{P} - \frac{\mathbf{a}_1}{\sqrt{a_{11}}} \frac{\mathbf{a}_1^H}{\sqrt{a_{11}}} \right)^2 = \mathbf{P} - \frac{\mathbf{a}_1}{\sqrt{a_{11}}} \frac{\mathbf{a}_1^H}{\sqrt{a_{11}}}$$

adódik, tehát \mathbf{P} valóban hermitikus projektor. Az eljárást folytatva, r lépés után a zérusmátrixhoz jutunk. Ezzel az r -edrangú hermitikus protektort r hermitikus diád összegeként állítottuk elő. Jelölje \mathbf{u}_k az egymás után levonásra kerülő diádok oszlopvektorait; ekkor

$$\mathbf{P} = \sum_{k=1}^r \mathbf{u}_k \mathbf{u}_k^H = \mathbf{U} \mathbf{U}^H,$$

ahol az 1.6.3 tétel alapján $\mathbf{U}^H \mathbf{U} = \mathbf{E}_r$. Tehát a diádok oszlopvektorai nemteljes unitér vektorrendszer alkotnak. ■

27. Példa. Bontsuk fel a

$$\mathbf{P} = \begin{bmatrix} \frac{160}{169} & \frac{12i}{169} & \frac{-36}{169} \\ \frac{-12i}{169} & \frac{153}{169} & \frac{-48i}{169} \\ \frac{-36}{169} & \frac{48i}{169} & \frac{25}{169} \end{bmatrix}$$

hermitikus projektort hermitikus diádok összegére.

Megoldás. A mátrix alakjából látható, hogy hermitikus, beszorzással pedig közvetlenül meggyőződhetünk arról, hogy projektor. Készítsük el egy minimális diadikus előállítását. Az egyszerűbb számolás érdekében generáló elemként célszerű a jobb alsó sarokelemet választanunk:

$$\mathbf{P} - \frac{1}{\frac{5}{13}} \begin{bmatrix} -\frac{36}{169} \\ -\frac{48i}{169} \\ \frac{25}{169} \end{bmatrix} \begin{bmatrix} -\frac{36}{169} & \frac{48i}{169} & \frac{25}{169} \end{bmatrix} \frac{1}{\frac{5}{13}} =$$

$$= \begin{bmatrix} \frac{2704}{169 \cdot 25} & \frac{2028i}{169 \cdot 25} & 0 \\ -\frac{2028i}{169 \cdot 25} & \frac{1521}{169 \cdot 25} & 0 \\ 0 & 0 & 0 \end{bmatrix} = \stackrel{(2)}{\mathbf{P}}.$$

Következő lépésben a bal felső sarokelemet választjuk generáló elemként:

$$\stackrel{(2)}{\mathbf{P}} - \frac{1}{\frac{52}{13 \cdot 5}} \begin{bmatrix} \frac{2704}{169 \cdot 25} \\ -\frac{2028i}{169 \cdot 25} \\ 0 \end{bmatrix} \begin{bmatrix} \frac{2704}{169 \cdot 25} & \frac{2028i}{169 \cdot 25} & 0 \end{bmatrix} \frac{1}{\frac{52}{13 \cdot 5}} = \mathbf{0}.$$

A diadikus felbontásra tehát a következő adódik:

$$\mathbf{P} = \begin{bmatrix} -\frac{36}{65} & \frac{52}{65} \\ -\frac{48}{65}i & -\frac{39}{65}i \\ \frac{5}{13} & 0 \end{bmatrix} \begin{bmatrix} -\frac{36}{65} & \frac{48}{65}i & \frac{5}{13} \\ \frac{52}{65} & \frac{39}{65}i & 0 \end{bmatrix} = \mathbf{U}\mathbf{U}^H.$$

A diádok oszlop-, ill. sorvektorai unitér vektorrendszert alkotnak, amiről közvetlenül meggyőződhetünk a fenti faktorizáció tényezőinek fordított sorrendű összeszorzásával:

$$\begin{bmatrix} -\frac{36}{65} & \frac{48}{65}i & \frac{5}{13} \\ \frac{52}{65} & \frac{39}{65}i & 0 \end{bmatrix} \begin{bmatrix} -\frac{36}{65} & \frac{52}{65} \\ -\frac{48}{65}i & -\frac{39}{65}i \\ \frac{5}{13} & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

* * *

1.6.2 Mátrixok általánosított inverze

A következőkben megmutatjuk, hogy projektorok segítségével hogyan általánosítható az inverz mátrix fogalma. Az 1.2.3 tétel szerint nonszinguláris kvadratikus \mathbf{A} mátrix \mathbf{X} inverzét egyértelműen meghatározza az $\mathbf{AX} = \mathbf{E}$, ill. $\mathbf{XA} = \mathbf{E}$ összefüggés. Mivel az egységmátrix (az egyetlen) n -edrangú

projektor, azt is mondhatjuk, hogy bármely n -edrangú mátrixnak és az inverzének (tetszőleges sorrendben vett) szorzata az n -edrangú projektort adja. Ennek alapján indokolható az a követelmény, hogy egy r -edrangú \mathbf{A} mátrix \mathbf{X} általánosított inverzét úgy definiáljuk, hogy az \mathbf{AX} és az \mathbf{XA} szorzat r -edrangú projektort adjon. Mint látni fogjuk, az inverz fogalma többféleképpen általánosítható. Ezek pontos definíciójához a következő mátrixegyenleteket használjuk fel:

$$(1.6.13) \quad \mathbf{AXA} = \mathbf{A},$$

$$(1.6.14) \quad \mathbf{XAX} = \mathbf{X},$$

$$(1.6.15) \quad (\mathbf{AX})^H = \mathbf{AX},$$

$$(1.6.16) \quad (\mathbf{XA})^H = \mathbf{XA}.$$

Ha az (1.6.13) egyenletet balról vagy jobbról az \mathbf{X} mátrixszal, ha pedig az (1.6.14) egyenletet balról vagy jobbról az \mathbf{A} mátrixszal megszorozzuk, akkor látható, hogy teljesül az a követelmény, mely szerint \mathbf{AX} és \mathbf{XA} projektor. Ha még az (1.6.15), ill. az (1.6.16) feltétel is teljesül, akkor az is következik, hogy \mathbf{AX} , ill. \mathbf{XA} hermitikus projektor. Ezek alapján a következőképpen definiáljuk az inverz mátrix fogalmának néhány általánosítását.

1.6.5 definíció. Az \mathbf{A} mátrix általánosított inverze $\mathbf{X} = \mathbf{A}^g$, ha teljesül az (1.6.13) egyenlet.

1.6.6 definíció. Az \mathbf{A} mátrix reflexív általánosított inverze $\mathbf{X} = \mathbf{A}^r$, ha teljesül az (1.6.13) és az (1.6.14) egyenlet. Ez az elnevezés arra utal, hogy $(\mathbf{A}^r)^r = \mathbf{A}$.

1.6.7 definíció. Az \mathbf{A} mátrix normált általánosított inverze $\mathbf{X} = \mathbf{A}^n$, ha teljesül az (1.6.13), (1.6.14) és az (1.6.15) egyenlet.

Megjegyzés. A normált általánosított inverzet másképpen jobbról gyengén általánosított inverznek is nevezik. Ennek megfelelően, az \mathbf{A} mátrix balról gyengén általánosított inverze $\mathbf{X} = \mathbf{A}^\omega$, ha teljesül az (1.6.13), (1.6.14) és az (1.6.16) egyenlet.

1.6.8 definíció. Az \mathbf{A} mátrix pszeudoinverze, vagy Moore–Penrose-féle* inverze $\mathbf{X} = \mathbf{A}^\dagger$, ha teljesülnek az (1.6.13)–(1.6.16) egyenletek.

E definíciókból látható, hogy a legkevesebb megszorítást az \mathbf{A}^g általánosított inverz tartalmazza, ami ezért általában nem is egyértelmű. Az alábbiakban csupán az \mathbf{A}^g inverz létezésére és az \mathbf{A}^\dagger pszeudoinverz egyértelműségére vonatkozóan bizonyítunk be egy-egy tételt. (Az általánosított inverzzel kapcsolatban lásd [3], [14], [15].)

1.6.5 tétel. Bármely \mathbf{A} mátrixhoz található \mathbf{A}^g általánosított inverz.

*E. H. Moore (1862–1932) USA-beli matematikus.

R. Penrose angol matematikus.

Bizonyítás. Legyen az \mathbf{A} mátrix rangja r . Az 1.3.11 tétel szerint ez a mátrix ekvivalens transzformációval a következő normálalakra hozható

$$\mathbf{PAQ} = \begin{bmatrix} \mathbf{E}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

Tetszőlegesen választott \mathbf{U} , \mathbf{V} és \mathbf{W} blokk segítségével képezzük most az alábbi \mathbf{A}^g mátrixot:

$$(1.6.17) \quad \mathbf{A}^g = \mathbf{Q} \begin{bmatrix} \mathbf{E}_r & \mathbf{U} \\ \mathbf{V} & \mathbf{W} \end{bmatrix} \mathbf{P}.$$

Közvetlen beszorzással meggyőződhetünk arról, hogy az \mathbf{A} és \mathbf{A}^g mátrixra teljesül az (1.6.13) egyenlet, vagyis \mathbf{A}^g az \mathbf{A} mátrix általánosított inverze. Ugyanis

$$\begin{aligned} \mathbf{AA}^g\mathbf{A} &= \mathbf{P}^{-1} \begin{bmatrix} \mathbf{E}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{Q}^{-1} \mathbf{Q} \begin{bmatrix} \mathbf{E}_r & \mathbf{U} \\ \mathbf{V} & \mathbf{W} \end{bmatrix} \mathbf{P} \mathbf{P}^{-1} \begin{bmatrix} \mathbf{E}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{Q}^{-1} = \\ &= \mathbf{P}^{-1} \begin{bmatrix} \mathbf{E}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{Q}^{-1} = \mathbf{A}, \end{aligned}$$

és ezzel a tételt bebizonyítottuk. ■

Az (1.6.13) egyenletre alkalmazva az 1.3.6 tételt, következik, hogy az r -edrangú \mathbf{A} mátrix \mathbf{A}^g általánosított inverzének a rangja legalább r :

$$\varrho(\mathbf{A}^g) \geq \varrho(\mathbf{A}).$$

Ugyancsak az (1.6.13) egyenletből következik, hogy az $m \times n$ típusú \mathbf{A} mátrix \mathbf{A}^g általánosított inverze $n \times m$ típusú mátrix.

Ha az (1.6.13) és az (1.6.14) egyenletekre alkalmazzuk az 1.3.6 tételt, akkor azt látjuk, hogy a reflexív általánosított inverz rangja már egyértelmű:

$$\varrho(\mathbf{A}^r) = \varrho(\mathbf{A}).$$

Ha az (1.6.13)–(1.6.16) egyenletek valamennyien teljesülnek, akkor az így definiált inverz egyértelműen meghatározott. Erre vonatkozik a következő tétel.

1.6.6 tétel. *Tetszőleges \mathbf{A} mátrix \mathbf{A}^\dagger Moore–Penrose-féle inverze (pseudo-inverze) egyértelmű.*

Bizonyítás. Egyszerűség kedvéért a bizonyítást valós elemű kvadratikus mátrixokra végezzük, megjegyezzük azonban, hogy lényegében ugyanígy bizonyítható a tétel komplex elemű tetszőleges mátrixokra is.

Képezzük az (1.6.13) egyenlet transzponáltját:

$$\mathbf{A}^\top \mathbf{X}^\top \mathbf{A}^\top = \mathbf{A}^\top,$$

és helyettesítsük be az (1.6.16) összefüggést:

$$(1.6.18) \quad \mathbf{X}\mathbf{A}\mathbf{A}^\top = \mathbf{A}^\top.$$

Ha az (1.6.14) egyenlet transzponáltját képezzük és az (1.6.15) összefüggést helyettesítjük be, akkor a következőt kapjuk:

$$\mathbf{X}^\top = \mathbf{X}^\top \mathbf{A}^\top \mathbf{X}^\top = \mathbf{A}\mathbf{X}\mathbf{X}^\top.$$

Mindkét oldal transzponáltját véve,

$$(1.6.19) \quad \mathbf{X} = (\mathbf{X}\mathbf{X}^\top)^\top \mathbf{A}^\top$$

adódik.

Most indirekt módszerrel bebizonyítjuk, hogy az \mathbf{A}^\dagger pszeudoinverz egyértelmű. Tegyük fel tehát, hogy van két pszeudoinverz, jelölje ezeket \mathbf{A}_1^\dagger és \mathbf{A}_2^\dagger . Mindkettő kielégíti az (1.6.18) és az (1.6.19) egyenletet, innen következik, hogy

$$(1.6.20) \quad (\mathbf{A}_1^\dagger - \mathbf{A}_2^\dagger)\mathbf{A}\mathbf{A}^\top = \mathbf{0},$$

továbbá a

$$\mathbf{B}_1 = \mathbf{A}_1^\dagger \mathbf{A}_1^{\dagger\top} \quad \text{és} \quad \mathbf{B}_2 = \mathbf{A}_2^\dagger \mathbf{A}_2^{\dagger\top}$$

jelölésekkel:

$$(1.6.21) \quad (\mathbf{A}_1^\dagger - \mathbf{A}_2^\dagger) = (\mathbf{B}_1^\top - \mathbf{B}_2^\top)\mathbf{A}^\top.$$

A továbbiakban felhasználjuk a következő mátrixazonosságot:

$$(1.6.22) \quad (\mathbf{A}_1^\dagger - \mathbf{A}_2^\dagger)\mathbf{A}\mathbf{A}^\top(\mathbf{A}_1^\dagger - \mathbf{A}_2^\dagger)^\top = [(\mathbf{A}_1^\dagger - \mathbf{A}_2^\dagger)\mathbf{A}] [(\mathbf{A}_1^\dagger - \mathbf{A}_2^\dagger)\mathbf{A}]^\top.$$

Az (1.6.20) összefüggés behelyettesítésével

$$(1.6.23) \quad [(\mathbf{A}_1^\dagger - \mathbf{A}_2^\dagger)\mathbf{A}] [(\mathbf{A}_1^\dagger - \mathbf{A}_2^\dagger)\mathbf{A}]^\top = \mathbf{0}$$

adódik. Képezzük az így nyert mátrix nyomát, ami (1.6.10) szerint $(\mathbf{A}_1^\dagger - \mathbf{A}_2^\dagger)\mathbf{A}$ elemei abszolút értékének négyzetösszege, és vegyük figyelembe, hogy ez csak akkor lehet zérus, ha

$$(1.6.24) \quad (\mathbf{A}_1^\dagger - \mathbf{A}_2^\dagger)\mathbf{A} = \mathbf{0}.$$

Ha most az (1.6.21) egyenletet jobbról megszorozzuk egy olyan \mathbf{C} mátrixszal, amelyre $\mathbf{A}^\top \mathbf{C} = \mathbf{0}$ – tehát amelynek oszlopai ortogonálisak az \mathbf{A} mátrix oszlopaira –, akkor

$$(1.6.25) \quad (\mathbf{A}_1^\dagger - \mathbf{A}_2^\dagger)\mathbf{C} = \mathbf{0}.$$

Válasszuk a \mathbf{C} mátrixot maximális rangú olyan mátrixnak, amely kielégíti az $\mathbf{A}^\top \mathbf{C} = 0$ egyenletet. Ekkor $\rho(\mathbf{C}) = n - r$, mivel az r -edrangú \mathbf{A} mátrix valamennyi oszlopára ortogonális lineárisan független vektorok maximális száma $n - r$. Tehát az \mathbf{A} és \mathbf{C} mátrixok oszlopai közül összesen n lineárisan független vektor választható ki, amelyek az (1.6.24) és (1.6.25) összefüggések szerint valamennyien ortogonálisak az $\mathbf{A}_1^\dagger - \mathbf{A}_2^\dagger$ mátrix minden sorára. Ebből következik, hogy $\mathbf{A}_1^\dagger - \mathbf{A}_2^\dagger$ minden sora zérus, azaz $\mathbf{A}_1^\dagger = \mathbf{A}_2^\dagger$ és ezzel a tételt bebizonyítottuk. ■

Az alábbi tételben egy egyszerű eljárást adunk a Moore–Penrose-féle inverz kiszámítására.

1.6.7 tétel. Legyen \mathbf{A} egy $m \times n$ típusú valós elemű mátrix, amelynek a rangja $\rho(\mathbf{A}) = r$. Ha a mátrix egy minimális diadikus előállítása

$$(1.6.26) \quad \mathbf{A} = \mathbf{U}\mathbf{V}^\top,$$

akkor a Moore–Penrose-féle inverze:

$$(1.6.27) \quad \mathbf{A}^\dagger = \mathbf{V}(\mathbf{V}^\top \mathbf{V})^{-1}(\mathbf{U}^\top \mathbf{U})^{-1} \mathbf{U}^\top.$$

Bizonyítás. Szorzással közvetlenül meggyőződhetünk róla, hogy (1.6.26) és (1.6.27) kielégítik az (1.6.13)–(1.6.16) összefüggéseket, amelyek az 1.6.6 tétel értelmében egyértelműen meghatározzák a Moore–Penrose-féle inverzet. ■

Megjegyzés. Felhívjuk a figyelmet arra, hogy noha az (1.6.26) alakú előállítás nem egyértelmű, a segítségével nyert (1.6.27) alakú Moore–Penrose-féle inverz egyértelműen adódik!

28. Példa. Határozzuk meg az

$$\mathbf{A} = \begin{bmatrix} 3 & 11 & 10 & 5 \\ 2 & 7 & 6 & 2 \\ 1 & 3 & 2 & -1 \end{bmatrix}$$

mátrix Moore–Penrose-féle inverzét.

Megoldás. Az \mathbf{A} mátrix egy minimális diadikus előállítása

$$\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 2 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 3 & 2 & -1 \\ 0 & 1 & 2 & 4 \end{bmatrix}.$$

Elvégezve a számításokat az (1.6.27) formulában, kapjuk, hogy

$$\mathbf{A}^\dagger = \frac{1}{612} \begin{bmatrix} -9 & 16 & 43 \\ -16 & 36 & 84 \\ -4 & 8 & -4 \\ 53 & -64 & -223 \end{bmatrix}.$$

* * *

1.7 LINEÁRIS EGYENLETRENDSZEREK

Ebben a szakaszban a lineáris egyenletrendszerek általános elméletét és a megoldási módszerek alapelveit tárgyaljuk, a mátrixalgebra elemeinek felhasználásával. Vizsgáljuk az egyenletrendszer megoldhatóságának feltételét és ismertetjük, hogy ennek teljesülése esetén hogyan konstruálható az egyenletrendszer általános megoldása. A tárgyalás középpontjába az együtthatókból alkotott mátrix minimális diadikus előállítását állítjuk; ily módon

- (a) egyszerű kritériumot kapunk az egyenletrendszer megoldhatóságára;
- (b) megszabadítjuk az egyenletrendszert a fölös egyenletektől;
- (c) a redukált rendszer rekurzív módon egyszerűen megoldhatóvá válik.

Ezután belátjuk, hogy ha az együtthatókból alkotott mátrix kvadrati-kus, akkor speciális esetként adódnak az előtanulmányokból ismert megoldási módszerek; majd megvizsgáljuk azokat az egyenletrendszereket, amelyeknek az együtthatómátrixa projektor.

A legáltalánosabb lineáris egyenletrendszer:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m, \end{aligned}$$

ahol az x_j ($j = 1, 2, \dots, n$) mennyiségek ismeretlenek, az a_{ij} ($i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$) számok az egyenletrendszer együtthatói és a b_i ($i = 1, 2, \dots, m$) értékek adott számok.

A mátrix fogalmanak és a mátrixműveleteknek a segítségével ez az egyenletrendszer az

$$(1.7.1) \quad \mathbf{Ax} = \mathbf{b}$$

mátrix alakban írható, ahol

$$\mathbf{A} = [a_{ij}] \quad (i = 1, 2, \dots, m; \ j = 1, 2, \dots, n)$$

az egyenletrendszer együtthatóiból álló *együtthatómátrix*,

$$\mathbf{x} = [x_j] \quad (j = 1, 2, \dots, n) \quad \text{és} \quad \mathbf{b} = [b_i] \quad (i = 1, 2, \dots, m)$$

n , ill. m elemű oszlopvektorok.

A mátrix alakban felírt egyenletrendszer mátrixegyenletként is tekinthető, ezért a lineáris egyenletrendszer helyett gyakran röviden egyenletet mondunk. Hangsúlyozni kell, hogy a lineáris egyenletrendszerek itt következő általános tárgyalása során nem tesszük fel, hogy az ismeretlenek n száma és az egyenletek m száma megegyezik, az \mathbf{A} együtthatómátrix tehát tetszőleges téglalap alakú mátrix lehet!

Az (1.7.1) lineáris egyenletrendszert $\mathbf{b} \neq \mathbf{0}$ esetén *inhomogén*, $\mathbf{b} = \mathbf{0}$ esetén pedig *homogén* lineáris egyenletrendszernek nevezzük. Először az utóbbival foglalkozunk, mert – mint látni fogjuk – az inhomogén lineáris egyenletrendszerek vizsgálata mindig visszavezethető homogén lineáris egyenletrendszerek vizsgálatára.

1.7.1 Homogén lineáris egyenletrendszer

Tekintsük az

$$(1.7.2) \quad \mathbf{Ax} = \mathbf{0}$$

alakú egyenletrendszert, ahol \mathbf{A} egy $m \times n$ típusú, r -edrangú mátrix ($r \leq \min(m, n)$). Készítsük el az \mathbf{A} mátrix egy minimális diadikus előállítását (1.3.4) szerint:

$$\mathbf{A} = \sum_{k=1}^r \mathbf{u}_k \mathbf{v}_k^T,$$

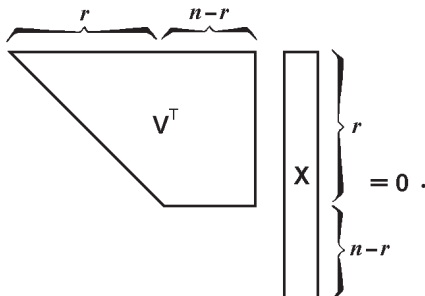
helyettesítsük be az (1.7.2) egyenletbe:

$$(1.7.3) \quad \sum_{k=1}^r \mathbf{u}_k \mathbf{v}_k^T \mathbf{x} = \mathbf{0}.$$

Az 1.3.4 tétel szerint a minimális diadikus felbontás \mathbf{u}_k oszlopvektorai lineárisan független vektorok, ezért az (1.7.3) összefüggésből következik, hogy $\mathbf{v}_k^T \mathbf{x} = 0$ ($k = 1, 2, \dots, r$). Ez annyit jelent, hogy az $\mathbf{Ax} = \mathbf{UV}^T \mathbf{x} = \mathbf{0}$ egyenlet ekvivalens az alábbi, ún. *redukált egyenlettel*:

$$(1.7.4) \quad \mathbf{V}^T \mathbf{x} = \mathbf{0}.$$

Ha még azt is figyelembe vesszük, hogy a diadikus felbontással nyert \mathbf{V}^T mátrix trapéz alakú, akkor megállapíthatjuk, hogy az eredetileg m egyenletből álló egyenletrendszert olyan r egyenletből álló rendszerre redukáltuk, amelynek együtthatómátrixa az ismeretleneket két csoportra osztja. Sématis-
kusan vázolva:



mátrixban:

$$[\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_{n-r}] = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1,n-r} \\ x_{21} & x_{22} & \dots & x_{2,n-r} \\ \dots & \dots & \dots & \dots \\ x_{r1} & x_{r2} & \dots & x_{r,n-r} \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix} = \mathbf{X}_{n-r}.$$

Ezek a megoldásvektorok lineárisan függetlenek. Ugyanis felírva a c_1, c_2, \dots, c_{n-r} mennyiségekkel alkotott lineáris kombinációjukat:

$$\sum_{i=1}^{n-r} \mathbf{x}_i c_i = \begin{bmatrix} \sum_{i=1}^{n-r} x_{1i} c_i \\ \sum_{i=1}^{n-r} x_{2i} c_i \\ \vdots \\ \sum_{i=1}^{n-r} x_{ri} c_i \\ c_1 \\ c_2 \\ \vdots \\ c_{n-r} \end{bmatrix} = \mathbf{X}_{n-r} \mathbf{c} \quad (\text{ahol } \mathbf{c} = [c_i], \mathbf{x}_i = [x_{ki}]),$$

ez csak akkor lehet zérus, ha $c_1 = c_2 = \dots = c_{n-r} = 0$. Az \mathbf{X}_{n-r} mátrix oszlopvektorai tehát az adott homogén lineáris egyenletrendszer lineárisan független partikuláris megoldásai, amelyek lineáris kombinációi az egyenletrendszer *valamennyi* megoldását szolgáltatják. Ugyanis a szabad ismeretlenek értékének tetszőleges

$$(1.7.5) \quad x_{r+1} = t_1, \quad x_{r+2} = t_2, \quad \dots, \quad x_n = t_{n-r}$$

választása esetén az \mathbf{x}_i megoldások

$$\sum_{i=1}^{n-r} \mathbf{x}_i t_i = \mathbf{X}_{n-r} \mathbf{t} \quad (\mathbf{t} = [t_i])$$

alakú lineáris kombinációja kielégíti a szabad ismeretlenekre tett (1.7.5) kikötést és az adott homogén lineáris egyenletet. Másrészt, ha $\mathbf{y} = [y_i]$ az

(1.7.2) egyenlet egy tetszőleges megoldása, akkor az $\mathbf{y} = \mathbf{X}_{n-r}\mathbf{c}$ összefüggésből $c_i = y_{r+i}$ ($i = 1, 2, \dots, n-r$) választás esetén \mathbf{y} előállítható az \mathbf{x}_i vektorok lineáris kombinációjaként.

Az

$$(1.7.6) \quad \mathbf{x} = \mathbf{X}_{n-r}\mathbf{t}$$

megoldást ezért a homogén lineáris egyenletrendszer *általános megoldásának* nevezzük.

1.7.2 Inhomogén lineáris egyenletrendszer

Térjünk most rá az $\mathbf{Ax} = \mathbf{b}$ inhomogén lineáris egyenletrendszer megoldására. Ezt a következőképpen vezetjük vissza homogén lineáris egyenletrendszer megoldására. A jobb oldali \mathbf{b} vektort a bal oldalra hozzuk:

$$\mathbf{Ax} + \mathbf{b}(-1) = \mathbf{0}.$$

Ha most az \mathbf{x} vektort kiegészítjük (-1) -gyel $(n+1)$ -edik elemként, továbbá az \mathbf{A} mátrixot kiegészítjük a \mathbf{b} vektorral mint $(n+1)$ -edik oszloppal, akkor a fenti egyenletrendszert az alábbi alakban írhatjuk fel:

$$[\mathbf{A} \quad \mathbf{b}] \begin{bmatrix} \mathbf{x} \\ -1 \end{bmatrix} = \mathbf{0}.$$

Bevezetve a következő jelölést:

$$\tilde{\mathbf{A}} = [\mathbf{A} \quad \mathbf{b}]; \quad \tilde{\mathbf{x}} = \begin{bmatrix} \mathbf{x} \\ -1 \end{bmatrix},$$

a feladatot visszavezettük az

$$(1.7.7) \quad \tilde{\mathbf{A}}\tilde{\mathbf{x}} = \mathbf{0} \quad (\tilde{\mathbf{a}}_{n+1} = \mathbf{b}, \tilde{x}_{n+1} = -1)$$

homogén lineáris egyenletrendszerre. $\tilde{\mathbf{A}}$ diadikus felbontásával a feladat az előzőekben tárgyalt módon megoldható. Egyetlen igen lényeges körülmény azonban mindig figyelembe kell venni. Ahhoz, hogy az egyenletrendszer megoldható – *kompatibilis* – legyen, vagyis ahhoz, hogy a pótlólag bevezetett \tilde{x}_{n+1} ismeretlen értéke – előírásunknak megfelelően – (-1) lehessen, \tilde{x}_{n+1} *szabad ismeretlen* kell, hogy legyen. Ezért a diadikus felbontás során generáló elemet sohasem szabad az utolsó (az $n+1$ -edik) oszlopból választani. Ha a diadikus felbontás valamelyik diádjának sorvektorában az $n+1$ -edik elem kivételével valamennyi elem zérus, akkor \tilde{x}_{n+1} nem lehet szabad ismeretlen, tehát az egyenletrendszernek nincs megoldása, az egyenletrendszer *nem kompatibilis*.

(egyenletei ellentmondóak). Mivel ebben az esetben az $\tilde{\mathbf{A}}$ mátrix e diadikus felbontása során található egy $r + 1$ -edik generáló elem, $\tilde{\mathbf{A}}$ rangja $r + 1$:

$$\varrho[\mathbf{A} \ \mathbf{b}] = r + 1.$$

A kapott eredményt a következő tételben fogalmazzuk meg:

1.7.1 tétel. Az (1.7.1) egyenletrendszer kompatibilitásának szükséges és elégséges feltétele az, hogy a \mathbf{b} vektorral bővítve az \mathbf{A} mátrixot, a mátrix rangja ne változzék, vagyis

$$(1.7.8) \quad \varrho[\mathbf{A} \ \mathbf{b}] = \varrho(\mathbf{A})$$

teljesüljön.

Megjegyezzük, hogy ha a kompatibilitási feltétel speciálisan úgy teljesül, hogy az együtthatómátrix rangja megegyezik az ismeretlenek számával, tehát

$$\varrho(\mathbf{A}) = \varrho[\mathbf{A} \ \mathbf{b}] = n,$$

akkor az egyenletrendszernek egyetlen megoldása van.

Ha a kompatibilitási feltétel teljesül, és a homogénra redukált rendszer általános megoldása $\tilde{\mathbf{x}} = \tilde{\mathbf{X}}_{n+1-r} \tilde{\mathbf{t}}$, akkor ebből úgy térhetünk vissza az eredeti inhomogén lineáris egyenletrendszer megoldására, hogy az \tilde{x}_{n+1} szabad ismeretlennek a -1 értéket adjuk.

Részletesen felírva, a homogén rendszer általános megoldása a következő:

$$\begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \vdots \\ \tilde{x}_r \\ \tilde{x}_{r+1} \\ \vdots \\ \tilde{x}_n \\ \tilde{x}_{n+1} = -1 \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1,n+1-r} \\ x_{21} & x_{22} & \dots & x_{2,n+1-r} \\ \dots & \dots & \dots & \dots \\ x_{r1} & x_{r2} & \dots & x_{r,n+1-r} \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} \tilde{t}_1 \\ \tilde{t}_2 \\ \vdots \\ \tilde{t}_{n+1-r} \end{bmatrix}.$$

Innen a $\tilde{t}_{n+1-r} = -1$ helyettesítéssel, és a

$$-x_{i,n+1-r} = x_{i0} \quad (i = 1, 2, \dots, r)$$

jelölés bevezetésével az adott inhomogén lineáris egyenletrendszer általános megoldására az alábbi adódik:

$$(1.7.9) \quad \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \\ x_{r+1} \\ x_{r+2} \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_{10} \\ x_{20} \\ \vdots \\ x_{r0} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1,n-r} \\ x_{21} & x_{22} & \dots & x_{2,n-r} \\ \dots & \dots & \dots & \dots \\ x_{r1} & x_{r2} & \dots & x_{r,n-r} \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_{n-r} \end{bmatrix},$$

ahol t_1, t_2, \dots, t_{n-r} tetszőleges paraméterek.

Az inhomogén egyenletrendszer megoldása tehát két részből áll: az első tag az egyenletrendszernek a paraméterek $t_1 = t_2 = \dots = t_{n-r} = 0$ értékei mellett adódó megoldása, amit az inhomogén lineáris egyenletrendszer egy *partikuláris megoldásának* nevezünk; a másik tag az adott inhomogén egyenletrendszerhez tartozó ($\mathbf{b} = \mathbf{0}$ esetén adódó) homogén lineáris egyenletrendszer *általános megoldása*. Összegezve tehát kimondhatjuk a következő tételt.

1.7.2 tétel. *Az inhomogén lineáris egyenletrendszer általános megoldását egy partikuláris megoldásának és a hozzá tartozó homogén lineáris egyenletrendszer általános megoldásának összege adja.*

Az alábbiakban néhány példán szemléltetjük a fenti elméleti megfontolásokat.

29. Példa. Határozzuk meg a

$$\begin{bmatrix} 3 & 11 & 10 & 5 & 10 \\ 2 & 7 & 6 & 2 & 5 \\ 1 & 3 & 2 & -1 & 2 \\ 4 & 10 & 4 & -12 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 10 \\ 8 \\ 5 \\ \lambda \end{bmatrix}$$

inhomogén lineáris egyenletrendszer megoldásait a λ paraméter függvényében.

Megoldás. Először elkészítjük a jobb oldali vektorral kiegészített együtthatómátrix egy minimális diadikus előállítását.

Megjegyezzük, hogy az ilyen egyszerű, egész együtthatós feladatokban – a törtekkel való számolás elkerülésére – a levonásra kerülő diádok generáló elemét lehetőleg úgy választjuk, hogy az oszlopában vagy sorában álló többi elemnek osztója legyen (ebben az esetben ugyanis a generáló elemmel a diád

oszlopának, ill. sorának elemei eloszthatók). Ha nem a sarokelemet választjuk generáló elemként, akkor a faktorizáció nem trapéz alapú ugyan, de erre nincs is szükségünk, mivel a továbbiakban csupán a kötött és a szabad ismeretleneket kell szétválasztani. A kötött ismeretleneket azoknak az oszlopoknak az indexe határozza meg, amelyekből a generáló elemeket választjuk, abban a sorrendben, ahogy a diádokat egymás után levonjuk. Erre azért kell ügyelni, mert a kötött ismeretleneket a szabad ismeretlenek függvényében éppen fordított sorrendben – rekurzíve – tudjuk majd kifejezni. Fontos megemlíteni, hogy az utolsó oszlopot mindenképpen utoljára kell hagyni a generáló elem kiválasztásához, mert csak így derülhet ki, lehet-e az utólag behozott ismeretlen szabad ismeretlen. Egyébként a bonyolult kifejezésekkel végzendő számolás elkerülése érdekében célszerű a határozatlan paramétert tartalmazó oszlopokból minél később generáló elemet választani.

Az alábbi felbontás során lépésenként bekeretezzük a generáló elemet.

$$\begin{bmatrix} 3 & 11 & 10 & 5 & 10 & 10 \\ 2 & 7 & 6 & 2 & 5 & 8 \\ \boxed{1} & 3 & 2 & -1 & 2 & 5 \\ 4 & 10 & 4 & -12 & 4 & \lambda \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \\ 1 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 3 & 2 & -1 & 2 & 5 \end{bmatrix},$$

$$\begin{bmatrix} 0 & 2 & 4 & 8 & 4 & -5 \\ 0 & \boxed{1} & 2 & 4 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & -4 & -8 & -4 & \lambda - 20 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \\ 0 \\ -2 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 & 4 & 1 & -2 \end{bmatrix},$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & \boxed{2} & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 & \lambda - 24 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 2 & -1 \end{bmatrix},$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{\lambda - 25} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \lambda - 25 \end{bmatrix}.$$

A jobb oldali vektorral kiegészített együtthatómátrix faktorizált alakja ezért

$$[\mathbf{A} \quad \mathbf{b}] = \begin{bmatrix} 3 & 2 & 1 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 4 & -2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 2 & -1 & 2 & 5 \\ 0 & 1 & 2 & 4 & 1 & -2 \\ 0 & 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & \lambda - 25 \end{bmatrix}.$$

Az adott inhomogén lineáris egyenletrendszer megoldása tehát ekvivalens az alábbi homogén lineáris egyenletrendszer megoldásával:

$$\begin{bmatrix} 1 & 3 & 2 & -1 & 2 & 5 \\ 0 & 1 & 2 & 4 & 1 & -2 \\ 0 & 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & \lambda - 25 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \mathbf{0}.$$

Az egyenletrendszer kompatibilitási feltétele, hogy x_6 szabad ismeretlen legyen. Ez pedig azt jelenti, hogy az együtthatómátrixnak nem lehet olyan sora, amelyben csak az utolsó elem zérustól különböző; tehát csak

$$\lambda - 25 = 0, \quad \text{vagyis} \quad \lambda = 25$$

esetén létezik az adott egyenletrendszernek megoldása.

A következő lépés az egyenletrendszer kötött és szabad ismeretleneinek szétválasztása. Mivel a generáló elemeket rendre az első, második és ötödik oszlopból választottuk, x_1 , x_2 és x_5 lesz kötött ismeretlen, x_3 , x_4 és x_6 pedig szabad ismeretlen. Az együtthatómátrix oszlopainak megfelelő átrendezésével az egyenletrendszer tehát a következő alakba írható:

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 2 & 2 & -1 & 5 \\ 0 & 1 & 1 & 2 & 4 & -2 \\ 0 & 0 & 2 & 0 & 0 & -1 \end{array} \right] \begin{bmatrix} x_1 \\ x_2 \\ x_5 \\ x_3 \\ x_4 \\ x_6 \end{bmatrix} = \mathbf{0}.$$

A kötött ismeretleneket rekurzív úton határozzuk meg, a szabad ismeretlenek függvényében; ehhez a mátrix alakban felírt egyenletet átírjuk egyenletrendszer alakjába és úgy rendezzük, hogy a bal oldalra a kötött, a jobb oldalra a szabad ismeretlenek kerüljenek:

$$\begin{aligned} 2x_5 &= x_6, \\ x_2 + x_5 &= -2x_3 - 4x_4 + 2x_6, \\ x_1 + 3x_2 + 2x_5 &= -2x_3 + x_4 - 5x_6. \end{aligned}$$

Innen

$$\begin{aligned} x_5 &= \frac{1}{2} x_6; \\ x_2 &= -2x_3 - 4x_4 + \frac{3}{2} x_6; \\ x_1 &= 4x_3 + 13x_4 - \frac{21}{2} x_6. \end{aligned}$$

Ezek után a szabad ismeretlenek értékeit kétféleképpen választjuk meg:

$$(a) \quad x_3 = 1, \quad x_4 = 0, \quad x_6 = 0,$$

$$(b) \quad x_3 = 0, \quad x_4 = 1, \quad x_6 = 0.$$

Ezek behelyettesítésével a kötött ismeretlenek:

$$(a) \quad x_1 = 4, \quad x_2 = -2, \quad x_5 = 0$$

és

$$(b) \quad x_1 = 13, \quad x_2 = -4, \quad x_5 = 0.$$

Az adott egyenletrendszerhez tartozó homogén lineáris egyenletrendszerre így az alábbi két, lineárisan független partikuláris megoldást kapjuk:

$$(a) \quad \begin{bmatrix} 4 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad (b) \quad \begin{bmatrix} 13 \\ -4 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

A homogén lineáris egyenletrendszer általános megoldását ezek lineáris kombinációja adja:

$$\begin{bmatrix} 4 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} t_1 + \begin{bmatrix} 13 \\ -4 \\ 0 \\ 1 \\ 0 \end{bmatrix} t_2 = \begin{bmatrix} 4 & 13 \\ -2 & -4 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}.$$

Végül a szabad ismeretleneknek az alábbi értékeket adjuk:

$$(c) \quad x_3 = 0, \quad x_4 = 0, \quad x_6 = -1;$$

így megkapjuk az adott inhomogén egyenletrendszer egy partikuláris megoldását:

$$\mathbf{x}_0 = \begin{bmatrix} \frac{21}{2} \\ \frac{3}{2} \\ -\frac{3}{2} \\ 0 \\ 0 \\ -\frac{1}{2} \end{bmatrix}.$$

A keresett megoldást e kettő összege adja:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 4 & 13 \\ -2 & -4 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} + \begin{bmatrix} \frac{21}{2} \\ -\frac{3}{2} \\ 0 \\ 0 \\ -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 4t_1 + 13t_2 + \frac{21}{2} \\ -2t_1 - 4t_2 - \frac{3}{2} \\ t_1 \\ t_2 \\ -\frac{1}{2} \end{bmatrix}.$$

* * *

30. Példa. Határozzuk meg a

$$\begin{bmatrix} 2 & 5 & 1 & 3 \\ 4 & 6 & 3 & 5 \\ 4 & 14 & 1 & 7 \\ 2 & -3 & 3 & \lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 4 \\ 7 \end{bmatrix}$$

egyenletrendszer megoldását a λ paraméter függvényében.

Megoldás. Elkészítjük a jobb oldali vektorral kiegészített együtthatómátrix egy minimális diadikus felbontását:

$$[\mathbf{A} \quad \mathbf{b}] = \begin{bmatrix} 1 & 0 & 0 \\ 3 & -1 & 0 \\ 1 & 1 & 0 \\ 3 & -2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 5 & 1 & 3 & 2 \\ 2 & 9 & 0 & 4 & 2 \\ 0 & 0 & 0 & \lambda - 1 & 5 \end{bmatrix}.$$

A redukált homogén lineáris egyenletrendszer mátrix alakja tehát

$$\begin{bmatrix} 2 & 5 & 1 & 3 & 2 \\ 2 & 9 & 0 & 4 & 2 \\ 0 & 0 & 0 & \lambda - 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \mathbf{0}.$$

Az egyenletrendszer megoldhatóságának feltétele, hogy x_5 szabad ismeretlen legyen. Az együtthatómátrixnak tehát nem lehet olyan sora, amelyben az utolsó elem kivételével minden elem zérus. Az egyenletrendszernek ezért csak akkor van megoldása, ha

$$\lambda - 1 \neq 0, \text{ vagyis } \lambda \neq 1.$$

A kötött ismeretlenek (sorrendben) x_3, x_1, x_4 , a szabad ismeretlenek x_2

és x_5 . A megoldandó egyenletrendszer (a rekurzív megoldás sorrendjében):

$$\begin{aligned}(\lambda - 1)x_4 &= -5x_5; \\ 2x_1 + 4x_4 &= -9x_2 - 2x_5; \\ x_3 + 2x_1 + 3x_4 &= -5x_2 - 2x_5.\end{aligned}$$

A keresett megoldás tehát az alábbi alakban adódik:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -\frac{9}{2} \\ 1 \\ 4 \\ 0 \end{bmatrix} t + \begin{bmatrix} \frac{\lambda - 11}{\lambda - 1} \\ 0 \\ 5 \\ \frac{5}{\lambda - 1} \end{bmatrix}$$

ahol $\lambda \neq 1$.

* * *

31. Példa. Határozzuk meg az alábbi egyenletrendszer megoldását a $\lambda_1, \lambda_2, \dots, \lambda_n$ paraméterek függvényében.

$$\begin{bmatrix} \lambda_1 & 1 & 1 & \dots & 1 & 1 \\ 1 & \lambda_2 & 1 & \dots & 1 & 1 \\ 1 & 1 & \lambda_3 & \dots & 1 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & \lambda_{n-1} & 1 \\ 1 & 1 & 1 & \dots & 1 & \lambda_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix}.$$

Megoldás. Három fő esetet kell megkülönböztetni:

- (a) egyik λ_k értéke sem egyenlő 1-gyel;
- (b) egyetlen λ_k értéke 1, a többi nem egyenlő 1-gyel;
- (c) p számú λ_k értéke 1, a többi nem egyenlő 1-gyel.

(a) Legyen $\lambda_k \neq 1$ minden k -ra. Jelölje a λ_k elemekből alkotott diagonálmátrixot \mathbf{A} , a csupa 1 elemből álló oszlopvektort \mathbf{e} . Ezekkel a jelölésekkel az együtthatómátrix felírható mint $(\mathbf{A} - \mathbf{E}) + \mathbf{e}\mathbf{e}^T$, tehát mint az egyetlen $\mathbf{e}\mathbf{e}^T$ diáddal módosított $\mathbf{A} - \mathbf{E}$ diagonálmátrix.

Ekkor két alesetet különböztetünk meg:

(α) Ha $1 + \mathbf{e}^T(\mathbf{A} - \mathbf{E})^{-1}\mathbf{e} \neq 0$, akkor az (1.5.29) alatti Sherman–Morrison-formula szerint az együtthatómátrix invertálható és az inverz az alábbi képlet alapján számítható:

$$[(\mathbf{A} - \mathbf{E}) + \mathbf{e}\mathbf{e}^T]^{-1} = (\mathbf{A} - \mathbf{E})^{-1} - \frac{(\mathbf{A} - \mathbf{E})^{-1}\mathbf{e}\mathbf{e}^T(\mathbf{A} - \mathbf{E})^{-1}}{1 + \mathbf{e}^T(\mathbf{A} - \mathbf{E})^{-1}\mathbf{e}}.$$

Így az egyenletrendszer megoldására a következő adódik:

$$\mathbf{x} = [(\mathbf{A} - \mathbf{E}) + \mathbf{e}\mathbf{e}^\top]^{-1} \mathbf{e} = \frac{1}{1 + \mathbf{e}^\top(\mathbf{A} - \mathbf{E})^{-1}\mathbf{e}} (\mathbf{A} - \mathbf{E})^{-1} \mathbf{e}.$$

Figyelembe véve, hogy

$$\mathbf{e}^\top(\mathbf{A} - \mathbf{E})^{-1}\mathbf{e} = \sum_{v=1}^n \frac{1}{\lambda_v - 1},$$

az egyes ismeretlenek értéke

$$x_k = \frac{1}{1 + \sum_{v=1}^n \frac{1}{\lambda_v - 1}} \cdot \frac{1}{\lambda_k - 1} \quad (k = 1, 2, \dots, n).$$

(β) Ha $1 + \mathbf{e}^\top(\mathbf{A} - \mathbf{E})^{-1}\mathbf{e} = 0$, azaz

$$1 + \sum_{v=1}^n \frac{1}{\lambda_v - 1} = 0 \quad (\lambda_v \neq 1),$$

akkor az együtthatómátrix szinguláris. Megmutatjuk, hogy az egyenletrendszernek ekkor nincs megoldása, mert a kompatibilitási feltétel nem teljesül. Erről legegyszerűbben a következőképpen győződhetünk meg. A jobb oldali \mathbf{e} vektorral kiegészített együtthatómátrixot balról megszorozzuk a nemszinguláris \mathbf{R} mátrixszal, ahol

$$\mathbf{R} = \begin{bmatrix} 1 & & & -1 \\ & 1 & & -1 \\ & & 1 & -1 \\ \dots & \dots & \dots & \dots \\ & & & 1 & -1 \\ & & & & 1 \end{bmatrix}.$$

A szorzás eredménye a következő mátrixot adja:

$$\begin{bmatrix} \lambda_1 - 1 & 0 & \dots & 0 & 1 - \lambda_n & 0 \\ 0 & \lambda_2 - 1 & \dots & 0 & 1 - \lambda_n & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_{n-1} - 1 & 1 - \lambda_n & 0 \\ 1 & 1 & \dots & 1 & \lambda_n & 1 \end{bmatrix}.$$

Innen kiolvasható, hogy az utolsó előtti oszlop elhagyásával adódó n -edrendű minormátrix determinánsa

$$\prod_{k=1}^{n-1} (\lambda_k - 1) \neq 0,$$

tehát a kiegészített együtthatómátrix rangja n . Ez azt jelenti, hogy a kiegészített mátrix rangja nem egyenlő az eredeti mátrix rangjával (hiszen a feltétel értelmében az szinguláris, tehát az 1.7.1 tétel alapján az egyenletrendszer nem kompatibilis, vagyis valóban nincs megoldása).

Megjegyzés. Érdekes külön megemlíteni, hogy abban a speciális esetben, amikor valamennyi λ_k értéke egyenlő, akkor az együtthatómátrix $\lambda_k = 1 - n$ ($k = 1, 2, \dots, n$) esetén válik szinguláris, tehát ez esetben nincs megoldása az egyenletrendszernek.

(b) Ha valamelyik λ_k értéke 1 – legyen például $\lambda_j = 1$ –, a többi pedig nem egyenlő 1-gyel, akkor az együtthatómátrix továbbra is nonszinguláris, tehát az egyenletrendszernek egyetlen megoldása van. Ezt a fenti megoldásból úgy nyerhetjük, hogy a törtet bővítjük a $\lambda_j - 1$ kifejezéssel, és elvégezzük a $\lambda_j \rightarrow 1$ határátmenetet. Így az $x_j = 1$; ha $k \neq j$ eredményre jutunk.

(c) Legyen $p > 1$ számú λ_k értéke 1, a többi pedig nem egyenlő 1-gyel. Az általánosság megszorítása nélkül feltehető, hogy $\lambda_1 = \lambda_2 = \dots = \lambda_p = 1$. Az együtthatómátrix ekkor szinguláris, a rangja $n - p + 1$. A jobb oldali vektorral kiegészített együtthatómátrix diadikus felbontása a következő alakú:

$$\begin{bmatrix} 1 & 1 \dots 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & 1 \dots 1 & 1 & 1 & \dots & 1 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 1 \dots 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & 1 \dots 1 & \lambda_{p+1} & 1 & \dots & 1 & 1 \\ 1 & 1 \dots 1 & 1 & \lambda_{p+2} \dots 1 & 1 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 1 \dots 1 & 1 & 1 & \dots & \lambda_n & 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \dots 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 0 & 0 \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 \dots & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 \dots & \lambda_{p+1} - 1 & 0 & \dots & 0 & 0 \\ 0 & 0 \dots & 0 & \lambda_{p+2} - 1 \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 \dots & 0 & 0 & \dots & \lambda_n - 1 & 0 \end{bmatrix}.$$

A redukált egyenletrendszer tehát a következő:

$$\begin{bmatrix} 1 & 1 & 1 \dots & 1 & 1 & \dots & 1 & 1 \\ 0 & 0 & 0 \dots & \lambda_{p+1} - 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 \dots & 0 & \lambda_{p+2} - 1 \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 \dots & 0 & 0 & \dots & \lambda_n - 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \\ x_{n+1} \end{bmatrix} = \mathbf{0}.$$

Mivel x_{n+1} szabad ismeretlennek adódik, az egyenletrendszernek van megoldása. Közvetlenül kiolvasható, hogy

$$x_{p+1} = x_{p+2} = \dots = x_n = 0.$$

Szabad ismeretlennek választható x_2, x_3, \dots, x_p , tehát az általános megoldás:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_p \\ x_{p+1} \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} -1 & -1 & \dots & -1 \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_{p-1} \end{bmatrix}.$$

* * *

1.7.3 Lineáris egyenletrendszer kvadratikus együtthatómátrixszal

A következőkben azzal a speciális esettel foglalkozunk, amikor az egyenletrendszer együtthatómátrixa kvadratikus (vagyis az ismeretlenek száma és az egyenletek száma megegyezik, $n = m$). Ekkor még két alesetet kell megkülönböztetnünk aszerint, hogy az együtthatómátrix nemszinguláris, ill. szinguláris.

(a) **Nemszinguláris együtthatómátrix.** Abban az esetben, ha nemszinguláris az együtthatómátrix (azaz ha $r = n$), ennek inverzével balról megszorozva az egyenletet, megkapjuk az egyenletrendszer *egyetlen* megoldását.

Homogén lineáris egyenletrendszer esetén az $\mathbf{Ax} = \mathbf{0}$ (ahol $|\mathbf{A}| \neq 0$) egyenletből $\mathbf{x} = \mathbf{0}$ következik, ami az egyenletrendszer triviális megoldása.

Inhomogén lineáris egyenletrendszer esetén az

$$(1.7.10) \quad \mathbf{Ax} = \mathbf{b} \quad (\text{ahol } |\mathbf{A}| \neq 0)$$

egyenletből

$$(1.7.11) \quad \mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

adódik. Ha az inverz helyére behelyettesítjük az (1.2.34) összefüggést, akkor az (1.7.11) megoldás az alábbi alakban írható:

$$(1.7.12) \quad \mathbf{x} = \frac{(\text{adj } \mathbf{A})\mathbf{b}}{|\mathbf{A}|}.$$

diadikus felbontásával kapott faktorizáció tényezőit, vagyis az $\tilde{\mathbf{U}}$ alsó háromszögmátrixot és a $\tilde{\mathbf{V}}^T$ ($n + 1$ oszlopot tartalmazó) felső trapézmátrix „főátló” feletti elemeit – gépi tárolókapacitás megtakarítása végett – egyetlen mátrixszá „összetoljuk”:

$$\tilde{\mathbf{A}} = \tilde{\mathbf{U}} \tilde{\mathbf{V}}^T$$

Az $\tilde{\mathbf{A}}$ mátrix elemeit az $\tilde{\mathbf{U}}$ és $\tilde{\mathbf{V}}^T$ tényezők elemeinek segítségével a következő alakban írhatjuk fel:

$$a_{ik} = u_{i1}v_{1k} + u_{i2}v_{2k} + \cdots + u_{i,k-1}v_{k-1,k} + u_{ik}, \quad \text{ha } i \geq k,$$

és

$$a_{ik} = u_{i1}v_{1k} + u_{i2}v_{2k} + \cdots + u_{i,i-1}v_{i-1,k} + u_{ii}v_{ik}, \quad \text{ha } i < k.$$

Innen $\tilde{\mathbf{U}}$ és $\tilde{\mathbf{V}}^T$ elemeire az

$$(1.7.14) \quad u_{ik} = a_{ik} - u_{i1}v_{1k} - \cdots - u_{i,k-1}v_{k-1,k}, \quad \text{ha } i \geq k,$$

és

$$(1.7.15) \quad v_{ik} = \frac{1}{u_{ii}}(a_{ik} - u_{i1}v_{1k} - \cdots - u_{i,i-1}v_{i-1,k}), \quad \text{ha } i < k,$$

rekurziós összefüggéseket nyerjük. Ezeket a képleteket a következőképpen alkalmazzuk. Először \tilde{U} első oszlopát:

$$(1.7.16) \quad u_{i1} = a_{i1} \quad (i = 1, 2, \dots, n),$$

majd $\tilde{\mathbf{V}}^\top$ első sorát írjuk fel:

$$(1.7.17) \quad v_{1k} = \frac{1}{u_{11}} a_{1k} \quad (k = 2, 3, \dots, n).$$

Ezután \tilde{U} második oszlopának elemeit

$$(1.7.18) \quad u_{i2} = a_{i2} - u_{i1}v_{12} \quad (i = 2, 3, \dots, n)$$

szerint, majd $\tilde{\mathbf{V}}^\top$ második sorának elemeit

$$(1.7.19) \quad v_{2k} = \frac{1}{u_{22}}(a_{2k} - u_{21}v_{1k}) \quad (k = 3, 4, \dots, n)$$

alapján számítjuk, és ezt folytatjuk, felváltva haladva oszloponként és soronként, amíg a teljes „trianguláris faktorizációt” el nem érjük*:

$$(1.7.20) \quad \left[\begin{array}{c|cccc} u_{11} & v_{12} & v_{13} & \dots & v_{1n} & v_{1,n+1} \\ u_{21} & u_{22} & v_{23} & \dots & v_{2n} & v_{2,n+1} \\ u_{31} & u_{32} & u_{33} & \dots & v_{3n} & v_{3,n+1} \\ \vdots & & & & & \\ u_{n1} & u_{n2} & u_{n3} & \dots & u_{nn} & v_{n,n+1} \end{array} \right].$$

Megjegyezzük, hogy a közölt algoritmus szerint az adott \mathbf{A} mátrix elemeit fokozatosan „felülírhatjuk” a faktorizációval nyert (1.7.20) „összetolt” mátrix elemeivel, így az eljárás nem igényel újabb tárolókapacitást.

A gyakorlatban az együtthatók táblázatát kiegészítjük az egyes oszlopok elemeinek összegét tartalmazó $(n+1)$ -edik sorral, és az egyes sorok elemeinek összegét tartalmazó $(n+2)$ -edik oszloppal; ezek a kontroll-szummák. Alkalmazásuk biztosítja a számítás „menet közbeni” ellenőrzését: úgy számolunk velük, mintha az adott feladathoz tartozó együtthatók lennének, a trianguláris faktorizáció végén nyert $(n+1)$ -edik sor az $\tilde{\mathbf{U}}$ oszlopainak kontroll-szummája, az $(n+2)$ -edik oszlop pedig $\tilde{\mathbf{V}}^T$ sorainak kontroll-szummája (nem feledkezve meg a főátlóban explicite nem szereplő egységekről!).

A mondottak szemléltetésére felvázoljuk az eljárás során fellépő mátrixokat. Az egyenletrendszer jobb oldalán álló vektorral kiegészített együtthatómátrix trianguláris faktorizációja legyen a következő:

$$\begin{array}{c} n+1 \\ \hline \begin{array}{c} n \\ \hline \tilde{\mathbf{A}} \end{array} \end{array} = \begin{array}{c} \tilde{\mathbf{U}} \end{array} \begin{array}{c} \begin{array}{c} 1 \\ \hline 1 \\ \hline \vdots \\ \hline 1 \end{array} \\ \tilde{\mathbf{V}}^T \end{array}.$$

Az oszlopelemek összegezését úgy vesszük figyelembe, hogy az n -edrendű egységmátrixot egy $(n+1)$ -edik összegező sorvektorral egészítjük ki (ezzel balról szorzunk), a sorelemek összegezéséhez pedig az $(n+1)$ -edrendű egységmátrixot egy $(n+2)$ -edik összegező oszlopvektorral egészítjük ki (és ezzel jobbról szorzunk):

* Az így kapott trianguláris faktorizációt az irodalomban általában **LU**-faktorizációnak nevezik, utalva arra, hogy **L** alsó háromszögmátrix (*lower triangular matrix*), **U** pedig felső háromszögmátrix (*upper triangular matrix*). (Lásd [20].)

$$\begin{array}{c}
 \begin{array}{|c|} \hline \mathbf{E}_n \\ \hline 1 \quad \cdot \quad \cdot \quad \cdot \quad 1 \\ \hline \end{array}
 \quad
 \begin{array}{|c|} \hline \tilde{\mathbf{A}} \\ \hline \end{array}
 \quad
 \begin{array}{|c|} \hline \mathbf{E}_{n+1} \\ \hline 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \\ \hline \end{array}
 = \\
 \\
 =
 \begin{array}{|c|} \hline \mathbf{E}_n \\ \hline 1 \quad \cdot \quad \cdot \quad \cdot \quad 1 \\ \hline \end{array}
 \quad
 \begin{array}{|c|} \hline \tilde{\mathbf{U}} \\ \hline \end{array}
 \quad
 \begin{array}{|c|} \hline 1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 1 \\ \hline \tilde{\mathbf{V}}^T \\ \hline \end{array}
 \quad
 \begin{array}{|c|} \hline \mathbf{E}_{n+1} \\ \hline 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \\ \hline \end{array}
 \end{array}$$

A szorzásokat elvégezve látható, hogy az egyenlet bal oldalán olyan $(n+1) \times (n+2)$ típusú mátrixot kapunk, amelynek $(n+1)$ -edik sorában $\tilde{\mathbf{A}}$ oszlopelemeinek összege (ezeket alsó indexszel jelöltük), $(n+2)$ -edik oszlopában pedig a sorelemek összege áll (megkülönböztetésül ezeket felső indexszel láttuk el; a jobb alsó sarokban álló kettős szumma pedig arra utal, hogy itt $\tilde{\mathbf{A}}$ valamennyi elemének az összege szerepel). Az egyenlet jobb oldalán, az első tényező utolsó sorában $\tilde{\mathbf{U}}$ oszlopelemeinek összege a második tényező utolsó oszlopában pedig $\tilde{\mathbf{V}}^T$ sorelemeinek összege áll:

$$\begin{array}{|c|} \hline \tilde{\mathbf{A}} \\ \hline \Sigma_1 \quad \Sigma_2 \quad \cdot \quad \cdot \quad \Sigma_n \\ \hline \end{array}
 \quad
 \begin{array}{|c|} \hline \Sigma^1 \\ \Sigma^2 \\ \cdot \\ \cdot \\ \Sigma^n \\ \hline \Sigma \Sigma \\ \hline \end{array}
 =
 \begin{array}{|c|} \hline \tilde{\mathbf{U}} \\ \hline \Sigma_1 \quad \Sigma_2 \quad \cdot \quad \cdot \quad \Sigma_n \\ \hline \end{array}
 \quad
 \begin{array}{|c|} \hline 1 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \\ \hline \tilde{\mathbf{V}}^T \\ \hline \end{array}
 \quad
 \begin{array}{|c|} \hline \Sigma^1 \\ \Sigma^2 \\ \cdot \\ \cdot \\ \Sigma^n \\ \hline \end{array}$$

Vegyük most figyelembe, hogy az egyenlet bal oldalán álló mátrix rangja n , a faktORIZÁCIÓ során tehát n diád levonása során alakul ki az „összetolt” mátrix. Vagyis utolsó két lépésként a bal oldalon álló mátrix $(n+1)$ -edik sorának $(n+1)$ -edik eleméből (Σ_{n+1} -ből), illetve $(n+2)$ -edik eleméből ($\Sigma \Sigma$ -ből) vonjuk le a jobb oldal első tényező $(n+1)$ -edik sorának és a második tényező $(n+1)$ -edik, illetve $(n+2)$ -edik oszlopának (skaláris) szorzatát, így az „összetolt” mátrix utolsó sorában ez a két utolsó elem szükségképpen zérust kell adjon minden esetben. (Numerikus feladatok megoldásakor végső kontrollként jó szolgálatot tehet ennek a két feltételnek az ellenőrzése is!)

Mivel az (1.7.7) egyenletrendszer az együtthatómátrix trianguláris faktORIZÁCIÓJÁVAL a $\tilde{\mathbf{V}}^T \tilde{\mathbf{x}} = \mathbf{0}$ egyenletrendszerre redukálható, ez az utolsó egyenletből kiindulva, rekurzív úton megoldható. Behelyettesítve az $x_{n+1} = -1$ értéket, így

$$\begin{aligned} x_n &= v_{n,n+1}, \\ x_i &= v_{i,n+1} - v_{i,i+1}x_{i+1} - \cdots - v_{in}x_n \quad (i = n-1, n-2, \dots, 2, 1). \end{aligned}$$

Ennél a módszernél a faktORIZÁCIÓ körülbelül $\frac{1}{3}n^3$, az ismeretlenek rekurzív kiszámítása pedig körülbelül n^2 szorzás elvégzését teszi szükségessé. Meg kell jegyezni, hogy a kerekítési hibák folytán – különösen a sok ismeretlent tartalmazó egyenletrendszerek esetén – az itt leírt módszerrel nyert megoldás a pontos megoldás egy első közelítésének tekintendő.

32. Példa. Oldjuk meg az alábbi lineáris egyenletrendszert a Gauss–Crout-féle eliminációs módszerrel:

$$\begin{aligned} 2x_1 - x_2 - x_3 + 3x_4 + 2x_5 &= 6, \\ 6x_1 - 2x_2 + 3x_3 - x_5 &= -3, \\ -4x_1 + 2x_2 + 3x_3 - 3x_4 - 2x_5 &= -5, \\ 2x_1 + 4x_3 - 7x_4 - 3x_5 &= -8, \\ x_2 + 8x_3 - 5x_4 - x_5 &= -3. \end{aligned}$$

Megoldás. A feladat megoldását az alábbi táblázatos elrendezés szemlélteti:

							Σ
	2	-1	-1	3	2	6	11
	6	-2	3	0	-1	-3	3
	-4	2	3	-3	-2	-5	-9
	2	0	4	-7	-3	-8	-12
	0	1	8	-5	-1	-3	0
Σ	6	0	17	-12	-5	-13	-7
	2	$-\frac{1}{2}$	$-\frac{1}{2}$	$\frac{3}{2}$	1	3	$\frac{11}{2}$
	6	1	6	-9	-7	-21	-30
	-4	0	1	3	2	7	13
	2	1	-1	2	2	7	10
	0	1	2	-2	6	3	4
Σ	6	3	2	0	6	0	0
$x_i =$	8	21	-2	1	3	-1	0

A kontrollszummák nem jeleznek hibát, tehát az alábbi faktorizációhoz jutottunk:

$$\begin{bmatrix} 2 & & & & & \\ & 6 & 1 & & & \\ & -4 & 0 & 1 & & \\ & 2 & 1 & -1 & 2 & \\ & 0 & 1 & 2 & -2 & 6 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & \frac{3}{2} & 1 & 3 \\ & 1 & 6 & -9 & -7 & -21 \\ & & 1 & 3 & 2 & 7 \\ & & & 1 & 2 & 7 \\ & & & & 1 & 3 \end{bmatrix}.$$

Ennek alapján a redukált egyenletrendszer rekurzív módon megoldható:

$$\begin{aligned} x_5 &= 3, \\ x_4 &= 7 - 2x_5 = 1, \\ x_3 &= 7 - 2x_5 - 3x_4 = -2, \\ x_2 &= -21 + 7x_5 + 9x_4 - 6x_3 = 21, \\ x_1 &= 3 - x_5 - \frac{3}{2}x_4 + \frac{1}{2}x_3 + \frac{1}{2}x_2 = 8. \\ &\qquad\qquad\qquad * \quad * \quad * \end{aligned}$$

Abban a speciális esetben, amikor az egyenletrendszer együtthatómátrixa *sávmátrix*, azaz amikor

$$(1.7.21) \quad a_{ij} = 0, \quad \text{ha } i - j > q \text{ és } j - i > p,$$

ahol p és q pozitív egész, valamint a mátrix bal felső sarokminorai zérustól különbözőek, akkor a trianguláris faktorizáció „öröklí” ezt a tulajdonságot.

Mielőtt ezt az állítást bebizonyítanánk, emlékeztetünk az 1.3.1 tételre. Ott megmutattuk, hogy amennyiben egy n -edrendű nonszinguláris mátrix minimális diadikus felbontása során a bal felső sarokelem minden lépésben zérustól különböző és ezt választjuk a levonandó diádok generáló elemeként, akkor a diadikus felbontás ekvivalens a mátrix trianguláris faktorizációjával és a mátrix bal felső sarokminorai szükségképpen zérustól különbözőek. Most bebizonyítjuk a tétel megfordítását (lásd [11]).

1.7.3 tétel. *Ha az $\mathbf{A} = [a_{ij}]$ nonszinguláris mátrix bal felső sarokminorai zérustól különbözőek:*

$$(1.7.22) \quad |\mathbf{A}_{1,2,\dots,k}^{1,2,\dots,k}| \neq 0 \quad (k = 1, 2, \dots, n),$$

akkor létezik a mátrix (1.3.3) alakú minimális diadikus felbontása, ami ekvivalens a mátrix (1.3.4) alatti trianguláris faktorizációjával.

Bizonyítás. A bizonyítást teljes indukcióval végezzük; $n = 1$ esetén az állítás triviálisan teljesül, hiszen $a_{11} \neq 0$ és felírható $a_{11} = a_{11} \cdot 1$ alakban. Tegyük

fel tehát, hogy az állítás $n = k - 1$ esetén fennáll, és ebből következtetünk arra, hogy $n = k$ esetén is teljesül. Jelölje \mathbf{A}_{k-1} a $(k - 1)$ -edrendű bal felső sarokblokkot, amelynek faktorizációja a feltevés értelmében

$$\mathbf{A}_{k-1} = \mathbf{U}_{k-1} \mathbf{V}_{k-1}^T,$$

ahol \mathbf{U}_{k-1} $(k - 1)$ -edrendű nonszinguláris alsó, \mathbf{V}_{k-1}^T pedig $(k - 1)$ -edrendű nonszinguláris felső háromszögmátrix. Be kell látnunk, hogy ha $|\mathbf{A}_k| \neq 0$, akkor lehetséges az $\mathbf{A}_k = \mathbf{U}_k \mathbf{V}_k^T$ trianguláris faktorizáció. Legyen \mathbf{A}_k particionált alakja

$$\mathbf{A}_k = \begin{bmatrix} \mathbf{A}_{k-1} & \mathbf{f} \\ \mathbf{g}^T & a_{kk} \end{bmatrix},$$

és tegyük fel, hogy a faktorizációval nyert tényezők particionált alakja a következő:

$$\mathbf{U}_k = \begin{bmatrix} \mathbf{U}_{k-1} & \mathbf{0} \\ \mathbf{d}^T & u_{kk} \end{bmatrix}, \quad \mathbf{V}_k^T = \begin{bmatrix} \mathbf{V}_{k-1}^T & \mathbf{c} \\ \mathbf{0} & 1 \end{bmatrix}.$$

A beszorzást elvégezve innen

$$\begin{bmatrix} \mathbf{A}_{k-1} & \mathbf{f} \\ \mathbf{g}^T & a_{kk} \end{bmatrix} = \begin{bmatrix} \mathbf{U}_{k-1} \mathbf{V}_{k-1}^T & \mathbf{U}_{k-1} \mathbf{c} \\ \mathbf{d}^T \mathbf{V}_{k-1}^T & \mathbf{d}^T \mathbf{c} + u_{kk} \end{bmatrix}$$

adódik, azaz

$$(1.7.23) \quad \mathbf{U}_{k-1} \mathbf{c} = \mathbf{f}, \quad \mathbf{d}^T \mathbf{V}_{k-1}^T = \mathbf{g}^T \quad \text{és} \quad a_{kk} = u_{kk} + \mathbf{d}^T \mathbf{c}.$$

Mivel pedig \mathbf{U}_{k-1} és \mathbf{V}_{k-1}^T nonszinguláris, így az (1.7.23) összefüggésekből \mathbf{c} és \mathbf{f} egyértelműen meghatározható;

$$\mathbf{c} = \mathbf{U}_{k-1}^{-1} \mathbf{f} \quad \text{és} \quad \mathbf{d}^T = \mathbf{g}^T (\mathbf{V}_{k-1}^T)^{-1},$$

és ezzel

$$a_{kk} = u_{kk} + \mathbf{g}^T (\mathbf{V}_{k-1}^T)^{-1} \mathbf{U}_{k-1}^{-1} \mathbf{f},$$

ahonnan

$$u_{kk} = a_{kk} - \mathbf{g}^T \mathbf{A}_{k-1}^{-1} \mathbf{f}.$$

Figyelembe véve az 1.5.3 tételt és az (1.5.20) összefüggést, az \mathbf{A}_k determinánsára tett feltétel értelmében

$$|\mathbf{A}_k| = |\mathbf{A}_{k-1}| (a_{kk} - \mathbf{g}^T \mathbf{A}_{k-1}^{-1} \mathbf{f}) = |\mathbf{A}_{k-1}| \cdot u_{kk} \neq 0,$$

azaz $u_{kk} \neq 0$. Ezzel bebizonyítottuk, hogy $|\mathbf{U}_k| \neq 0$, tehát a trianguláris faktorizáció létezik. ■

Ezek szerint, ha nemszinguláris mátrix bal felső sarokminorai zérustól különbözőek, akkor a generáló elemek megfelelő megválasztásával biztosítható a mátrix trianguláris faktorizációja. Ennek felhasználásával most bebizonyítjuk a *sávmátrixok* trianguláris faktorizációjára vonatkozó következő tételt.

1.7.4 tétel. Ha $[a_{ij}]$ nemszinguláris sávmátrix, azaz

$$a_{ij} = 0, \text{ ha } j - i > p \text{ és } i - j > q,$$

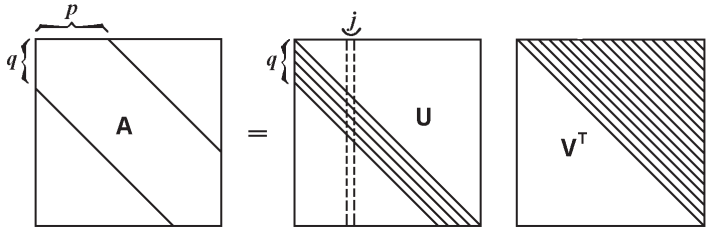
ahol p és q pozitív egész, továbbá a mátrix bal felső sarokminorai zérustól különbözőek, akkor a mátrix $\mathbf{A} = \mathbf{U}\mathbf{V}^T$ trianguláris faktorizációjában

$$u_{ij} = 0, \text{ ha } i - j > q, \text{ és } v_{ij} = 0, \text{ ha } j - i > p.$$

Bizonyítás. Mivel \mathbf{V}^T felső háromszögmátrix, ezért az \mathbf{U} mátrix j -edik oszlopa $\mathbf{U} = \mathbf{A}(\mathbf{V}^T)^{-1}$ miatt az \mathbf{A} mátrix első j oszlopának lineáris kombinációja. Tehát ugyanúgy, mint az \mathbf{A} mátrix j -edik oszlopát, az jellemzi, hogy

$$u_{ij} = 0, \text{ ha } i - j > q;$$

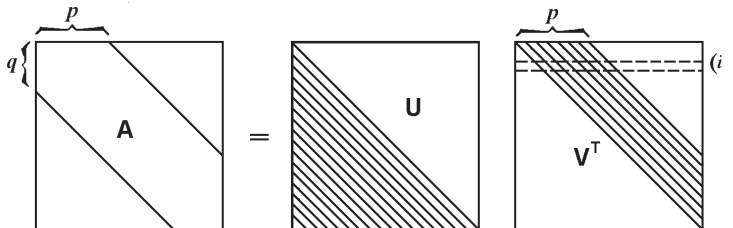
ez szemléltethető a mátrixok szerkezetének felvázolásával:



Hasonlóképpen, mivel \mathbf{U} alsó háromszögmátrix, ezért a \mathbf{V}^T mátrix i -edik sora $\mathbf{V}^T = \mathbf{U}^{-1}\mathbf{A}$ miatt az \mathbf{A} mátrix első i sorának lineáris kombinációja. Tehát az \mathbf{A} mátrix i -edik sorához hasonlóan ezt is az jellemzi, hogy

$$v_{ij} = 0, \text{ ha } j - i > p;$$

felvázolva:



Ezzel bebizonyítottuk, hogy a sávmátrixok sávszerkezetét a trianguláris faktorizációjuk valóban „öröklí”. ■

33. Példa. Határozzuk meg az alábbi kontinuáns mátrix trianguláris faktorizációját:

$$\begin{bmatrix} 1 & 2 & & & \\ 1 & 4 & 6 & & \\ & 1 & 6 & 12 & \\ & & 1 & 8 & 20 \\ & & & 1 & 10 \end{bmatrix}.$$

Megoldás. A faktorizáció táblázata:

$$\begin{bmatrix} 1 & \boxed{2} & 0 & 0 & 0 \\ 1 & 2 & \boxed{3} & 0 & 0 \\ 0 & 1 & 3 & \boxed{4} & 0 \\ 0 & 0 & 1 & 4 & \boxed{5} \\ 0 & 0 & 0 & 1 & 5 \end{bmatrix},$$

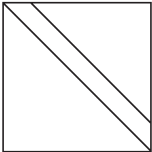
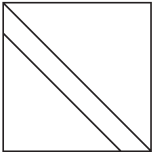
azaz

$$\mathbf{U} = \begin{bmatrix} 1 & & & & \\ 1 & 2 & & & \\ & 1 & 3 & & \\ & & 1 & 4 & \\ & & & 1 & 5 \end{bmatrix} \quad \text{és} \quad \mathbf{V}^T = \begin{bmatrix} 1 & 2 & & & \\ & 1 & 3 & & \\ & & 1 & 4 & \\ & & & 1 & 5 \\ & & & & 1 \end{bmatrix}.$$

* * *

Amint ebből a példából is kitűnik, a kontinuáns mátrixok trianguláris faktorizációja azt jelenti, hogy az alsó, ill. felső háromszögmátrix alakú tényezőben csak a főátlóban és az alsó, ill. felső kodiagonálisban lehetnek zérustól különböző elemek. Ezekre bevezetjük a *bidiagonális* mátrixok definícióját (lásd [9], [30]):

1.7.1 definíció. Ha egy mátrixnak csak a főátlójában és a felső (alsó) kodiagonálisában vannak zérustól különböző elemek, akkor azt felső (alsó) bidiagonális mátrixnak nevezzük.

Sematikusan vázolva: $\mathbf{B} =$  $\mathbf{B}^T =$ .

Ezek szerint tehát azt mondhatjuk, hogy egy kontinuáns mátrix mindig felbontható egy alsó és egy felső bidiagonális mátrix szorzatára és megfordítva:

egy alsó és egy felső bidiagonális mátrix szorzata mindig kontinuáns (tridia-gonális) mátrixot eredményez (ez a fordított sorrendben végzett szorzatra is igaz: egy felső és egy alsó bidiagonális mátrix szorzata is kontinuáns mátrixra vezet).

A Gauss–Crout-módszer alkalmazásával egyszerű rekurziós képletek adódnak a kontinuáns mátrix együtthatójú egyenletrendszerek megoldására. A műszaki és a természettudományok sok feladata vezet ilyen egyenletrend-szerre. Tekintsük tehát az

$$\begin{bmatrix} a_1 & b_1 & & & \\ c_1 & a_2 & b_2 & & \\ & c_2 & a_3 & b_3 & \\ & & & \ddots & \\ & & & c_{n-2} & a_{n-1} & b_{n-1} \\ & & & & c_{n-1} & a_n \end{bmatrix} \mathbf{x} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_{n-1} \\ d_n \end{bmatrix}$$

egyenletrendszert, amely röviden a

$$\begin{aligned} c_{i-1}x_{i-1} + a_i x_i + b_i x_{i+1} &= d_i, \quad 1 \leq i \leq n, \\ c_0 &= 0, \quad b_n = 0 \end{aligned}$$

alakban írható. Feltehető, hogy az együtthatómátrix irreducibilis, mert különben az egyenletrendszer szétesne kisebb méretű, ugyanilyen típusú egyenletrendszerekre. Tegyük fel továbbá, hogy az együtthatómátrix bal felső sarokminorai zérustól különbözőek. Ekkor az 1.7.4 tétel biztosítja, hogy a faktorizáció során megmarad a sávszerkezet, vagyis ebben az esetben a \mathbf{V}^T tényező felső bidiagonális mátrix, és az egyenletrendszer igen egyszerű rekur-zióval megoldható. Az (1.7.14) és (1.7.15) képletek szerint $n = 4$ esetére az alábbiak szerint végezhetjük el a faktorizációt.

$\begin{array}{ccccc} a_1 & b_1 & 0 & 0 & d_1 \\ c_1 & a_2 & b_2 & 0 & d_2 \\ 0 & c_2 & a_3 & b_3 & d_3 \\ 0 & 0 & c_3 & a_4 & d_4 \end{array}$	
$\begin{array}{ccccc} \beta_1 & \frac{b_1}{\beta_1} & 0 & 0 & \frac{d_1}{\beta_1} = \gamma_1 \\ c_1 & \beta_2 & \frac{b_2}{\beta_2} & 0 & \gamma_2 \\ 0 & c_2 & \beta_3 & \frac{b_3}{\beta_3} & \gamma_3 \\ 0 & 0 & c_3 & \beta_4 & \gamma_4 \end{array}$	$\begin{aligned} \beta_1 &= a_1, & \gamma_1 &= \frac{d_1}{\beta_1}, \\ \beta_2 &= a_2 - c_1 \frac{b_1}{\beta_1}, & \gamma_2 &= \frac{d_2 - c_1 \gamma_1}{\beta_2}, \\ \beta_3 &= a_3 - c_2 \frac{b_2}{\beta_2}, & \gamma_3 &= \frac{d_3 - c_2 \gamma_2}{\beta_3}, \\ \beta_4 &= a_4 - c_3 \frac{b_3}{\beta_3}, & \gamma_4 &= \frac{d_4 - c_3 \gamma_3}{\beta_4}. \end{aligned}$
$\begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & \end{array}$	

Általánosan (lásd [17])

$$(1.7.24) \quad \begin{aligned} \beta_1 &= a_1, & \beta_i &= a_i - \frac{c_{i-1}b_{i-1}}{\beta_{i-1}} \quad (i = 2, 3, \dots, n), \\ \gamma_1 &= \frac{d_1}{\beta_1}, & \gamma_i &= \frac{d_i - c_{i-1}\gamma_{i-1}}{\beta_i} \quad (i = 2, 3, \dots, n). \end{aligned}$$

A megoldásra pedig az alábbi rekurziós összefüggés adódik:

$$x_n = \gamma_n, \quad x_i = \gamma_i - x_{i+1} \frac{b_i}{\beta_i} \quad (i = n-1, \dots, 2, 1).$$

34. Példa. Oldjuk meg az alábbi egyenletrendszert a Gauss–Crout-féle eliminációs módszerrel:

$$\begin{aligned} x_1 + 2x_2 &= 3, \\ 2x_1 + 7x_2 + 12x_3 &= 15, \\ x_2 + 5x_3 + 2x_4 &= 4, \\ 3x_3 + 11x_4 + 15x_5 &= 13, \\ x_4 + 13x_5 &= 22. \end{aligned}$$

Megoldás. A feladat megoldását az alábbi táblázatos elrendezés szemlélteti:

$$\begin{array}{cccccc|c} 1 & 2 & 0 & 0 & 0 & 3 & 6 \\ 2 & 7 & 12 & 0 & 0 & 15 & 36 \\ 0 & 1 & 5 & 2 & 0 & 4 & 12 \\ 0 & 0 & 3 & 11 & 15 & 13 & 42 \\ 0 & 0 & 0 & 1 & 13 & 22 & 36 \\ \hline 3 & 10 & 20 & 14 & 28 & 57 & 132 \end{array}.$$

$$\begin{array}{cccccc|c} 1 & 2 & 0 & 0 & 0 & 3 & 6 \\ 2 & 3 & 4 & 0 & 0 & 3 & 8 \\ 0 & 1 & 1 & 2 & 0 & 1 & 4 \\ 0 & 0 & 3 & 5 & 3 & 2 & 6 \\ 0 & 0 & 0 & 1 & 10 & 2 & 3 \\ 3 & 4 & 4 & 6 & 10 & 0 & 0 \end{array}.$$

A kontrollszummák nem jeleznek hibát, tehát az alábbi faktorizációhoz jutottunk:

$$\begin{bmatrix} 1 & & & & & \\ 2 & 3 & & & & \\ & 1 & 1 & & & \\ & & 3 & 5 & & \\ & & & 1 & 10 & \end{bmatrix} \begin{bmatrix} 1 & 2 & & & & 3 \\ & 1 & 4 & & & 3 \\ & & 1 & 2 & & 1 \\ & & & 1 & 3 & 2 \\ & & & & 1 & 2 \end{bmatrix}.$$

Ennek alapján a redukált egyenletrendszer rekurzív módon megoldható:

$$\begin{aligned}x_5 &= 2, \\x_4 &= 2 - 3x_3 = -4, \\x_3 &= 1 - 2x_4 = 9, \\x_2 &= 3 - 4x_3 = -33, \\x_1 &= 3 - 2x_2 = 69.\end{aligned}$$

* * *

Mint már említettük, a kerekítési hibák halmozódása miatt a Gauss–Crout-módszerrel kapott megoldást a pontos megoldás *közelítésének* kell tekintenünk. Az alábbiakban olyan iterációs módszert ismertetünk, amellyel ez a közelítés finomítható, és segítségével elvileg tetszőleges – gyakorlatilag csupán a számítás elvégzéséhez felhasznált számítógéptől függő – pontosság elérhető.

Jelölje $\mathbf{x}^{(1)}$ az

$$(1.7.25) \quad \mathbf{Ax} = \mathbf{b}$$

egyenletrendszer megoldásának első közelítését, és $\mathbf{r}^{(1)}$ az

$$(1.7.26) \quad \mathbf{r}^{(1)} = \mathbf{b} - \mathbf{Ax}^{(1)}$$

különbséggel definiált hibavektort. Ekkor az

$$(1.7.27) \quad \mathbf{Au} = \mathbf{r}^{(1)}$$

egyenlet megoldása az $\mathbf{x}^{(1)}$ közelítés korrekcióját adja, mert az (1.7.27) egyenletet behelyettesítve az (1.7.26) egyenletbe, (1.7.25) figyelembevételével a pontos megoldásra

$$(1.7.28) \quad \mathbf{x} = \mathbf{x}^{(1)} + \mathbf{u}$$

adódik. Az (1.7.27) egyenlet megoldása viszont az \mathbf{A} mátrix már ismert faktorizációja segítségével nyerhető, pontosabban ezzel az (1.7.27) egyenlet megoldásának egy $\mathbf{u}^{(1)}$ közelítését kapjuk. Várható, hogy az így számított

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \mathbf{u}^{(1)}$$

az adott (1.7.25) egyenletnek egy jobb közelítése lesz, mint $\mathbf{x}^{(1)}$. Az eljárást ezután megismételjük, vagyis képezzük az

$$\mathbf{r}^{(k)} = \mathbf{b} - \mathbf{Ax}^{(k)},$$

$$\begin{aligned}\mathbf{A}\mathbf{u}^{(k)} &= \mathbf{r}^{(k)}, \\ \mathbf{x}^{(k+1)} &= \mathbf{x}^{(k)} + \mathbf{u}^{(k)}\end{aligned}$$

iterációval definiált $\mathbf{x}^{(k)}$ vektorsorozatot. Amennyiben az $\mathbf{x}^{(1)}$ első közelítés „elég” közel van a pontos megoldáshoz, akkor az így nyert $\mathbf{x}^{(k)}$ vektorsorozat konvergens és a határértéke az adott egyenletrendszer pontos megoldása.

Megjegyzés. A konvergencia vizsgálata meghaladja e könyv kereteit, ezért arra itt nem térhetünk ki.

Az eljárás szemléltetésére oldjuk meg a következő feladatot.

35. Példa. Határozzuk meg az

$$\begin{bmatrix} 1,0303 & 0,99030 \\ 0,99030 & 0,95285 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2,4944 \\ 2,3988 \end{bmatrix}$$

egyenletrendszer egy közelítő megoldását és finomítsuk utóiterációs módszerrel.

Megoldás. Az adott egyenletrendszer megoldása nyolc tizedesjegyre kerekítve:

$$\mathbf{A}^{-1}\mathbf{b} = \begin{bmatrix} 1,22402691 \\ 1,24536512 \end{bmatrix}.$$

Az adott egyenletrendszeren elvégezve a Gauss-eliminációt, az alábbi mátrixot kapjuk:

$$\left[\begin{array}{cc|c} 1,0303 & 0,99030 & 2,4944 \\ 0 & 0,00099 & 0,0012 \end{array} \right].$$

Ebből kiszámíthatjuk az $\mathbf{x}^{(1)}$ és $\mathbf{r}^{(1)}$ vektorokat:

$$\begin{aligned}\mathbf{x}^{(1)} &= \begin{bmatrix} 1,2560 \\ 1,2121 \end{bmatrix}, \\ \mathbf{r}^{(1)} &= \mathbf{b} - \mathbf{A}\mathbf{x}^{(1)} = \begin{bmatrix} 0,00000057 \\ 0,00003371 \end{bmatrix}.\end{aligned}$$

Az $\mathbf{A}\mathbf{u}^{(1)} = \mathbf{r}^{(1)}$ egyenletrendszer megoldásából

$$\mathbf{u}^{(1)} = \begin{bmatrix} -0,0322 \\ 0,0335 \end{bmatrix} \quad \text{és} \quad \mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \mathbf{u}^{(1)} = \begin{bmatrix} 1,2238 \\ 1,2456 \end{bmatrix}$$

adódik. A számításokat tovább folytatva:

$$\mathbf{r}^{(2)} = \begin{bmatrix} 0,00000118 \\ 0,00000090 \end{bmatrix},$$

$$\mathbf{u}^{(2)} = \begin{bmatrix} 0,0002 \\ -0,0002 \end{bmatrix},$$

$$\mathbf{x}^{(3)} = \mathbf{x}^{(2)} + \mathbf{u}^{(2)} = \begin{bmatrix} 1,2240 \\ 1,2454 \end{bmatrix},$$

majd

$$\mathbf{r}^{(3)} = \begin{bmatrix} -0,00000682 \\ -0,00000659 \end{bmatrix},$$

$$\mathbf{u}^{(3)} = \begin{bmatrix} 0,00002717 \\ -0,00003515 \end{bmatrix}.$$

Az $\mathbf{x}^{(4)} = \mathbf{x}^{(3)} + \mathbf{u}^{(3)}$ és az $\mathbf{x}^{(3)}$ közelítések egybevetéséből látható, hogy azok már az első öt tizedesjegyre megegyeznek, mivel $\mathbf{u}^{(3)}$ elemeinek abszolút értéke kisebb, mint $0,5 \cdot 10^{-5}$. Itt általában meg szoktunk állni. További javítás csak úgy érhető el, ha az $\mathbf{x}^{(k)}$ vektorból több jegyet ismerünk.

* * *

(b) **Szinguláris együtthatómátrix.** Abban az esetben, ha az egyenletrendszer együtthatómátrixa szinguláris kvadratikusan mátrix (vagyis $r < n$), akkor az általános tárgyalás során követett módszert alkalmazzuk, tehát az egyenletrendszert az együtthatómátrix minimális diadikus felbontásának segítségével redukáljuk.

Ezen belül egy speciális esettel azonban érdemes külön foglalkozni, mégpedig azzal az esettel, amikor a *homogén lineáris egyenletrendszer együtthatómátrixának rangja eggyel kisebb, mint a rendszáma*. Tekintsük tehát az

$$(1.7.29) \quad \mathbf{Ax} = \mathbf{0} \quad (|\mathbf{A}| = 0; \varrho(\mathbf{A}) = n - 1)$$

egyenletrendszert. Írjuk fel az (1.2.27) azonosságot, és vegyük figyelembe, hogy \mathbf{A} szinguláris; ekkor az

$$(1.7.30) \quad \mathbf{A} \operatorname{adj} \mathbf{A} = \mathbf{0}$$

összefüggést kapjuk, tehát a Sylvester-féle nullitási tételből (lásd az 1.3.13 tételt) következik, hogy

$$\varrho(\mathbf{A}) + \varrho(\operatorname{adj} \mathbf{A}) \leq n,$$

azaz $\varrho(\mathbf{A}) = n - 1$ miatt $\varrho(\operatorname{adj} \mathbf{A}) \leq 1$, de ha $\varrho(\operatorname{adj} \mathbf{A}) = 0$ lenne, az azt jelentené, hogy \mathbf{A} rangja $n - 2$, ami ellentmondásra vezetett. Ezért

$$(1.7.31) \quad \varrho(\operatorname{adj} \mathbf{A}) = 1,$$

és így adj \mathbf{A} oszlopai csak állandó szorzóban különbözhetnek egymástól. Tehát az ismeretlenek az adj \mathbf{A} mátrix bármely oszlopának elemeivel arányosak, vagyis

$$(1.7.32) \quad x_k = tA_{ik} \quad (k = 1, 2, \dots, n; i = 1, 2, \dots, n);$$

vagy – más megfogalmazásban – az ismeretlenek aránya egyenlő az együtthatómátrix bármely sorának elemeihez tartozó előjeles aldeterminánsok arányával (feltéve, hogy közülük legalább egy zérustól különböző):

$$(1.7.33) \quad x_1 : x_2 : \dots : x_n = A_{i1} : A_{i2} : \dots : A_{in} \quad (i = 1, 2, \dots, n).$$

A szinguláris együtthatómátrixok körében különleges helyet foglalnak el a projektorok. Az alábbiakban megmutatjuk, hogy az ilyen egyenletek megoldása közvetlenül felírható.

Tekintsük tehát a

$$(1.7.34) \quad \mathbf{P}\mathbf{x} = \mathbf{b}$$

egyenletrendszert, ahol \mathbf{P} egy n -edrendű és r -edrangú projektor. Megoldását az (1.7.9) képlet alapján a homogén egyenlet általános és az inhomogén egyenlet egy partikuláris megoldásának összegeként állítjuk elő.

A *homogén* egyenlet,

$$(1.7.35) \quad \mathbf{P}\mathbf{x} = \mathbf{0},$$

általános megoldását a projektorok

$$(1.7.36) \quad \mathbf{P}(\mathbf{E} - \mathbf{P}) = \mathbf{0}$$

definiáló egyenletéből közvetlenül megkaphatjuk. Az (1.7.35) és (1.7.36) egyenletek egybevetéséből ugyanis látható, hogy $(\mathbf{E} - \mathbf{P})$ bármely oszlopa kielégíti a homogén egyenletet. Mivel az 1.6.2 tétel értelmében $\varrho(\mathbf{P}) = r$ esetén $\varrho(\mathbf{E} - \mathbf{P}) = n - r$, tehát az $\mathbf{E} - \mathbf{P}$ mátrixnak van $n - r$ lineárisan független oszlopvektora, és ezeknek lineáris kombinációja a homogén egyenlet általános megoldását adja. A tetszőleges elemű \mathbf{t} paramétervektorral az általános megoldás

$$\mathbf{x} = (\mathbf{E} - \mathbf{P})\mathbf{t}$$

alakban írható.

Az *inhomogén* egyenlet, $\mathbf{P}\mathbf{x} = \mathbf{b}$, egy partikuláris megoldása, ha létezik,

$$\mathbf{x} = \mathbf{b};$$

ugyanis az (1.7.34) egyenletet balról szorozva a \mathbf{P} mátrixszal, $\mathbf{P}^2\mathbf{x} = \mathbf{Pb}$ adódik, és $\mathbf{P}^2 = \mathbf{P}$ miatt a bal oldalak azonosságából következik, hogy

$$(1.7.37) \quad \mathbf{Pb} = \mathbf{b},$$

vagyis $\mathbf{x} = \mathbf{b}$ valóban kielégíti az egyenletet.

Az (1.7.37) összefüggésnek ezért teljesülnie kell adott \mathbf{P} és \mathbf{b} mellett, ha az egyenletrendszernek egyáltalán létezik megoldása, vagyis (1.7.37) az egyenletrendszer *kompatibilitási feltétele*.

Összegezve: az (1.7.34) inhomogén egyenletnek akkor és csak akkor van megoldása, ha az (1.7.37) összefüggés teljesül. Az általános megoldás pedig – a homogén egyenlet általános megoldásának és az inhomogén egyenlet egy partikuláris megoldásának összegeként felírva – a következő:

$$(1.7.38) \quad \mathbf{x} = \mathbf{b} + (\mathbf{E} - \mathbf{P})\mathbf{t}.$$

Ezek szerint az olyan lineáris egyenletrendszerek megoldása, amelyeknek együtthatómátrixa projektor, igen egyszerű. Tehát ha sikerül az $\mathbf{Ax} = \mathbf{0}$ egyenletrendszer együtthatómátrixát olyan

$$\mathbf{A} = \mathbf{UP}$$

szorzat alakjában felírni, ahol \mathbf{P} projektor és az $\mathbf{UPx} = \mathbf{0}$ egyenletből

$$\mathbf{Px} = \mathbf{0}$$

következik, akkor ezzel a feladat lényegében megoldottnak tekinthető.

2. fejezet

A LINEÁRIS ALGEBRA ALAPJAI

*ARKHIMÉDÉSZ: A matematika csak azoknak enged bepillantást titkaiba,
akik szépségéért lelkesedve, tiszta tudásvágygal közelednek hozzá.
(Rényi Alfréd: Dialógus a matematika alkalmazásairól)*

Ebben a fejezetben bevezetjük az Olvasót a lineáris algebra absztrakt tárgyalásmódjába. Mint azt látjuk majd, a lineáris algebra az algebrának, a geometriának, az analízisnek számos fogalmát elvonatkoztatja konkrét tartalmától, és megragadva azt, ami bennük közös, olyan általános tételeket mond ki és bizonyít, amelyek, éppen az említett absztrakció révén, lehetővé teszik a matematika látszólag egymástól távol eső területein fellépő problémák egységes tárgyalását (lásd pl. [6], [8]).

Első olvasásra talán úgy tűnik, hogy néhány ismert fogalom definíciójának szerepeltetése (pl. vektor hossza, két vektor által bezárt szög stb.) felesleges ezen a helyen. Fel kell azonban hívni a figyelmet arra, hogy éppen az általános és absztrakt tárgyalásmód teszi szükségessé, hogy az ismert fogalmakat is nagyon pontosan újra definiáljuk. A korábban szerzett ismeretek pedig lehetővé teszik, hogy az absztrakt fogalmakat és tételeket – többnyire a geometriai szemlélet segítségével – valamilyen konkrét megnyilvánulási formájukhoz kapcsoljuk.

Először definiáljuk a lineáris tér, a dimenzió, a bázis és a koordináták fogalmát, majd a skaláris szorzat értelmezésével bevezetjük az euklideszi tér fogalmát. Ezután röviden kitérünk a bilineáris és kvadratikussá alakokra, majd részletesen foglalkozunk a lineáris transzformációk elméletével. Megmutatjuk, hogy a mátrixok alkotják a bilineáris alakok és a lineáris transzformációk tárgyalásának matematikai apparátusát, majd külön pontban megvizsgáljuk a bázis transzformációjának hatását. Ezzel lehetővé válik, hogy a lineáris transzformációknak a koordináta-rendszer (vagyis a bázis) megválasztásától független, invariáns tulajdonságait vizsgáljuk. Így jutunk a lineáris algebra legfontosabb problémaköréhez, az ún. sajátérték-feladathoz. Végül a fejezet utolsó pontjában, az adjungált transzformáció fogalmának bevezetése után, a lineáris transzformációk néhány speciális osztályának tulajdonságait elemezzük.

2.1 A LINEÁRIS TÉR

A matematika különböző ágaiban igen gyakran találkozunk olyan fogalmakkal, amelyekre az összeadás és a számokkal való szorzás értelmezve van.

Így pl. a geometriában a háromdimenziós tér két vektorának összegét – mint ismeretes – a paralelogramma-szabállyal értelmezzük, egy vektornak egy számmal képezett szorzatán pedig a vektor megfelelő nyújtását, esetleg irányításának megváltoztatását értjük. Az algebraiban két, n számból álló rendszer (rendezett szám- n -es) összegét úgy értelmezzük, hogy az azonos sorszámú helyen álló elemeket összeadjuk, a számmal szorzást pedig úgy, hogy a szám- n -es minden egyes elemét megszorozzuk a számmal. Az analízisben értelmezik pl. a zárt intervallumban folytonos függvények összeadását és számmal szorzását.

Mint e példákból látható, az összeadás és a számmal való szorzás egészen különböző fogalmakra értelmezhető. Egységes tárgyalásuk és szemléletük kedvéért bevezetjük a *lineáris tér* fogalmát.

2.1.1 definíció. *Legyenek az $\alpha, \beta, \dots, \lambda, \dots$ számok egy K számtest elemei*, $\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots$ pedig egy \mathbf{R} halmaz elemei. Az \mathbf{R} halmazt **lineáris térnek** (vektortérnek, gyakran röviden csak **térnek**) nevezzük, ha*

(1) *bármely két $\mathbf{x}, \mathbf{y} \in \mathbf{R}$ elemhez egyértelműen hozzá van rendelve a halmaznak egy eleme, amelyet az \mathbf{x} és \mathbf{y} elem összegének nevezünk és $\mathbf{x} + \mathbf{y}$ -nal jelölünk; az összeadás tulajdonságai:*

- (a) *kommutatív: $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$;*
- (b) *asszociatív: $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$;*
- (c) *létezik olyan $\mathbf{0}$ zéruselem, hogy $\mathbf{x} + \mathbf{0} = \mathbf{x}$ minden \mathbf{x} -re;*
- (d) *létezik minden \mathbf{x} -hez a $-\mathbf{x}$ (inverz) elem úgy, hogy $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$;*

* *Testnek* nevezzük azt a legalább kételemű halmazt, amelyben értelmezve van egy összeadásnak és egy szorzásnak nevezett kétváltozós (bináris) művelet, a következő tulajdonságokkal:

(1) A test bármely két elemének van egyértelműen meghatározott összege, amely szintén a test eleme; az összeadás kommutatív és asszociatív; az összeadás inverz művelete (kivonás) egyértelműen elvégezhető a testen belül – vagyis a különbség is a test eleme; létezik egy 0-val jelölt elem, amelyet bármely elemhez hozzáadva, azt változatlanul hagyja. (Röviden: A test az összeadás műveletére nézve Abel-csoportot alkot a 0 neutrális elemmel.)

(2) A test bármely két elemének van egyértelműen meghatározott szorzata a testen; a szorzás kommutatív és asszociatív: a szorzás inverz művelete (osztás) a 0-val való osztás kivételével egyértelműen elvégezhető a testen belül; létezik egy 1-gyel jelölt elem, amellyel bármely elemet megszorozva az nem változik meg. (Röviden: A test 0-tól különböző elemei multiplikatív csoportot alkotnak az 1 egységelemmel.)

(3) Az összeadásra és szorzásra érvényes a disztributivitási törvény. Belátható, hogy testet alkotnak pl. a közönséges összeadás és szorzás műveletével a racionális számok, a valós számok, a komplex számok. Így beszélhetünk a racionális számtestről, a valós számtestről, a komplex számtestről.

(2) bármely $x \in R$ elemhez és bármely $\lambda \in K$ számhoz egyértelműen hozzá van rendelve az R halmaznak egy eleme, amelyet a λ szám és az x elem szorzatának nevezünk és λx -szel jelölünk; a számmal való szorzás tulajdonságai:

(a) a számtest 1 egységelemének és az x elemnek szorzata maga az x elem:

$$1 \cdot x = x;$$

(b) több számmal végzett szorzásra fennáll, hogy $\alpha(\beta x) = \alpha\beta(x)$.

(3) a számmal való szorzás és az elemek összeadása disztributív:

$$(a) (\alpha + \beta)x = \alpha x + \beta x; \quad (b) \alpha(x + y) = \alpha x + \alpha y.$$

Ha K a valós számtest, akkor a teret **valós lineáris térnek**, ha pedig a komplex számtest, akkor **komplex lineáris térnek** nevezzük.

A lineáris tér elemeinek összeadásától és számmal való szorzásától csupán azt kívánjuk meg, hogy a fenti tulajdonságoknak eleget tegyenek, de hogy miként van értelmezve az összeadás és a számmal való szorzás, arra nézve nem teszünk semmilyen megkötést.

Nem nehéz meggyőződni arról, hogy pl. a komplex számtest felett értelmezett legfeljebb n -edfokú polinomok halmaza lineáris teret alkot a szokásos összeadásra és számmal való szorzásra nézve, de a pontosan n -edfokú polinomok halmaza már nem alkot lineáris teret, mivel két n -edfokú polinom összege n -nél alacsonyabb fokú polinom is lehet. Ugyancsak lineáris teret alkotnak a síkvektorok a szokásos vektorösszeadás és számmal való szorzás műveletére nézve.

A lineáris tér elemeit *vektoroknak* is nevezzük.

Ezt az elnevezést, valamint a lineáris algebra számos más elnevezését, amelyek szintén a geometriából származnak, az indokolja, hogy a lineáris algebra legegyszerűbb példáit a geometria szolgáltatja. Éppen ezért a lineáris algebra számos absztrakt fogalmának és tételének jobb megértéséhez is hozzásegít a geometria szemléletessége, amire sokszor fogunk hivatkozni. A félreértések elkerülése végett ilyenkor „geometriai térről” fogunk beszélni.

A következőkben bevezetjük a vektorok lineáris összefüggésének és függetlenségének, valamint a lineáris tér dimenziójának fogalmát*.

2.1.2 definíció. Az R lineáris tér $x_1, x_2, x_3, \dots, x_n$ vektorait akkor mondjuk **lineárisan függetleneknek**, ha az

$$(2.1.1) \quad \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n = \mathbf{0}$$

összefüggés csak $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$ esetén áll fenn.

Ha viszont található olyan $\alpha_1, \alpha_2, \dots, \alpha_n$ értékrendszer, hogy az α_i értékek közül legalább egy zérustól különböző, és a (2.1.1) összefüggés fennáll, akkor az x_1, x_2, \dots, x_n vektorok **lineárisan összefüggőek**.

*Vö. az 1.3.4 definícióval.

A

$$(2.1.2) \quad \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \beta_3 \mathbf{x}_3 + \cdots + \beta_n \mathbf{x}_n$$

kifejezést az $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n$ vektorok **lineáris kombinációjának** nevezzük.

2.1.1 tétel. Ha az \mathbf{R} lineáris tér $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ vektorai lineárisan összefüggőek, akkor közülük legalább egy kifejezhető a többiek lineáris kombinációjaként.

Bizonyítás. Ha az $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ vektorok lineárisan összefüggőek, akkor teljesül a (2.1.1) egyenlőség úgy, hogy pl. $\alpha_1 \neq 0$. Fejezzük ki az \mathbf{x}_1 vektort a (2.1.1) egyenletből:

$$\mathbf{x}_1 = -\frac{\alpha_2}{\alpha_1} \mathbf{x}_2 - \frac{\alpha_3}{\alpha_1} \mathbf{x}_3 \cdots - \frac{\alpha_n}{\alpha_1} \mathbf{x}_n;$$

innen a

$$\beta_2 = -\frac{\alpha_2}{\alpha_1}; \quad \beta_3 = -\frac{\alpha_3}{\alpha_1}; \quad \dots; \quad \beta_n = -\frac{\alpha_n}{\alpha_1}$$

jelölésekkel láthatjuk, hogy az \mathbf{x}_1 vektor kifejezhető az $\mathbf{x}_2, \dots, \mathbf{x}_n$ vektorok lineáris kombinációjaként. ■

2.1.3 definíció. Egy \mathbf{R} lineáris teret **n -dimenziós**nak nevezünk, ha létezik benne n lineárisan független vektor, de nem létezik ennél nagyobb számú lineárisan független vektor. Jelölje ezt a teret \mathbf{R}_n .

Például a geometriai tér vektorai között lehet találni három lineárisan független vektort, de bármely négy – vagy több – vektor már lineárisan összefüggő. Ezért a geometriából ismert háromdimenziós tér a lineáris algebra értelmezése szerint is háromdimenziós!

Algebrai példaként a rendezett szám- n -esek terében meg lehet adni n lineárisan független vektort, pl.:

$$\begin{aligned} &(1, 0, \dots, 0); \\ &(0, 1, \dots, 0); \\ &\dots\dots\dots \\ &(0, 0, \dots, 1); \end{aligned}$$

de bármely m vektor, ahol $m > n$, már lineárisan összefüggő. Tehát a rendezett szám- n -esek tere n -dimenziós.

A legfeljebb n -edfokú polinomok terében megadható $n + 1$ lineárisan független polinom, pl.: $1, t, t^2, \dots, t^n$, de meg lehet mutatni, hogy ennél több nem található. Ezért ez a tér $(n + 1)$ -dimenziós.

Megjegyzés. Nem nehéz példát megadni olyan lineáris térre, amelyikben bármilyen N esetén található N -nél több lineárisan független vektor. (Ilyen pl. a $[0, 1]$ intervallumban

folytonos függvények tere.) Az ilyen lineáris teret végtelen dimenziósnek nevezzük. E könyv keretei között azonban csak véges dimenziójú lineáris terekkel foglalkozunk.

2.1.4 definíció. Az n -dimenziós R_n lineáris tér bármely n számú lineárisan független e_1, e_2, \dots, e_n vektorából álló rendszert a tér egy **bázisának**, e vektorok mindegyikét pedig **bázisvektornak** nevezzük.

Például geometriai térben bázist alkot bármely három, nem egy síkban fekvő vektor.

2.1.2 tétel. Az n -dimenziós R_n lineáris tér bármely x vektora egyértelműen előállítható a bázis vektoraival lineáris kombinációjaként.

Bizonyítás. Legyen az n -dimenziós lineáris tér egy bázisa e_1, e_2, \dots, e_n és vegyük hozzá az x vektort. Mivel ez összesen $n + 1$ vektor, ezért lineárisan összefüggők, vagyis az

$$(2.1.3) \quad \alpha_0 x + \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n = 0$$

egyenlőségben nem lehet mindegyik α_i zérus. Mivel α_0 biztosan nem zérus, mert különben (2.1.3) alapján az e_1, e_2, \dots, e_n vektorok nem lennének lineárisan függetlenek, az x vektor tehát kifejezhető a (2.1.3) egyenletből:

$$x = -\frac{\alpha_1}{\alpha_0} e_1 - \frac{\alpha_2}{\alpha_0} e_2 - \dots - \frac{\alpha_n}{\alpha_0} e_n.$$

Az egyértelműséget indirekt módon bizonyítjuk. Tegyük fel, hogy az x vektornak létezik két előállítása a bázis vektoraival:

$$x = \xi_1 e_1 + \xi_2 e_2 + \dots + \xi_n e_n, \quad \text{és} \quad x = \xi'_1 e_1 + \xi'_2 e_2 + \dots + \xi'_n e_n.$$

Képezzük e két összefüggés különbségét:

$$(\xi_1 - \xi'_1) e_1 + (\xi_2 - \xi'_2) e_2 + \dots + (\xi_n - \xi'_n) e_n = 0.$$

Az e_1, e_2, \dots, e_n vektorok lineárisan függetlenek, amiből következik, hogy

$$\xi_1 - \xi'_1 = \xi_2 - \xi'_2 = \dots = 0,$$

tehát

$$\xi_1 = \xi'_1; \quad \xi_2 = \xi'_2; \quad \dots; \quad \xi_n = \xi'_n,$$

és ezzel az előállítás egyértelműségét bebizonyítottuk. ■

2.1.5 definíció. Ha e_1, e_2, \dots, e_n az n -dimenziós tér egy bázisa, és

$$(2.1.4) \quad x = \xi_1 e_1 + \xi_2 e_2 + \dots + \xi_n e_n,$$

akkor a $\xi_1, \xi_2, \dots, \xi_n$ számokat az \mathbf{x} vektor $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisra vonatkozó koordinátáinak nevezzük.

Ez a definíció összhangban áll azzal, ahogyan az analitikus geometriában egy vektor (nem feltétlenül derékszögű koordináta-rendszerben vett) koordinátáit definiáljuk.

Adott \mathbf{x} vektor koordinátái tehát a bázis megválasztásától függenek.

A koordináták definíciójából következik, hogy a vektorok összeadásakor a koordináták összeadódnak, valamint hogy a $\lambda \mathbf{x}$ vektor koordinátái az \mathbf{x} vektor megfelelő koordinátáinak λ -szorosai.

A legfeljebb n -edfokú polinomok terében például egy

$$P(t) = a_0 + a_1 t + a_2 t^2 + \dots + a_n t^n$$

polinom koordinátái az

$$\mathbf{e}_1 = 1; \quad \mathbf{e}_2 = t; \quad \mathbf{e}_3 = t^2; \quad \dots; \quad \mathbf{e}_{n+1} = t^n$$

bázisra vonatkozóan az $a_0, a_1, a_2, \dots, a_n$ együtthatók. Ha viszont tekintjük az

$$\tilde{\mathbf{e}}_1 = 1; \quad \tilde{\mathbf{e}}_2 = t - a; \quad \tilde{\mathbf{e}} = (t - a)^2; \quad \dots; \quad \tilde{\mathbf{e}}_{n+1} = (t - a)^n$$

vektorokat, melyekről könnyű belátni, hogy a legfeljebb n -edfokú polinomok terében bázist alkotnak, akkor a $P(t)$ polinom koordinátái erre a bázisra vonatkozóan a megfelelő Taylor-polinom együtthatói. Ugyanis a $P(t)$ polinomot az a hely körül sorbafejtve,

$$P(t) = P(a) + P'(a)(t-a) + \frac{P''(a)}{2!}(t-a)^2 + \dots + \frac{P^{(n)}(a)}{n!}(t-a)^n,$$

tehát a koordináták most valóban a következők:

$$P(a); \quad P'(a); \quad \frac{P''(a)}{2!}; \quad \dots; \quad \frac{P^{(n)}(a)}{n!}.$$

Az eddig vizsgált feladatokban előfordultak olyan lineáris terek, amelyek a vizsgált tulajdonságok szempontjából (például bázis, dimenzió, koordináták) nem különböztek egymástól. Most megadjuk annak a definícióját, hogy milyen lineáris tereket tekintünk majd egymással „egyenrangúnak”, azaz *izomorf*nak.

2.1.6 definíció. Az \mathbf{R} és az \mathbf{R}' lineáris teret **izomorf**nak nevezzük, ha $\mathbf{x} \in \mathbf{R}$ és $\mathbf{x}' \in \mathbf{R}'$ vektorok között olyan kölcsönösen egyértelmű megfeleltetést lehet létrehozni, amely kielégíti a következő feltételeket: ha az \mathbf{x} vektornak az \mathbf{x}' vektor, az \mathbf{y} vektornak az \mathbf{y}' vektor felel meg, akkor

$$(2.1.5) \quad \text{az } \mathbf{x} + \mathbf{y} \text{ vektornak az } \mathbf{x}' + \mathbf{y}' \text{ vektor,}$$

(2.1.6) $a \lambda \mathbf{x}$ vektornak a $\lambda \mathbf{x}'$ vektor

felel meg.

Könnyen meggyőződhetünk arról, hogy különböző dimenziójú lineáris terek nem lehetnek izomorfak egymással (lásd [8]).

2.1.3 tétel. *Adott n esetén az n -dimenziós terek izomorfak.*

Bizonyítás. Legyen \mathbf{R} és \mathbf{R}' két n -dimenziós lineáris tér. Válasszunk az \mathbf{R} -ben egy $\mathbf{e}_1, \dots, \mathbf{e}_n$ bázist, és \mathbf{R}' -ben egy $\mathbf{e}'_1, \dots, \mathbf{e}'_n$ bázist. Tekintsünk egy tetszőleges \mathbf{x} vektort az \mathbf{R} lineáris térben:

$$(2.1.7) \quad \mathbf{x} = \xi_1 \mathbf{e}_1 + \dots + \xi_n \mathbf{e}_n.$$

Feleltessük meg az \mathbf{x} vektornak az

$$\mathbf{x}' = \xi_1 \mathbf{e}'_1 + \dots + \xi_n \mathbf{e}'_n$$

vektort az \mathbf{R}' lineáris térben. Ez a megfeleltetés kölcsönösen egyértelmű, hiszen az \mathbf{x} vektor (2.1.7) előállítása egyértelmű. Ezért a ξ_i ($i = 1, \dots, n$) számok, és így az \mathbf{x}' vektor is egyértelműen meghatározott. Hasonló megfontolással adódik, hogy minden \mathbf{x}' vektornak egy és csak egy \mathbf{x} vektor felel meg. Ezek alapján (2.1.5) és (2.1.6) teljesül, tehát az \mathbf{R} és az \mathbf{R}' lineáris tér izomorf egymással. ■

Ebből a tételből következik, hogy véges dimenziójú lineáris tér egyetlen lényeges adata a dimenziója.

2.1.7 definíció. *Az \mathbf{R} tér $\mathbf{R}^{(1)}$ alterének nevezzük \mathbf{R} elemeinek olyan részhalmazát, amely az \mathbf{R} -ben értelmezett összeadás és számmal való szorzás műveletére maga is lineáris teret alkot.*

Ha például \mathbf{R} a geometriai tér és egy olyan síkot tekintünk, amely átmegy a koordináta-rendszer kezdőpontján, akkor e sík vektorainak összessége az \mathbf{R} tér egy $\mathbf{R}^{(1)}$ alterét alkotja.

A rendezett szám- n -esek terében alteret alkotnak pl. mindazok a $(\xi_1, \xi_2, \dots, \xi_n)$ vektorok, amelyekre $\xi_1 = 0$. Belátható, hogy általánosabban ugyanezen lineáris térben alteret alkotnak az összes olyan vektorok, amelyekre teljesül a $c_1 \xi_1 + c_2 \xi_2 + \dots + c_n \xi_n = 0$ feltétel, ahol c_1, c_2, \dots, c_n rögzített számok.

Minden \mathbf{R} térben alteret képezhetünk úgy, hogy kiválasztjuk az \mathbf{R} tér tetszőleges $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r$ vektorait és képezzük ezek lineáris kombinációit. Az így nyert $\mathbf{R}^{(1)}$ alteret az $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r$ vektorok által generált alternek nevezzük. Lineárisan független, k számú $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k$ vektor által generált $\mathbf{R}^{(1)}$ alterné k -dimenziós, és az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k$ vektorok benne bázist alkotnak.

2.2 AZ EUKLIDESZI TÉR

Mint már hangsúlyoztuk, a lineáris terek legszemléletesebb példája a geometriai tér. Mivel azonban a lineáris tér definíciójához csak a vektorok összeadásának és számmal való szorzásának értelmezésére van szükség, ez nem teszi még lehetővé, hogy a geometria számos egyszerű fogalmának és feladatának megfelelő általánosítását elvégezzük. Ez indokolja, hogy a geometriából ismert skaláris szorzat fogalmát bevezessük a lineáris terek elméletében is. A skaláris szorzat segítségével definiáljuk az euklideszi teret, amelyben lehetővé válik a vektorok hosszának, két pont távolságának, valós euklideszi térben két vektor szögének az értelmezése. Bevezetjük az ortonormált bázis fogalmát és rámutatunk annak néhány előnyös tulajdonságára. Előljáróban is hangsúlyozni kívánjuk, hogy fogalmainkat általában a komplex lineáris térben értelmezzük és egyes esetekben rámutatunk arra, hogy ezek hogyan módosulnak valós lineáris tér alkalmazása esetén.

2.2.1 definíció. Az \mathbf{R} komplex lineáris térben **skaláris szorzat*** van értelmezve, ha a tér bármely két $\mathbf{x}, \mathbf{y} \in \mathbf{R}$ vektorához hozzárendelünk egy (\mathbf{x}, \mathbf{y}) -*nal jelölt, \mathbf{x} és \mathbf{y} skaláris szorzatának nevezett komplex számot a következő tulajdonságokkal:*

- (2.2.1) (a) $(\mathbf{x}, \mathbf{y}) = \overline{(\mathbf{y}, \mathbf{x})}$;
 (b) $(\lambda \mathbf{x}, \mathbf{y}) = \lambda(\mathbf{x}, \mathbf{y})$ (ahol λ komplex szám);
 (c) $(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) = (\mathbf{x}_1, \mathbf{y}) + (\mathbf{x}_2, \mathbf{y})$;
 (d) $(\mathbf{x}, \mathbf{x}) > 0$, ha $\mathbf{x} \neq \mathbf{0}$ és $(\mathbf{x}, \mathbf{x}) = 0$, ha $\mathbf{x} = \mathbf{0}$.

Az (a) tulajdonságból következik, hogy (\mathbf{x}, \mathbf{x}) mindig valós.

Az (a) és (b) tulajdonságból következik, hogy ha a skaláris szorzat második tényezőjéből emeljük ki a skalár szorzót, akkor az a kiemelés során konjugáltjára változik. Ugyanis

$$(2.2.2) \quad (\mathbf{x}, \lambda \mathbf{y}) = \overline{(\lambda \mathbf{y}, \mathbf{x})} = \bar{\lambda} \overline{(\mathbf{y}, \mathbf{x})} = \bar{\lambda}(\mathbf{x}, \mathbf{y}).$$

Valós lineáris térben a skaláris szorzatot valós számként értelmezzük és így az (a) tulajdonság az egyszerű

$$(2.2.3) \quad (\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x})$$

alakot veszi fel.

A skaláris szorzat értelmezésével vezetjük be az euklideszi tér fogalmát.

2.2.2 definíció. Az olyan lineáris teret, amelyben a (2.2.1) feltételi egyenleteket kielégítő szorzat van értelmezve, **komplex euklideszi** térnek** vagy

*A skaláris szorzatot szokták *belső szorzatnak* is nevezni.

**Eukleidész (i. e. 365–300) alexandriai matematikus.

unitér térnek nevezzük és \mathbf{E} -vel jelöljük. Ha valós lineáris térben van definiálva a skaláris szorzat, vagyis (2.2.3) érvényes, akkor **valós euklideszi térről** beszélünk.

A valós euklideszi térre konkrét példa a szokásos értelemben használt geometriai tér, amelyben az euklideszi geometria érvényes (és amelyet ezért szintén euklideszi térnek szoktak nevezni).

Ebben a térben – a skaláris szorzat megszokott geometriai jelentésének megfelelően – két vektor skaláris szorzatát úgy értelmezhetjük, mint hosszuknak és a közbezárt szög koszinuszának a szorzatát (ezzel nyilvánvalóan valós euklideszi teret értelmezünk).

A komplex szám- n -esek lineáris terében a $(\xi_1, \xi_2, \dots, \xi_n)$ és $(\eta_1, \eta_2, \dots, \eta_n)$ vektorok skaláris szorzatát a következőképpen értelmezhetjük:

$$(2.2.4) \quad \bar{\eta}_1 \xi_1 + \bar{\eta}_2 \xi_2 + \dots + \bar{\eta}_n \xi_n,$$

vagy az 1. fejezetben alkalmazott jelöléssel [vö. (1.2.11)-gyel]:

$$(2.2.5) \quad \mathbf{y}^H \mathbf{x},$$

ahol

$$\mathbf{x} = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix}.$$

Az $[a, b]$ zárt intervallumban folytonos függvények terében az f és g függvényeknek mint a lineáris tér „vektorainak” skaláris szorzatát az

$$(2.2.6) \quad (f, g) = \int_a^b f(t) \overline{g(t)} dt$$

szorzatintegrállal értelmezhetjük.

Könnyen igazolható, hogy mindhárom példában az értelmezett skaláris szorzat kielégíti a (2.2.1) követelményeket.

A leggyakrabban használt skaláris szorzat a geometriai tér skaláris szorzatának egyenes általánosítása.

2.2.3 definíció. Legyen $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ az \mathbf{R}_n komplex vagy valós lineáris tér egy bázisa és a tér minden

$$\mathbf{x} = \sum_{i=1}^n x_i \mathbf{u}_i \quad \mathbf{y} = \sum_{i=1}^n y_i \mathbf{u}_i$$

vektorpárjához legyen hozzárendelve az

$$(2.2.7) \quad (\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \overline{y_i} x_i$$

érték. Ekkor a (2.2.7) alatti (\mathbf{x}, \mathbf{y}) az $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ bázisra vonatkozó **standard skaláris szorzat**.

Annak igazolását, hogy egy $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ bázisra vonatkozó standard skaláris szorzat valóban skaláris szorzat, azaz kielégíti a (2.2.1) feltételi egyenleteket, az Olvasóra bízunk. Megemlítjük, hogy példaként a szám- n -esek terében a (2.2.4) alatt definiált skaláris szorzat az

$$(2.2.8) \quad \mathbf{e}_1 = (1, 0, \dots, 0) \quad \mathbf{e}_2 = (0, 1, \dots, 0) \dots \mathbf{e}_n = (0, 0, \dots, 1)$$

bázisra vonatkozó standard skaláris szorzat.

A skaláris szorzat segítségével a következőképpen értelmezzük a vektor hosszát és két pont távolságát.

2.2.4 definíció. Ha $\mathbf{x} \neq \mathbf{0}$ az \mathbf{E} komplex euklideszi tér egy vektora, akkor az

$$(2.2.9) \quad |\mathbf{x}| = \sqrt{(\mathbf{x}, \mathbf{x})}$$

pozitív számot az \mathbf{x} vektor hosszának nevezzük.

A vektor hosszának 2.2.4 definíciója alapján a következőképpen értelmezhető két vektor távolsága.

2.2.5 definíció. Ha \mathbf{x} és \mathbf{y} az \mathbf{E} komplex euklideszi tér vektorai, akkor \mathbf{x} és \mathbf{y} távolságát a

$$(2.2.10) \quad d = |\mathbf{x} - \mathbf{y}|$$

képlettel értelmezzük.

Vagyis \mathbf{x} és \mathbf{y} távolsága az $\mathbf{x} - \mathbf{y}$ különbségvektor hossza.

A geometriában ennek megfelelője két pont távolsága, vagyis az e pontokba mutató helyvektorok különbségének hossza.

A vektorok lineáris függetlenségének eldöntésére szolgál a következő tétel, amelynek a bizonyítására a 2.8.4 pontban visszatérünk.

2.2.1 tétel. Ha $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ az n -dimenziós \mathbf{E}_n euklideszi tér vektorai, akkor a velük képzett

$$(2.2.11) \quad G(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k) = \begin{vmatrix} (\mathbf{x}_1, \mathbf{x}_1) & (\mathbf{x}_1, \mathbf{x}_2) & \dots & (\mathbf{x}_1, \mathbf{x}_k) \\ (\mathbf{x}_2, \mathbf{x}_1) & (\mathbf{x}_2, \mathbf{x}_2) & \dots & (\mathbf{x}_2, \mathbf{x}_k) \\ \dots & \dots & \dots & \dots \\ (\mathbf{x}_k, \mathbf{x}_1) & (\mathbf{x}_k, \mathbf{x}_2) & \dots & (\mathbf{x}_k, \mathbf{x}_k) \end{vmatrix}$$

ún. Gram-féle* determináns nemnegatív, és zérussal akkor és csak akkor egyenlő, ha az $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ vektorok lineárisan összefüggők.

Mielőtt a következő alapfogalmat – a valós euklideszi tér két vektorának hajlásszögét – definiálnánk, bebizonyítunk egy egyenlőtlenséget, amit a matematika különböző területein igen gyakran alkalmaznak. Ez az ún. Cauchy–Bunyakovszkij–Schwarz-egyenlőtlenség, amit a szög fogalmának bevezetéséhez is fel fogunk használni.

2.2.2 tétel (Cauchy–Bunyakovszkij–Schwarz-egyenlőtlenség).** Ha \mathbf{x} és \mathbf{y} a komplex euklideszi tér tetszőleges két vektora, akkor fennáll az alábbi egyenlőtlenség:

$$(2.2.12) \quad |(\mathbf{x}, \mathbf{y})| \leq |\mathbf{x}| \cdot |\mathbf{y}|.$$

Az egyenlőség jele csak akkor érvényes, ha \mathbf{x} és \mathbf{y} lineárisan összefüggők. Bizonyítás. Alkalmazzuk a 2.2.1 tételt $k = 2$ esetén. Legyen $\mathbf{x}_1 = \mathbf{x}$ és $\mathbf{x}_2 = \mathbf{y}$:

$$(2.2.13) \quad \begin{vmatrix} (\mathbf{x}, \mathbf{x}) & (\mathbf{x}, \mathbf{y}) \\ (\mathbf{y}, \mathbf{x}) & (\mathbf{y}, \mathbf{y}) \end{vmatrix} \geq 0,$$

azaz $(\mathbf{x}, \mathbf{x}) \cdot (\mathbf{y}, \mathbf{y}) - (\mathbf{x}, \mathbf{y})(\mathbf{y}, \mathbf{x}) \geq 0$, $|\mathbf{x}|^2 \cdot |\mathbf{y}|^2 - |(\mathbf{x}, \mathbf{y})|^2 \geq 0$, ahonnan a (2.2.12) következik. Mivel (2.2.13)-ban az egyenlőség akkor és csak akkor áll fenn, ha \mathbf{x} és \mathbf{y} lineárisan összefüggők, ezért (2.2.12)-ben az egyenlőségre ugyanez érvényes. ■

A Cauchy–Bunyakovszkij–Schwarz-egyenlőtlenség jelentősége igen nagy, ezért felírjuk speciális alakját a rendezett szám- n -esek terében és az $[a, b]$ zárt intervallumban folytonos függvények terében is. Az előző esetben a (2.2.4) összefüggés alapján

$$(2.2.14) \quad \left| \sum_{i=1}^n \bar{\eta}_i \xi_i \right| \leq \sqrt{\sum_{i=1}^n |\xi_i|^2 \sum_{j=1}^n |\eta_j|^2};$$

az utóbbi esetben – (2.2.6) felhasználásával –

$$(2.2.15) \quad \left| \int_a^b f(t) \overline{g(t)} dt \right| \leq \sqrt{\int_a^b |f(t)|^2 dt \int_a^b |g(t)|^2 dt}.$$

*J. P. Gram (1850–1916) dán matematikus.

**A. L. Cauchy (1789–1857) francia matematikus, V. J. Bunyakovszkij (1804–1889) orosz matematikus, H. A. Schwarz (1843–1921) német matematikus.

Legyen most \mathbf{x} és \mathbf{y} a valós euklideszi tér két vektora. Térjünk vissza a Cauchy–Bunyakovszkij–Schwarz-egyenlőtlenség (2.2.12) alakjára. Végigosztva $|\mathbf{x}|$ és $|\mathbf{y}|$ szorzatával, a következő alakot kapjuk:

$$(2.2.16) \quad -1 \leq \frac{(\mathbf{x}, \mathbf{y})}{|\mathbf{x}| \cdot |\mathbf{y}|} \leq 1.$$

Ebből már következik, hogy az egyenlőtlenségben szereplő törtkifejezés fel fogható pl. egy szög koszinuszaként, mivel az általa felvehető értékek a $[-1, 1]$ intervallumba esnek. Az \mathbf{x} és \mathbf{y} vektor által bezárt szög tehát definiálható úgy, hogy koszinusza a (2.2.16) egyenlőtlenségben fellépő tört legyen.

2.2.6 definíció. Ha \mathbf{x} és \mathbf{y} az \mathbf{E} valós euklideszi tér vektorai, akkor az általuk bezárt szöget α

$$(2.2.17) \quad \cos \alpha = \frac{(\mathbf{x}, \mathbf{y})}{|\mathbf{x}| \cdot |\mathbf{y}|}$$

összefüggéssel értelmezzük.

Ebből a definícióból közvetlenül adódik két vektor ortogonalitásának a fogalma: a geometriai tér két vektoráról akkor mondjuk, hogy egymásra merőlegesek (ortogonálisak), ha az általuk bezárt szög $\frac{\pi}{2}$, azaz, ha a skaláris szorzatuk zérus. Ennek megfelelően kapjuk valós és komplex euklideszi térre egyaránt a következő definíciót.

2.2.7 definíció. Az \mathbf{E} komplex euklideszi tér \mathbf{x} és \mathbf{y} vektorát akkor nevezzük ortogonálisnak, ha

$$(2.2.18) \quad (\mathbf{x}, \mathbf{y}) = 0.$$

Az \mathbf{x} és \mathbf{y} vektor ortogonalítására a geometriából ismert $\mathbf{x} \perp \mathbf{y}$ jelölést alkalmazzuk.

A bevezetett fogalmak és műveletek alkalmazásaként bebizonyítjuk, hogy az euklideszi térben is érvényes a geometriából ismert Pitagorasz-tétel és a háromszög-egyenlőtlenség.

2.2.3 tétel (Pitagorasz-tétel).* Ha \mathbf{x} és \mathbf{y} az \mathbf{E} komplex euklideszi tér egymásra ortogonális két vektora, akkor

$$(2.2.19) \quad |\mathbf{x} + \mathbf{y}|^2 = |\mathbf{x}|^2 + |\mathbf{y}|^2.$$

Bizonyítás. Írjuk fel $|\mathbf{x} + \mathbf{y}|$ négyzetét, a skaláris szorzat segítségével:

$$|\mathbf{x} + \mathbf{y}|^2 = (\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = (\mathbf{x}, \mathbf{x}) + (\mathbf{x}, \mathbf{y}) + (\mathbf{y}, \mathbf{x}) + (\mathbf{y}, \mathbf{y}).$$

*Püthagorasz (kb. i.e. 580–500) görög matematikus és filozófus.

Az \mathbf{x} és \mathbf{y} vektorok ortogonalitása miatt

$$(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x}) = 0, \quad \text{tehát} \quad |\mathbf{x} + \mathbf{y}|^2 = (\mathbf{x}, \mathbf{x}) + (\mathbf{y}, \mathbf{y}) = |\mathbf{x}|^2 + |\mathbf{y}|^2,$$

és ezzel a tételt bebizonyítottuk. ■

A geometriai tér \mathbf{x} és \mathbf{y} vektora az ortogonalitás miatt lineárisan független, tehát egy síkot határoz meg. Az \mathbf{x} és az \mathbf{y} vektor e síkban egy derékszögű háromszög befogói, az átfogó pedig $\mathbf{x} + \mathbf{y}$.

2.2.4 tétel (háromszög-egyenlőtlenség). Az E komplex lineáris tér tetszőleges \mathbf{x} és \mathbf{y} vektorára fennáll az

$$(2.2.20) \quad |\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|$$

egyenlőtlenség.

Bizonyítás. Határozzuk meg az $|\mathbf{x} + \mathbf{y}|^2$ kifejezést:

$$(2.2.21) \quad |\mathbf{x} + \mathbf{y}|^2 = (\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = (\mathbf{x}, \mathbf{x}) + (\mathbf{x}, \mathbf{y}) + (\mathbf{y}, \mathbf{x}) + (\mathbf{y}, \mathbf{y}) = \\ = (\mathbf{x}, \mathbf{x}) + (\mathbf{x}, \mathbf{y}) + \overline{(\mathbf{x}, \mathbf{y})} + (\mathbf{y}, \mathbf{y}).$$

Mivel konjugált komplex számok összege valós részük kétszerese, továbbá egy komplex szám valós része nem lehet nagyobb a szám abszolút értékénél, ezért

$$(\mathbf{x}, \mathbf{y}) + \overline{(\mathbf{x}, \mathbf{y})} = 2 \operatorname{Re}(\mathbf{x}, \mathbf{y}) \leq 2|(\mathbf{x}, \mathbf{y})|.$$

Itt alkalmazhatjuk a Cauchy–Bunyakovszkij–Schwarz-féle egyenlőtlenséget, s így az $(\mathbf{x}, \mathbf{y}) + \overline{(\mathbf{x}, \mathbf{y})}$ mennyiségre az

$$(\mathbf{x}, \mathbf{y}) + \overline{(\mathbf{x}, \mathbf{y})} \leq 2|(\mathbf{x}, \mathbf{y})| \leq 2|\mathbf{x}| \cdot |\mathbf{y}|$$

becslést kapjuk. Ha $(\mathbf{x}, \mathbf{y}) + \overline{(\mathbf{x}, \mathbf{y})}$ helyébe e becslést írjuk, a (2.2.21) egyenlőségből az

$$|\mathbf{x} + \mathbf{y}|^2 \leq |\mathbf{x}|^2 + 2|\mathbf{x}| \cdot |\mathbf{y}| + |\mathbf{y}|^2 = (|\mathbf{x}| + |\mathbf{y}|)^2$$

egyenlőtlenséget nyerjük. Minthogy itt az egyenlőtlenség mindkét oldalán nemnegatív szám négyzete áll, így az egyenlőtlenség a gyökvonás után is érvényes, azaz

$$|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|,$$

amit bizonyítanunk kellett. ■

Ortogonalis vektorok segítségével lehetővé válik az euklideszi térben speciális bázis, ún. ortogonalis bázis bevezetése, ami itt ugyanazt a szerepet tölti be, mint az analitikus geometriában a derékszögű koordináta-rendszer.

2.2.8 definíció. Az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ vektorok az n -dimenziós euklideszi térben **ortogonális bázist** alkotnak, ha páronként ortogonálisak, azaz ha

$$(2.2.22) \quad (\mathbf{e}_i, \mathbf{e}_j) = 0, \quad \text{ha } i \neq j.$$

Ha az ortogonális bázis vektorai egységnyi hosszúságúak, azaz ha

$$(2.2.23) \quad (\mathbf{e}_i, \mathbf{e}_j) = \delta_{ij},$$

akkor az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ vektorok **ortogonális normált bázist**, röviden **ortonormált bázist** alkotnak.

Ez a definíció természetesen csak akkor van összhangban a bázis definíciójával, ha a (2.2.22) tulajdonságú $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ vektorok valóban bázist alkotnak, vagyis lineárisan függetlenek. Az alábbi tételben ezt bizonyítjuk be, majd a következőkben azt, hogy az n -dimenziós euklideszi térben léteznek ortonormált bázisok, és egyúttal megmutatjuk, hogy tetszőleges bázisból hogyan nyerhető ortonormált bázis.

2.2.5 tétel. Ha $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ az n -dimenziós euklideszi térben egymásra páronként ortogonális (zérustól különböző) vektorok, azaz fennáll a (2.2.22) összefüggés, akkor $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ vektorok lineárisan függetlenek.

Bizonyítás. Azt kell belátni, hogy

$$(2.2.24) \quad \lambda_1 \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 + \dots + \lambda_n \mathbf{e}_n = \mathbf{0}$$

csak akkor állhat fenn, ha $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$. Szorozzuk meg ebből a célból (2.2.24) mindkét oldalát (jobbról) skalárisan rendre az \mathbf{e}_i ($i = 1, 2, \dots, n$) vektorokkal:

$$\lambda_1 (\mathbf{e}_1, \mathbf{e}_i) + \lambda_2 (\mathbf{e}_2, \mathbf{e}_i) + \dots + \lambda_i (\mathbf{e}_i, \mathbf{e}_i) + \dots + \lambda_n (\mathbf{e}_n, \mathbf{e}_i) = 0 \quad (i = 1, 2, \dots, n).$$

A (2.2.22) összefüggés szerint $(\mathbf{e}_i, \mathbf{e}_j) = 0$, ha $i \neq j$. Az összeg tehát csak akkor lehet 0, ha a $\lambda_i (\mathbf{e}_i, \mathbf{e}_i)$ tag szintén 0:

$$\lambda_i (\mathbf{e}_i, \mathbf{e}_i) = 0.$$

Mivel $(\mathbf{e}_i, \mathbf{e}_i) \neq 0$, ezért ez csak úgy teljesülhet, ha

$$\lambda_i = 0 \quad (i = 1, 2, \dots, n);$$

ezzel bebizonyítottuk, hogy az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ vektorok valóban lineárisan függetlenek, tehát bázist alkotnak az n -dimenziós euklideszi térben. ■

2.2.6 tétel. Minden n -dimenziós euklideszi térben létezik ortonormált bázis (Schmidt-féle* ortogonalizálási eljárás).

*E. Schmidt (1876–1959) német matematikus.

Bizonyítás. Legyen az n -dimenziós euklideszi tér egy tetszőleges bázisa $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$. Legyen az új bázis első vektora $\mathbf{e}_1 = \mathbf{f}_1$, és keressük a következő vektorát

$$\mathbf{e}_2 = \mathbf{f}_2 + \lambda \mathbf{e}_1$$

alakban. A λ együttható értékét abból a feltételből határozzuk meg, hogy \mathbf{e}_2 ortogonális az \mathbf{e}_1 vektorra, vagyis

$$(\mathbf{e}_2, \mathbf{e}_1) = (\mathbf{f}_2 + \lambda \mathbf{e}_1, \mathbf{e}_1) = (\mathbf{f}_2, \mathbf{e}_1) + \lambda(\mathbf{e}_1, \mathbf{e}_1) = 0.$$

Innen $\lambda = -\frac{(\mathbf{f}_2, \mathbf{e}_1)}{(\mathbf{e}_1, \mathbf{e}_1)}$. Tegyük most fel, hogy már ismerjük az ortogonális bázis $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{k-1}$ vektorait és keressük \mathbf{e}_k vektorát

$$(2.2.25) \quad \mathbf{e}_k = \mathbf{f}_k + \lambda_1 \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 + \dots + \lambda_{k-1} \mathbf{e}_{k-1}$$

alakban. A $\lambda_1, \lambda_2, \dots, \lambda_{k-1}$ együtthatók értékét abból a feltételből határozzuk meg, hogy \mathbf{e}_k ortogonális legyen az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{k-1}$ vektorok mindegyikére, vagyis teljesüljenek az alábbi összefüggések:

$$(\mathbf{f}_k + \lambda_1 \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 + \dots + \lambda_{k-1} \mathbf{e}_{k-1}, \mathbf{e}_1) = 0$$

$$(\mathbf{f}_k + \lambda_1 \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 + \dots + \lambda_{k-1} \mathbf{e}_{k-1}, \mathbf{e}_2) = 0$$

.....

$$(\mathbf{f}_k + \lambda_1 \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 + \dots + \lambda_{k-1} \mathbf{e}_{k-1}, \mathbf{e}_{k-1}) = 0.$$

Kihasználva az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{k-1}$ vektorok ortogonalitását, ebből következik, hogy

$$(\mathbf{f}_k, \mathbf{e}_1) + \lambda_1(\mathbf{e}_1, \mathbf{e}_1) = 0$$

$$(\mathbf{f}_k, \mathbf{e}_2) + \lambda_2(\mathbf{e}_2, \mathbf{e}_2) = 0$$

.....

$$(\mathbf{f}_k, \mathbf{e}_{k-1}) + \lambda_{k-1}(\mathbf{e}_{k-1}, \mathbf{e}_{k-1}) = 0,$$

$$\text{ahonnan } \lambda_1 = -\frac{(\mathbf{f}_k, \mathbf{e}_1)}{(\mathbf{e}_1, \mathbf{e}_1)}; \quad \lambda_2 = -\frac{(\mathbf{f}_k, \mathbf{e}_2)}{(\mathbf{e}_2, \mathbf{e}_2)}; \quad \dots; \quad \lambda_{k-1} = -\frac{(\mathbf{f}_k, \mathbf{e}_{k-1})}{(\mathbf{e}_{k-1}, \mathbf{e}_{k-1})}.$$

Be kell még látnunk, hogy \mathbf{e}_k nem lehet a zérusvektor. Vegyük figyelembe, hogy a fenti konstrukciónál az \mathbf{e}_i vektor előállításában csak az $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_i$ vektorok lépnek fel ($i = 1, 2, \dots, k-1$). Így az \mathbf{e}_k vektor előállítása (2.2.25) alapján

$$(2.2.26) \quad \mathbf{e}_k = \mathbf{f}_k + a_1 \mathbf{f}_1 + a_2 \mathbf{f}_2 + \dots + a_{k-1} \mathbf{f}_{k-1}$$

alakban írható, ahol a_1, a_2, \dots, a_{k-1} a (2.2.25) alatti lineáris kombináció λ_i együtthatóiból számítható. Ebben az előállításban azonban \mathbf{f}_k együtthatója nem zérus, így az $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_k$ vektorok lineáris függetlensége miatt a (2.2.26) alatti lineáris kombináció nem állíthatja elő a zérusvektort, tehát valóban $\mathbf{e}_k \neq 0$.

Az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ ortogonális bázisból az

$$\mathbf{e}'_i = \frac{\mathbf{e}_i}{|\mathbf{e}_i|} = \frac{\mathbf{e}_i}{\sqrt{(\mathbf{e}_i, \mathbf{e}_i)}} \quad (i = 1, 2, \dots, n)$$

normálással kapott $\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n$ bázis ortonormált. ■

1. Példa. A számhármassok euklideszi terének legyen egy bázisa $\mathbf{f}_1 = (1, 1, 1)$, $\mathbf{f}_2 = (1, 0, 1)$, $\mathbf{f}_3 = (0, 1, 1)$. A Schmidt-féle ortogonalizálási eljárással képezzünk ezekből egy ortonormált bázist.

Megoldás. Legyen az új bázis első vektora $\mathbf{e}_1 = \mathbf{f}_1$ és keressük a következő vektorát $\mathbf{e}_2 = \mathbf{f}_2 + \lambda_{21}\mathbf{e}_1$ alakban, ahol $\lambda_{21} = -\frac{(\mathbf{f}_2, \mathbf{e}_1)}{(\mathbf{e}_1, \mathbf{e}_1)}$. Mivel $(\mathbf{e}_1, \mathbf{e}_1) = 3$, $(\mathbf{f}_2, \mathbf{e}_1) = 2$, így $\lambda_{21} = -\frac{2}{3}$, és

$$\mathbf{e}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{2}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{bmatrix}; \quad (\mathbf{e}_2, \mathbf{e}_2) = \frac{2}{3}.$$

Hasonlóképpen $\mathbf{e}_3 = \mathbf{f}_3 + \lambda_{31}\mathbf{e}_1 + \lambda_{32}\mathbf{e}_2$, ahol

$$\lambda_{31} = -\frac{(\mathbf{f}_3, \mathbf{e}_1)}{(\mathbf{e}_1, \mathbf{e}_1)} \quad \text{és} \quad \lambda_{32} = -\frac{(\mathbf{f}_3, \mathbf{e}_2)}{(\mathbf{e}_2, \mathbf{e}_2)}.$$

Mivel $(\mathbf{f}_3, \mathbf{e}_1) = 2$, $(\mathbf{f}_3, \mathbf{e}_2) = -\frac{1}{3}$, így $\lambda_{31} = -\frac{2}{3}$, $\lambda_{32} = \frac{1}{2}$, tehát

$$\mathbf{e}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \end{bmatrix}; \quad (\mathbf{e}_3, \mathbf{e}_3) = \frac{1}{2}.$$

A keresett ortonormált bázis:

$$\left[\frac{\mathbf{e}_1}{\sqrt{(\mathbf{e}_1, \mathbf{e}_1)}} \quad \frac{\mathbf{e}_2}{\sqrt{(\mathbf{e}_2, \mathbf{e}_2)}} \quad \frac{\mathbf{e}_3}{\sqrt{(\mathbf{e}_3, \mathbf{e}_3)}} \right] = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

* * *

2. Példa. Határozzuk meg a $[-1, +1]$ intervallumon az ortogonális és normált k -adfokú polinomokat (az ún. *Legendre-féle polinomokat**) $k=0, 1, 2, \dots, 5$ esetén!

Megoldás. Mivel egy páros és egy páratlan függvény szorzata páratlan függvény, amelynek a 0-ra szimmetrikusan elhelyezkedő $[-1, +1]$ intervallumra vett integrálja 0, ezért elegendő, ha külön megkeressük a páratlan, ill. a páros Legendre-polinomokat a Schmidt-féle ortogonalizálási eljárással.

a) Legyen a páros polinomok bázisa

$$f_1 = 1, \quad f_2 = x^2, \quad f_3 = x^4.$$

Az új bázis első vektora legyen $e_1 = f_1 = 1$ és keressük a következő vektorát $e_2 = f_2 + \lambda_{21}e_1$ alakban, ahol $\lambda_{21} = -\frac{(f_2e_1)}{(e_1e_1)}$. Mivel

$$(e_1e_1) = \int_{-1}^{+1} dx = 2, \quad (f_2e_1) = \int_{-1}^{+1} x^2 dx = \frac{2}{3}, \quad \lambda_{21} = -\frac{1}{3}, \text{ így}$$

$$e_2 = x^2 - \frac{1}{3}; \quad (e_2e_2) = \int_{-1}^{+1} \left(x^2 - \frac{1}{3}\right)^2 dx = \frac{8}{45} = \left(\frac{2}{3}\sqrt{\frac{2}{5}}\right)^2.$$

Hasonlóképpen

$$e_3 = f_3 + \lambda_{31}e_1 + \lambda_{32}e_2, \text{ ahol } \lambda_{31} = -\frac{(f_3e_1)}{(e_1e_1)} \text{ és } \lambda_{32} = -\frac{(f_3e_2)}{(e_2e_2)}.$$

Mivel

$$(f_3e_1) = \int_{-1}^{+1} x^4 dx = \frac{2}{5}; \quad (f_3e_2) = \int_{-1}^{+1} x^4 \left(x^2 - \frac{1}{3}\right) dx = \frac{16}{105}, \text{ így}$$

$$\lambda_{31} = -\frac{1}{5}, \quad \lambda_{32} = -\frac{6}{7}, \text{ tehát } e_3 = x^4 - \frac{1}{5} - \frac{6}{7}\left(x^2 - \frac{1}{3}\right) = x^4 - \frac{6}{7}x^2 + \frac{3}{35};$$

$$(e_3e_3) = \int_{-1}^{+1} \left(x^4 - \frac{6}{7}x^2 + \frac{3}{35}\right)^2 dx = \left(\frac{8}{105}\sqrt{2}\right)^2.$$

b) Legyen a páratlan polinomok bázisa

$$f_1 = x, \quad f_2 = x^3, \quad f_3 = x^5.$$

* A-M. Legendre (1752–1833) francia matematikus.

Az új bázis első vektora legyen $\mathbf{e}_1 = \mathbf{f}_1 = x$ és keressük a következő vektorát $\mathbf{e}_2 = \mathbf{f}_2 + \lambda_{21}\mathbf{e}_1$ alakban, ahol $\lambda_{21} = -\frac{(\mathbf{f}_2\mathbf{e}_1)}{(\mathbf{e}_1\mathbf{e}_1)}$. Mivel $(\mathbf{e}_1\mathbf{e}_1) = \int_{-1}^{+1} x^2 dx = \frac{2}{3}$;

$$(\mathbf{f}_2\mathbf{e}_1) = \int_{-1}^{+1} x^4 dx = \frac{2}{5}; \quad \lambda_{21} = -\frac{3}{5}, \quad \text{így}$$

$$\mathbf{e}_2 = x^3 - \frac{3}{5}x; \quad (\mathbf{e}_2\mathbf{e}_2) = \int_{-1}^{+1} \left(x^3 - \frac{3}{5}x\right)^2 dx = \frac{8}{175} = \left(\frac{2}{5}\sqrt{\frac{2}{7}}\right)^2.$$

Hasonlóképpen $\mathbf{e}_3 = \mathbf{f}_3 + \lambda_{31}\mathbf{e}_1 + \lambda_{32}\mathbf{e}_2$, ahol $\lambda_{31} = -\frac{(\mathbf{f}_3\mathbf{e}_1)}{(\mathbf{e}_1\mathbf{e}_1)}$ és $\lambda_{32} = -\frac{(\mathbf{f}_3\mathbf{e}_2)}{(\mathbf{e}_2\mathbf{e}_2)}$. Mivel $(\mathbf{f}_3\mathbf{e}_1) = \int_{-1}^{+1} x^6 dx = \frac{2}{7}$; $(\mathbf{f}_3\mathbf{e}_2) = \int_{-1}^{+1} x^5 \left(x^3 - \frac{3}{5}x\right) dx = \frac{16}{315}$,
így

$$\mathbf{e}_3 = x^5 - \frac{3}{7}x - \frac{10}{9} \left(x^3 - \frac{3}{5}x\right) = x^5 - \frac{10}{9}x^3 + \frac{5}{21}x.$$

$$(\mathbf{e}_3\mathbf{e}_3) = \int_{-1}^{+1} \left(x^5 - \frac{10}{9}x^3 + \frac{5}{21}x\right)^2 dx = \left(\frac{8}{63}\sqrt{\frac{2}{11}}\right)^2.$$

A normált ortogonális polinomokat megkapjuk, ha az ortogonális bázis vektorait osztjuk az abszolút értékükkel (azaz négyzetintegráljuk négyzetgyökével). Így a normált Legendre-polinomokra az alábbi kifejezéseket kapjuk:

$$P_0(x) = \frac{1}{\sqrt{2}}; \quad P_1(x) = \sqrt{\frac{3}{2}}x; \quad P_2(x) = \frac{3}{2}\sqrt{\frac{5}{2}}\left(x^2 - \frac{1}{3}\right);$$

$$P_3(x) = \frac{5}{2}\sqrt{\frac{7}{2}}\left(x^3 - \frac{3}{5}x\right); \quad P_4(x) = \frac{105}{8\sqrt{2}}\left(x^4 - \frac{6}{7}x^2 + \frac{3}{35}\right);$$

$$P_5(x) = \frac{63}{8}\sqrt{\frac{11}{2}}\left(x^5 - \frac{10}{9}x^3 + \frac{5}{21}x\right).$$

* * *

Vizsgáljuk most meg, hogy ortonormált bázis választása mellett az n -dimenziós \mathbf{E} (komplex) euklideszi térben hogyan fejezhető ki a tér két vektorának skaláris szorzata a vektorok koordinátáinak segítségével. Legyen

$\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ egy adott bázis, és legyen \mathbf{x} és \mathbf{y} a két vektor, amelyek az adott bázisra vonatkozó koordinátáikkal felírva: $\mathbf{x} = \xi_1 \mathbf{e}_1 + \xi_2 \mathbf{e}_2 + \dots + \xi_n \mathbf{e}_n$, és $\mathbf{y} = \eta_1 \mathbf{e}_1 + \eta_2 \mathbf{e}_2 + \dots + \eta_n \mathbf{e}_n$. Ha most képezzük az (\mathbf{x}, \mathbf{y}) skaláris szorzatot, akkor

$$(2.2.27) \quad (\mathbf{x}, \mathbf{y}) = \left(\sum_{i=1}^n \xi_i \mathbf{e}_i, \sum_{j=1}^n \eta_j \mathbf{e}_j \right) = \sum_{i=1}^n \sum_{j=1}^n \overline{\eta_j} \xi_i (\mathbf{e}_i, \mathbf{e}_j)$$

adódik. Ez a kifejezés, mint látható, általában tartalmazza a bázisvektorok páronként vett $(\mathbf{e}_i, \mathbf{e}_j)$ skaláris szorzatát, valamint az \mathbf{x} és \mathbf{y} vektorok koordinátáiból képezett $\overline{\eta_j} \xi_i$ alakú szorzatokat. Ortonormált bázis választása esetén azonban a (2.2.27) kifejezés lényegesen egyszerűbb alakot ölt. Ebben az esetben ugyanis a bázis vektoraira fennáll a (2.2.23) összefüggés, és ennek behelyettesítésével a (2.2.27) skaláris szorzat az alábbi egyszerű alakot veszi fel:

$$(2.2.28) \quad (\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \overline{\eta_i} \xi_i.$$

Kimondható tehát a következő tétel.

2.2.7 tétel. *Ortonormált bázis választása esetén az n -dimenziós komplex euklideszi tér két vektorának skaláris szorzata egyenlő az első vektor koordinátáinak és a második vektor megfelelő koordinátái konjugáltjának szorzatösszegével.*

Megjegyzés. Figyelembe véve egy adott bázisra vonatkozó standard skaláris szorzat 2.2.3 definícióját, a tétel a következőképpen fogalmazható: *Az euklideszi tér skaláris szorzata bármely ortonormált bázisra vonatkozóan standard.*

Az ortonormált bázisokat tehát az tünteti ki, hogy rájuk vonatkozóan a skaláris szorzat standard. A 2.2.6 tétel alapján pedig mindig van olyan bázis, amelyre vonatkozóan a skaláris szorzat standard.

Vezessük be az \mathbf{x} és \mathbf{y} vektornak a tér egy adott $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisára vonatkozó $\xi_1, \xi_2, \dots, \xi_n$ ill. $\eta_1, \eta_2, \dots, \eta_n$ koordinátáiból alkotott

$$\mathbf{x} = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \quad \text{és} \quad \mathbf{y} = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix}$$

koordinátavektorokat. A 2.2.7 tétel szerint tehát, ha $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ ortonormált bázis, akkor az (\mathbf{x}, \mathbf{y}) skaláris szorzat az 1. fejezet jelöléseivel, a koordinátavektorok segítségével a következő alakban írható:

$$(2.2.29) \quad (\mathbf{x}, \mathbf{y}) = \mathbf{y}^H \mathbf{x}.$$

A valós euklideszi térben a konjugálás természetsszerűen feleslegessé válik, ekkor tehát a két vektor skaláris szorzata az

$$(2.2.30) \quad (\mathbf{x}, \mathbf{y}) = \mathbf{y}^T \mathbf{x}$$

képlettel számítható.

Most azt mutatjuk meg, hogy ortonormált bázis választása esetén, az euklideszi tér tetszőleges \mathbf{x} vektorának a koordinátáit is egyszerűen határozhatjuk meg. Legyen ugyanis $\mathbf{x} = \xi_1 \mathbf{e}_1 + \xi_2 \mathbf{e}_2 + \cdots + \xi_n \mathbf{e}_n$, és szorozzuk meg mindkét oldalt jobbról skalárisan az \mathbf{e}_1 vektorral. Azt kapjuk, hogy

$$(2.2.31) \quad \begin{aligned} (\mathbf{x}, \mathbf{e}_1) &= \xi_1 (\mathbf{e}_1, \mathbf{e}_1) + \xi_2 (\mathbf{e}_2, \mathbf{e}_1) + \cdots + \xi_n (\mathbf{e}_n, \mathbf{e}_1) = \xi_1, \\ \text{tehát} \quad \xi_1 &= (\mathbf{x}, \mathbf{e}_1); \quad \text{és ugyanígy} \quad \xi_2 = (\mathbf{x}, \mathbf{e}_2); \dots; \xi_n = (\mathbf{x}, \mathbf{e}_n). \end{aligned}$$

Ezek szerint egy vektornak egy ortonormált bázisra vonatkozó koordinátái a vektornak a megfelelő bázisvektorokkal képezett skaláris szorzatai. Az \mathbf{x} vektornak az \mathbf{e}_i egységvektorral képezett skaláris szorzatát – a geometriában szokásos elnevezést átvéve – az \mathbf{x} vektornak az \mathbf{e}_i vektorra való merőleges vetületének nevezzük. Azt kaptuk tehát, hogy egy vektornak egy ortonormált bázisra vonatkozó koordinátái a vektornak a bázisvektorokra (a koordináta-tengelyekre) vett vetületei.

3. Példa. Tekintsük a $[0, 2\pi]$ intervallumon az

$$(2.2.32) \quad 1, \cos t, \sin t, \cos 2t, \sin 2t, \dots, \cos nt, \sin nt$$

függvényrendszert. E függvények $P(t) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos kt + b_k \sin kt)$

alakú lineáris kombinációját *n-edrendű trigonometrikus polinomnak* nevezzük. A legfeljebb *n*-edrendű trigonometrikus polinomok halmaza egy $(2n+1)$ -dimenziós lineáris teret alkot. A skaláris szorzatot e lineáris térben, ha

$\mathbf{x} = P_1(t)$ és $\mathbf{y} = P_2(t)$, az $(\mathbf{x}, \mathbf{y}) = \int_0^{2\pi} P_1(t)P_2(t) dt$ integrállal definiáljuk.

A (2.2.32) függvényrendszer ortogonális bázist alkot, ugyanis

$$\int_0^{2\pi} \cos kt \cos lt dt = \delta_{kl} \pi, \quad \int_0^{2\pi} \sin kt \sin lt dt = \delta_{kl} \pi,$$

$$\int_0^{2\pi} \sin kt \cos lt dt = 0, \quad \int_0^{2\pi} 1 dt = 2\pi.$$

Innen az is következik, hogy az

$$\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos t, \frac{1}{\sqrt{\pi}} \sin t, \dots, \frac{1}{\sqrt{\pi}} \cos nt, \frac{1}{\sqrt{\pi}} \sin nt$$

függvények az n -edrendű trigonometrikus polinomok $(2n+1)$ -dimenziós térben ortonormált bázist alkotnak.

* * *

Ugyanúgy, mint a lineáris terek vizsgálatánál, itt is felmerül az a kérdés, hogy az euklideszi terek közül melyek a valóban különbözőek, és melyek tekinthetők „egyenrangúaknak”. Ehhez meg kell adnunk az euklideszi terek izomorfájának a definícióját.

2.2.9 definíció. Az \mathbf{E} és \mathbf{E}' euklideszi teret **izomorf**nak nevezzük, ha az $\mathbf{x} \in \mathbf{E}$ és $\mathbf{x}' \in \mathbf{E}'$ vektorok között olyan kölcsönösen egyértelmű megfeleltetést lehet létrehozni, amely kielégíti a (2.1.5), (2.1.6) összefüggéseket, és a megfelelő vektorpárok skaláris szorzatai egyenlők, azaz

$$(2.2.33) \quad (\mathbf{x}, \mathbf{y}) = (\mathbf{x}', \mathbf{y}').$$

Ha az \mathbf{E} és \mathbf{E}' euklideszi terek izomorfak, akkor az \mathbf{E} euklideszi térben bebizonyított tételnek megfelelő tétel igaz \mathbf{E}' -ben is.

2.2.8 tétel. Az összes n -dimenziós euklideszi terek izomorfak.

Bizonyítás. Bebizonyítjuk, hogy az összes n -dimenziós euklideszi tér izomorf egy speciálisan választott „standard” n -dimenziós térrel.

Ilyen standard n -dimenziós \mathbf{E}' euklideszi tér a szám- n -esek tere. Ebben a térben az $\mathbf{x}' = (\xi_1, \dots, \xi_n)$ és az $\mathbf{y}' = (\eta_1, \dots, \eta_n)$ vektorok skaláris szorzata az

$$(2.2.34) \quad (\mathbf{x}', \mathbf{y}') = \xi_1 \overline{\eta_1} + \dots + \xi_n \overline{\eta_n}$$

összefüggéssel van megadva.

Legyen adott egy n -dimenziós \mathbf{E} euklideszi tér. Válasszunk ebben a térben egy $\mathbf{e}_1, \dots, \mathbf{e}_n$ ortonormált bázist (minden euklideszi térben van ilyen bázis).

Feleltessük meg az $\mathbf{x} = \xi_1 \mathbf{e}_1 + \dots + \xi_n \mathbf{e}_n$ vektornak az \mathbf{E}' tér

$$\mathbf{x}' = (\xi_1, \dots, \xi_n)$$

vektorát!

Most megmutatjuk, hogy ez a megfeleltetés izomorf. A művelettartás ((2.1.5) és (2.1.6) összefüggés) nyilvánvaló, csak a skaláris szorzatok egyenlőségét kell igazolni. Tudjuk, hogy ortonormált bázis esetén a skaláris szorzat

$$(\mathbf{x}, \mathbf{y}) = \xi_1 \overline{\eta_1} + \dots + \xi_n \overline{\eta_n}.$$

Ez megegyezik az $(\mathbf{x}', \mathbf{y}') = \xi_1 \overline{\eta_1} + \cdots + \xi_n \overline{\eta_n}$ skaláris szorzattal:

$$(\mathbf{x}, \mathbf{y}) = (\mathbf{x}', \mathbf{y}'),$$

vagyis bebizonyítottuk a skaláris szorzatok egyenlőségét. ■

2.3 LINEÁRIS FÜGGVÉNYEK, BILINEÁRIS ÉS KVADRATIKUS ALAKOK

Ebben a szakaszban a komplex lineáris tér vektorainak egyszerű számértékű függvényeit vizsgáljuk, azaz olyan függvényeket, amelyek a lineáris térben minden vektorhoz, ill. vektorpárhoz egy komplex számot rendelnek hozzá. Először definiáljuk a komplex lineáris térben az első- és másodfajú lineáris függvényeket, majd bevezetjük a bilineáris és a kvadratikus alak fogalmát. Megmutatjuk, hogy adott bázis esetén a bilineáris alakhoz hozzárendelhető egy mátrix. Ezután azt mutatjuk meg, hogy a komplex lineáris térben minden bilineáris alakot egyértelműen meghatároz a hozzá tartozó kvadratikus alak. Bevezetjük az Hermite-féle bilineáris alak fogalmát, majd szükséges és elégséges feltételeket adunk arra, hogy egy bilineáris alak Hermite-féle legyen. Megismerkedünk az Hermite-féle kvadratikus alak, másképpen az hermitikus alak fogalmával és megmutatjuk, hogy az unitér tér meghatározható mint olyan komplex lineáris tér, amelyben értelmezve van egy pozitív definit hermitikus alak. Végül, speciális esetként, megkapjuk a valós lineáris térre vonatkozó eredményeket.

2.3.1 definíció. Legyen $f(\mathbf{x})$ az \mathbf{R} komplex lineáris tér \mathbf{x} vektorain értelmezett (komplex) számértékű függvény. Az $f(\mathbf{x})$ függvényt **elsőfajú lineáris függvénynek** nevezzük, ha eleget tesz az alábbi feltételeknek:

$$(2.3.1) \quad \begin{aligned} (a) \quad & f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y}) & \mathbf{x}, \mathbf{y} \in \mathbf{R}; \\ (b) \quad & f(\lambda \mathbf{x}) = \lambda f(\mathbf{x}) & \lambda \in K, \end{aligned}$$

ahol K jelöli a komplex számok testét.

Az $f(\mathbf{x})$ függvényt **másodfajú lineáris függvénynek** nevezzük, ha az alábbi feltételeknek tesz eleget:

$$(2.3.2) \quad \begin{aligned} (a) \quad & f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y}) & \mathbf{x}, \mathbf{y} \in \mathbf{R}; \\ (b) \quad & f(\lambda \mathbf{x}) = \overline{\lambda} f(\mathbf{x}) & \lambda \in K. \end{aligned}$$

Valós lineáris tér esetében az első- és másodfajú lineáris függvény fogalma egybeesik, ezért **lineáris függvényről** beszélünk.

Legyen $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ az n -dimenziós \mathbf{R}_n komplex lineáris tér egy bázisa. Az $\mathbf{x} \in \mathbf{R}_n$ vektor ekkor felírható a következő alakban:

$$(2.3.3) \quad \mathbf{x} = \xi_1 \mathbf{e}_1 + \xi_2 \mathbf{e}_2 + \cdots + \xi_n \mathbf{e}_n,$$

ahol $\xi_1, \xi_2, \dots, \xi_n$ az \mathbf{x} vektornak az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisra vonatkozó koordinátái.

A 2.3.1 definíció alapján az $f(\mathbf{x})$ elsőfajú lineáris függvény felírható mint

$$f(\mathbf{x}) = f\left(\sum_{i=1}^n \xi_i \mathbf{e}_i\right) = \sum_{i=1}^n \xi_i f(\mathbf{e}_i).$$

Bevezetve az

$$(2.3.4) \quad a_i = f(\mathbf{e}_i) \quad (i = 1, 2, \dots, n)$$

összefüggéssel definiált a_i számokat, az elsőfajú lineáris függvényre azt kapjuk, hogy

$$(2.3.5) \quad f(\mathbf{x}) = \xi_1 a_1 + \xi_2 a_2 + \dots + \xi_n a_n,$$

ahol az a_i számok csak a bázis megválasztásától függenek, a ξ_i számok pedig az \mathbf{x} vektornak ugyancsak a bázistól függő koordinátái. A másodfajú lineáris függvény

$$f(\mathbf{x}) = f\left(\sum_{i=1}^n \xi_i \mathbf{e}_i\right) = \sum_{i=1}^n \bar{\xi}_i f(\mathbf{e}_i)$$

alakban írható; a (2.3.4) összefüggések alapján ez esetben

$$(2.3.6) \quad f(\mathbf{x}) = \bar{\xi}_1 a_1 + \bar{\xi}_2 a_2 + \dots + \bar{\xi}_n a_n.$$

2.3.2 definíció. Legyen \mathbf{x} és \mathbf{y} az n -dimenziós \mathbf{R}_n komplex lineáris tér tetszőleges két vektora. Az $A(\mathbf{x}; \mathbf{y})$ függvény az \mathbf{x} és \mathbf{y} vektor **bilineáris alakja**, ha

- (a) rögzített \mathbf{y} esetén $A(\mathbf{x}; \mathbf{y})$ elsőfajú lineáris függvénye \mathbf{x} -nek;
- (b) rögzített \mathbf{x} esetén $A(\mathbf{x}; \mathbf{y})$ másodfajú lineáris függvénye \mathbf{y} -nak.

Részletesen felírva ez a következőt jelenti:

$$(2.3.7) \quad \begin{aligned} A(\mathbf{x}_1 + \mathbf{x}_2; \mathbf{y}) &= A(\mathbf{x}_1; \mathbf{y}) + A(\mathbf{x}_2; \mathbf{y}), \\ A(\lambda \mathbf{x}; \mathbf{y}) &= \lambda A(\mathbf{x}; \mathbf{y}), \\ A(\mathbf{x}; \mathbf{y}_1 + \mathbf{y}_2) &= A(\mathbf{x}; \mathbf{y}_1) + A(\mathbf{x}; \mathbf{y}_2), \\ A(\mathbf{x}; \lambda \mathbf{y}) &= \bar{\lambda} A(\mathbf{x}; \mathbf{y}). \end{aligned}$$

4. Példa. A komplex szám- n -esek n -dimenziós terében bilineáris alak az a függvény, amely a $(\xi_1, \xi_2, \dots, \xi_n)$ és $(\eta_1, \eta_2, \dots, \eta_n)$ szám- n -es párhoz az

$$(2.3.8) \quad A(\mathbf{x}; \mathbf{y}) = \sum_{j=1}^n \sum_{i=1}^n \bar{\eta}_i \xi_j a_{ij}$$

kifejezést rendeli, ahol a_{ij} ($i = 1, 2, \dots, n$) adott komplex számok. Könnyen belátható, hogy a (2.3.8) alatti kifejezésre a (2.3.7) alatt megadott feltételek teljesülnek. Ha a szám- n -eseket egy-egy n -elemű oszlopvektorba foglaljuk össze, akkor az 1. fejezet jelöléseivel [vö.: (1.2.21)-gyel] a (2.3.8) bilineáris alak a következő alakban írható:

$$(2.3.9) \quad A(\mathbf{x}; \mathbf{y}) = \mathbf{y}^H \mathbf{A} \mathbf{x},$$

ahol $\mathbf{A} = [a_{ij}]$ az adott a_{ij} számokból álló mátrix.

Legyen például az $A(\mathbf{x}; \mathbf{y})$ bilineáris alak megadva az

$$A(\mathbf{x}; \mathbf{y}) = 2\xi_1\bar{\eta}_1 + \xi_1\bar{\eta}_2 + \xi_2\bar{\eta}_1 + 3\xi_2\bar{\eta}_2 + \xi_3\bar{\eta}_3$$

képlettel. A (2.3.9) összefüggés szerint ez a következő alakban írható:

$$A(\mathbf{x}; \mathbf{y}) = \begin{bmatrix} \bar{\eta}_1 & \bar{\eta}_2 & \bar{\eta}_3 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix}.$$

* * *

5. Példa. Az $[a, b]$ zárt intervallumban folytonos függvények lineáris terében az f és g vektorok bilineáris alakja az alábbi integrál:

$$A(f; g) = \int_a^b \int_a^b K(s; t) f(s) \overline{g(t)} ds dt,$$

ahol $K(s; t)$ az s és t változó valamilyen folytonos függvénye.

* * *

A következőkben megmutatjuk, hogy az n -dimenziós komplex térben minden bilineáris alak egy mátrix segítségével írható fel, ha meg van adva a tér egy bázisa. Legyen $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ az n -dimenziós \mathbf{R}_n komplex lineáris tér egy bázisa, továbbá legyenek $(\xi_1, \xi_2, \dots, \xi_n)$, ill. $(\eta_1, \eta_2, \dots, \eta_n)$ az $\mathbf{x} \in \mathbf{R}_n$, ill. $\mathbf{y} \in \mathbf{R}_n$ vektor adott bázisra vonatkozó koordinátái. Az \mathbf{R}_n térben értelmezett $A(\mathbf{x}; \mathbf{y})$ bilineáris alak a (2.3.7) tulajdonságok felhasználásával a következőképpen írható:

$$A(\mathbf{x}; \mathbf{y}) = A \left(\sum_{j=1}^n \xi_j \mathbf{e}_j; \sum_{i=1}^n \eta_i \mathbf{e}_i \right) = \sum_{i=1}^n \sum_{j=1}^n \bar{\eta}_i \xi_j A(\mathbf{e}_j; \mathbf{e}_i).$$

Ha a bilineáris alaknak a bázisvektorokon felvett értékére bevezetjük az alábbi jelölést:

$$(2.3.10) \quad a_{ij} = A(\mathbf{e}_j; \mathbf{e}_i),$$

akkor a térben értelmezett bármely bilineáris alak felírható a következőképpen:

$$(2.3.11) \quad A(\mathbf{x}; \mathbf{y}) = \sum_{i=1}^n \sum_{j=1}^n \bar{\eta}_i a_{ij} \xi_j,$$

ahol az a_{ij} számokból alkotott $\mathbf{A} = [a_{ij}]$ mátrixot az $A(\mathbf{x}, \mathbf{y})$ bilineáris alak adott $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisra vonatkozó mátrixának nevezzük.

A (2.3.11) bilineáris alak tehát az

$$\mathbf{x} = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \quad \text{és} \quad \mathbf{y} = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix}$$

koordinátavektorokkal – az 1. fejezet jelöléseit használva – a következő alakban írható:

$$(2.3.13) \quad A(\mathbf{x}; \mathbf{y}) = \mathbf{y}^H \mathbf{A} \mathbf{x}.$$

6. Példa. Határozzuk meg az 1. példában megadott bilineáris alak mátrixát, ha a számhármások terében bázisnak választjuk az

$$\mathbf{e}_1 = (1, 1, 1) \quad \mathbf{e}_2 = (0, 1, 1) \quad \mathbf{e}_3 = (0, 0, 1)$$

vektorokat!

Megoldás. Az

$$A(\mathbf{x}; \mathbf{y}) = 2\xi_1\bar{\eta}_1 + \xi_1\bar{\eta}_2 + \xi_2\bar{\eta}_1 + 3\xi_2\bar{\eta}_2 + \xi_3\bar{\eta}_3$$

bilineáris alak adott bázishoz tartozó mátrixának elemei a (2.3.10) képlet szerint számíthatók ki:

$$\begin{aligned} a_{11} &= A(\mathbf{e}_1; \mathbf{e}_1) = 8, \\ a_{12} &= a_{21} = A(\mathbf{e}_2; \mathbf{e}_1) = 5, \\ a_{22} &= A(\mathbf{e}_2; \mathbf{e}_2) = 4, \\ a_{13} &= a_{31} = A(\mathbf{e}_3; \mathbf{e}_1) = 1, \\ a_{23} &= a_{32} = A(\mathbf{e}_3; \mathbf{e}_2) = 1, \\ a_{33} &= A(\mathbf{e}_3; \mathbf{e}_3) = 1. \end{aligned}$$

A bilineáris alak mátrixa tehát a következő:

$$\mathbf{A} = \begin{bmatrix} 8 & 5 & 1 \\ 5 & 4 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

* * *

2.3.3 definíció. Ha a komplex lineáris térben értelmezett $A(\mathbf{x}; \mathbf{y})$ bilineáris alakban \mathbf{y} helyére az \mathbf{x} vektort írjuk, akkor az így kapott $A(\mathbf{x}; \mathbf{x})$ kifejezést (a komplex lineáris térben) **kvadratikus alaknak** nevezzük.

2.3.1 tétel. A komplex lineáris térben minden bilineáris alakot egyértelműen meghatároz a hozzá tartozó kvadratikus alak.

Bizonyítás. Legyen $A(\mathbf{x}; \mathbf{x})$ egy kvadratikus alak a komplex lineáris térben és legyen \mathbf{x} és \mathbf{y} a tér két tetszőleges vektora. A bilineáris alakok (2.3.7) tulajdonságai alapján felírhatók az alábbi azonosságok (itt i a képzetes egységet jelöli):

$$(2.3.13) \quad A(\mathbf{x} + \mathbf{y}; \mathbf{x} + \mathbf{y}) = A(\mathbf{x}; \mathbf{x}) + A(\mathbf{y}; \mathbf{x}) + A(\mathbf{x}; \mathbf{y}) + A(\mathbf{y}; \mathbf{y}),$$

$$(2.3.14) \quad A(\mathbf{x} + i\mathbf{y}; \mathbf{x} + i\mathbf{y}) = A(\mathbf{x}; \mathbf{x}) + iA(\mathbf{y}; \mathbf{x}) - iA(\mathbf{x}; \mathbf{y}) + A(\mathbf{y}; \mathbf{y}),$$

$$(2.3.15) \quad A(\mathbf{x} - \mathbf{y}; \mathbf{x} - \mathbf{y}) = A(\mathbf{x}; \mathbf{x}) - A(\mathbf{y}; \mathbf{x}) - A(\mathbf{x}; \mathbf{y}) + A(\mathbf{y}; \mathbf{y}),$$

$$(2.3.16) \quad A(\mathbf{x} - i\mathbf{y}; \mathbf{x} - i\mathbf{y}) = A(\mathbf{x}; \mathbf{x}) - iA(\mathbf{y}; \mathbf{x}) + iA(\mathbf{x}; \mathbf{y}) + A(\mathbf{y}; \mathbf{y}).$$

Ha a (2.3.13) azonosságot 1-gyel, (2.3.14)-et i -vel (2.3.15)-öt -1 -gyel, (2.3.16)-ot $-i$ -vel megszorozzuk és az így kapott azonosságokat összeadjuk, akkor a jobb oldalon $A(\mathbf{x}; \mathbf{y})$ kivételével valamennyi tag kiesik és így az alábbi adódik:

$$(2.3.17) \quad A(\mathbf{x}; \mathbf{y}) = \frac{1}{4} \{ A(\mathbf{x} + \mathbf{y}; \mathbf{x} + \mathbf{y}) + iA(\mathbf{x} + i\mathbf{y}; \mathbf{x} + i\mathbf{y}) - \\ - A(\mathbf{x} - \mathbf{y}; \mathbf{x} - \mathbf{y}) - iA(\mathbf{x} - i\mathbf{y}; \mathbf{x} - i\mathbf{y}) \}.$$

A jobb oldalon álló kifejezés a kvadratikus alaknak az $\mathbf{x} + \mathbf{y}$, $\mathbf{x} - \mathbf{y}$, $\mathbf{x} + i\mathbf{y}$ és $\mathbf{x} - i\mathbf{y}$ vektorokra felvett értékeiből áll, tehát az $A(\mathbf{x}; \mathbf{y})$ bilineáris alakot a hozzá tartozó kvadratikus alak valóban egyértelműen meghatározza. ■

Szükségünk lesz a továbbiakban $A(\mathbf{y}; \mathbf{x})$ kifejezésére is. Ezt úgy nyerjük, ha a (2.3.13)–(2.3.16) azonosságokat rendre az 1, $-i$, -1 , i számokkal szorozzuk és összeadjuk. Ebből adódik, hogy

$$(2.3.18) \quad A(\mathbf{y}; \mathbf{x}) = \frac{1}{4} \{ A(\mathbf{x} + \mathbf{y}; \mathbf{x} + \mathbf{y}) - iA(\mathbf{x} + i\mathbf{y}; \mathbf{x} + i\mathbf{y}) - \\ - A(\mathbf{x} - \mathbf{y}; \mathbf{x} - \mathbf{y}) + iA(\mathbf{x} - i\mathbf{y}; \mathbf{x} - i\mathbf{y}) \}.$$

2.3.4 definíció. A komplex lineáris térben az $A(\mathbf{x}; \mathbf{y})$ bilineáris alakot *Hermite-félének* vagy **hermitikusnak** nevezzük, ha fennáll az

$$(2.3.19) \quad A(\mathbf{x}; \mathbf{y}) = \overline{A(\mathbf{y}; \mathbf{x})}$$

összefüggés.

2.3.2 tétel. Az $A(\mathbf{x}; \mathbf{y})$ bilineáris alak akkor és csak akkor hermitikus, ha mátrixa valamilyen bázisban hermitikus mátrix.

Bizonyítás. Tegyük fel, hogy az $A(\mathbf{x}; \mathbf{y})$ bilineáris alak hermitikus, és legyen az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisra vonatkozó mátrixa $\mathbf{A} = [a_{ij}]$. A (2.3.19) összefüggés alapján

$$a_{ij} = A(\mathbf{e}_j; \mathbf{e}_i) = \overline{A(\mathbf{e}_i; \mathbf{e}_j)} = \overline{a_{ji}}.$$

Most megfordítva: tegyük fel, hogy \mathbf{A} hermitikus mátrix, azaz $\mathbf{A}^H = \mathbf{A}$, vagyis $(a_{ij} = \overline{a_{ji}})$. Ekkor

$$A(\mathbf{x}; \mathbf{y}) = \sum_{j=1}^n \sum_{i=1}^n \overline{\eta_i} a_{ij} \xi_j = \overline{\sum_{j=1}^n \sum_{i=1}^n \eta_i a_{ji} \overline{\xi_j}} = \overline{A(\mathbf{y}; \mathbf{x})},$$

és ezzel a tételt bebizonyítottuk. ■

A tételből következik, hogy ha a bilineáris alak mátrixa valamilyen bázisban hermitikus mátrix, akkor bármely más bázisban is hermitikus.

2.3.5 definíció. Az hermitikus bilineáris alakhoz tartozó kvadratikus alakot **hermitikus alaknak** nevezzük.

A 2.3.2 tétellel ekvivalens tétel a következő:

2.3.3 tétel. Az $A(\mathbf{x}; \mathbf{y})$ bilineáris alak akkor és csak akkor hermitikus, ha a hozzá tartozó $A(\mathbf{x}; \mathbf{x})$ kvadratikus alak bármely \mathbf{x} vektorra valós.

Bizonyítás. Legyen $A(\mathbf{x}; \mathbf{y})$ hermitikus bilineáris alak, vagyis teljesüljön a (2.3.19) feltétel. Ha most \mathbf{y} helyére az \mathbf{x} vektort írjuk, akkor a feltétel alapján $A(\mathbf{x}; \mathbf{x}) = \overline{A(\mathbf{x}; \mathbf{x})}$, amiből következik, hogy $A(\mathbf{x}; \mathbf{x})$ valós.

Most megfordítva: legyen $A(\mathbf{x}; \mathbf{x})$ minden \mathbf{x} vektorra valós; tehát valós

$$A(\mathbf{x} + \mathbf{y}; \mathbf{x} + \mathbf{y}), \quad A(\mathbf{x} + i\mathbf{y}; \mathbf{x} + i\mathbf{y}), \quad A(\mathbf{x} - \mathbf{y}; \mathbf{x} - \mathbf{y}), \quad A(\mathbf{x} - i\mathbf{y}; \mathbf{x} - i\mathbf{y}).$$

A (2.3.17) és (2.3.18) azonosságok egybevetésével rögtön kiolvasható, hogy $A(\mathbf{x}; \mathbf{y})$ és $A(\mathbf{y}; \mathbf{x})$ egymásnak konjugáltjai, vagyis $A(\mathbf{x}; \mathbf{y})$ hermitikus. ■

Ebből a tételből közvetlenül adódik a 2.3.4 tétel.

2.3.4 tétel. Egy kvadratikus alak akkor és csak akkor hermitikus, ha csak valós értéket vehet fel.

Hermitikus alak például a komplex euklideszi tér egy \mathbf{x} vektorának önmagával való skaláris szorzata: (\mathbf{x}, \mathbf{x}) . Ugyanis a komplex euklideszi térben értelmezett skaláris szorzat (2.2.1) (a), (b), (c) tulajdonságai értelmében (\mathbf{x}, \mathbf{y}) hermitikus bilineáris alak, ezért (\mathbf{x}, \mathbf{x}) hermitikus alak.

2.3.6 definíció. *A komplex lineáris térben értelmezett $A(\mathbf{x}, \mathbf{x})$ hermitikus alakot pozitív definitnek nevezzük, ha minden $\mathbf{x} \neq 0$ vektorra fennáll az*

$$(2.3.20) \quad A(\mathbf{x}; \mathbf{x}) > 0$$

egyenlőtlenség.

Legyen most $A(\mathbf{x}; \mathbf{x})$ egy pozitív definit hermitikus alak, és $A(\mathbf{x}; \mathbf{y})$ a hozzá tartozó hermitikus bilineáris alak. A 2.3.2, 2.3.4 és a 2.3.6 definíció alapján ekkor

$$(2.3.21) \quad \begin{aligned} (a) \quad A(\mathbf{x}; \mathbf{y}) &= \overline{A(\mathbf{y}; \mathbf{x})}, \\ (b) \quad A(\mathbf{x}_1 + \mathbf{x}_2; \mathbf{y}) &= A(\mathbf{x}_1; \mathbf{y}) + A(\mathbf{x}_2; \mathbf{y}), \\ (c) \quad A(\lambda \mathbf{x}; \mathbf{y}) &= \lambda A(\mathbf{x}; \mathbf{y}), \\ (d) \quad A(\mathbf{x}; \mathbf{x}) &> 0, \quad \text{ha } \mathbf{x} \neq 0. \end{aligned}$$

Látjuk, hogy ezek a tulajdonságok megegyeznek a komplex lineáris térben bevezetett skaláris szorzat (2.2.1) alatti tulajdonságaival. Ebből következik, hogy a skaláris szorzat mindig pozitív definit hermitikus alakhoz tartozó bilineáris alak és minden ilyen bilineáris alak felfogható skaláris szorzatként, amellyel egy euklideszi teret definiálhatunk.

Ennek alapján tehát a következőképpen is értelmezhetjük a 2.2.2 definícióval bevezetett unitér vagy komplex euklideszi teret.

2.3.7 definíció. *Unitér térnek nevezzük az olyan komplex lineáris teret, amelyben ki van tüntetve egy rögzített $A(\mathbf{x}; \mathbf{x})$ pozitív definit hermitikus alak, és a hozzá tartozó $A(\mathbf{x}; \mathbf{y})$ bilineáris alakot az \mathbf{x} és \mathbf{y} vektor skaláris szorzataként értelmezzük.*

Beláttuk, hogy bármely $A(\mathbf{x}; \mathbf{x})$ pozitív definit hermitikus alakhoz tartozó $A(\mathbf{x}; \mathbf{y})$ bilineáris alak felfogható mint valamilyen euklideszi tér skaláris szorzata. Figyelembe véve a 2.2.3 definíciót, ha az $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ bázis olyan, hogy az $A(\mathbf{x}; \mathbf{y})$ bilineáris alak mint skaláris szorzat e bázisra vonatkozóan standard, akkor az $A(\mathbf{x}; \mathbf{y})$ bilineáris alaknak – s így az $A(\mathbf{x}; \mathbf{x})$ pozitív definit hermitikus alaknak – e bázisra vonatkozó mátrixa az egységmátrix.

2.3.8 definíció. *Ha az $A(\mathbf{x}; \mathbf{x})$ pozitív definit hermitikus alaknak, s így a hozzá tartozó $A(\mathbf{x}; \mathbf{y})$ bilineáris alaknak, az $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ bázisra vonatkozó mátrixa az \mathbf{E} egységmátrix, akkor az $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ bázist az $A(\mathbf{x}; \mathbf{x})$ hermitikus alak (ill. az $A(\mathbf{x}; \mathbf{y})$ bilineáris alak) standard bázisának nevezzük.*

A 2.2.6 tétel e definíció szerint éppen standard bázis létezését mondja ki. Így a 2.2.6 tétel általánosabb megfogalmazása hermitikus és bilineáris alakokra a következő.

2.3.5 tétel. *Bármely $A(\mathbf{x}; \mathbf{x})$ pozitív definit hermitikus alakhoz – s így a hozzá tartozó $A(\mathbf{x}; \mathbf{y})$ bilineáris alakhoz – megadható standard bázis.*

Megemlíttük még, hogy a komplex lineáris térben most megismert fogalmak és tételek értelemszerűen átvihetők a valós lineáris térre. Mivel a számértékű lineáris függvények esetén (lásd a 2.3.1 definíciót) az első- és másodfajú lineáris függvény fogalma egybeesik, ezáltal egyszerűbbé válik a bilineáris alak definíciója is; adott bázisban felírva a bilineáris alakot, elmarad a második helyen álló vektor koordinátáinak konjugálása.

Az *hermitikus* alak helyére a valós térben a *kvadratikus* alak lép, ennek mátrixa tetszőleges bázisban valós *szimmetrikus* mátrix*. Az hermitikus bilineáris alakot a valós térben szimmetrikus bilineáris alaknak nevezzük, tetszőleges bázisban ennek mátrixa is valós szimmetrikus mátrix. Egy adott $A(\mathbf{x}; \mathbf{x})$ kvadratikus alakhoz tartozó $A(\mathbf{x}; \mathbf{y})$ szimmetrikus bilineáris alakot a kvadratikus alak *poláris bilineáris alakjának* nevezzük.

A legfontosabb észrevétel, amit a valós térben definiált bilineáris és kvadratikus alakokra vonatkozó tételekkel kapcsolatban meg kell tennünk, az, hogy a 2.3.1 tétel nem érvényes minden változtatás nélkül a valós térben: nem igaz az, hogy minden bilineáris alakot egyértelműen meghatároz a hozzá tartozó kvadratikus alak. A tétel a valós térben a következőképpen módosul:

2.3.6 tétel. *A valós lineáris térben az $A(\mathbf{x}; \mathbf{y})$ poláris bilineáris alakot egyértelműen meghatározza a hozzá tartozó kvadratikus alak.*

Bizonyítás. A tétel bizonyításához kihasználjuk, hogy a poláris bilineáris alak szimmetrikus, azaz

$$(2.3.22) \quad A(\mathbf{x}; \mathbf{y}) = A(\mathbf{y}; \mathbf{x}).$$

A bilineáris alak értelmezéséből adódik, hogy

$$A(\mathbf{x} + \mathbf{y}; \mathbf{x} + \mathbf{y}) = A(\mathbf{x}; \mathbf{x}) + A(\mathbf{x}; \mathbf{y}) + A(\mathbf{y}; \mathbf{x}) + A(\mathbf{y}; \mathbf{y}),$$

és innen (2.3.22) behelyettesítésével és átrendezéssel

$$(2.3.23) \quad A(\mathbf{x}; \mathbf{y}) = \frac{1}{2} \{A(\mathbf{x} + \mathbf{y}; \mathbf{x} + \mathbf{y}) - A(\mathbf{x}; \mathbf{x}) - A(\mathbf{y}; \mathbf{y})\}.$$

Ebből már kiolvasható, hogy a kvadratikus alak valóban egyértelműen meghatározza poláris bilineáris alakját. ■

7. Példa. Tekintsük a (ξ_1, ξ_2, ξ_3) valós számhármassok terében megadott

$$A(\mathbf{y}; \mathbf{y}) = \xi_1 \eta_1 + 2\xi_2 \eta_2 + 3\xi_3 \eta_3$$

*Itt fel kell hívni az Olvasó figyelmét arra, hogy a *kvadratikus alak* kifejezés kétféle értelemben honosodott meg. A *komplex* lineáris térben a 2.3.3 definíció értelmében *tetszőleges* $A(\mathbf{x}; \mathbf{y})$ bilineáris alakhoz tartozó $A(\mathbf{x}; \mathbf{x})$ kifejezést nevezünk kvadratikus alaknak (ennek speciális esete az hermitikus alak, amelyhez hermitikus mátrix tartozik). A *valós* lineáris térben $A(\mathbf{x}; \mathbf{x})$ akkor kvadratikus alak, ha a hozzá tartozó mátrix valós szimmetrikus mátrix (ez tehát az hermitikus alak valós esete).

bilineáris alakot. Határozzuk meg ennek a mátrixát az

$$\mathbf{e}_1 = (2, -1, 1), \quad \mathbf{e}_2 = (1, 0, -2), \quad \mathbf{e}_3 = (0, 3, 1)$$

bázisra vonatkozóan!

Megoldás. A mátrix elemeit a (2.3.10) képlet segítségével számíthatjuk ki. Megjegyezzük, hogy mivel az adott bilineáris alak szimmetrikus, a keresett mátrix is szimmetrikus lesz.

$$\begin{aligned} a_{11} &= A(\mathbf{e}_1; \mathbf{e}_1) = 9, \\ a_{21} &= a_{12} = A(\mathbf{e}_1; \mathbf{e}_2) = -4, \\ a_{31} &= a_{13} = A(\mathbf{e}_1; \mathbf{e}_3) = -3, \\ a_{22} &= A(\mathbf{e}_2; \mathbf{e}_2) = 13, \\ a_{32} &= a_{23} = A(\mathbf{e}_2; \mathbf{e}_3) = -6, \\ a_{33} &= A(\mathbf{e}_3; \mathbf{e}_3) = 21. \end{aligned}$$

A bilineáris alak mátrixa az adott bázisra nézve tehát

$$\mathbf{A} = \begin{bmatrix} 9 & -4 & -3 \\ -4 & 13 & -6 \\ -3 & -6 & 21 \end{bmatrix}.$$

* * *

2.4 LINEÁRIS TRANSZFORMÁCIÓK

Ebben a szakaszban olyan lineáris függvényekkel foglalkozunk, amelyek az n -dimenziós \mathbf{R}_n komplex lineáris teret önmagába képezik le, azaz \mathbf{R}_n minden vektorához \mathbf{R}_n egy vektorát rendelik hozzá. Így jutunk a lineáris transzformáció fogalmához és megmutatjuk, hogy adott bázis esetén minden lineáris transzformációhoz egyértelműen hozzárendelhető egy mátrix. Ezért a bilineáris és kvadratikus alakok vizsgálatához hasonlóan, a lineáris transzformációk tárgyalásában is lényeges a mátrixelmélet alkalmazása. Defináljuk a lineáris transzformációk közötti műveleteket, majd megmutatjuk, hogy ezek a műveletek hogyan végezhetők el a lineáris transzformációkat adott bázisban reprezentáló mátrixokkal.

2.4.1 definíció. Ha az n -dimenziós \mathbf{R}_n komplex lineáris tér minden \mathbf{x} vektorának az \mathbf{R}_n tér egy $\mathbf{y} = \mathbf{A}(\mathbf{x})$ vektora felel meg és teljesülnek az alábbi feltételek:

$$(2.4.1) \quad \begin{aligned} \text{(a)} \quad \mathbf{A}(\mathbf{x}_1 + \mathbf{x}_2) &= \mathbf{A}(\mathbf{x}_1) + \mathbf{A}(\mathbf{x}_2); \\ \text{(b)} \quad \mathbf{A}(\lambda \mathbf{x}) &= \lambda \mathbf{A}(\mathbf{x}), \end{aligned}$$

akkor az $\mathbf{y} = \mathbf{A}(\mathbf{x})$ függvényt a tér **lineáris transzformációjának** nevezzük. A lineáris transzformáció jelölésére $\mathbf{A}(\mathbf{x})$ helyett ezentúl az \mathbf{Ax} szimbólumot használjuk*.

8. Példa. Tekintsük a geometriai térnek azt a transzformációját, amely a térnek az origón áthaladó valamilyen tengely körüli adott szögű elforgatása. Bármely \mathbf{x} vektornak az az \mathbf{Ax} vektor felel meg, amelybe az illető vektor az adott forgatás révén átmegy.

* * *

9. Példa. Tekintsük a komplex szám- n -esek lineáris terét. Minden $(\xi_1, \xi_2, \dots, \xi_n)$ szám- n -eshez azt az $\mathbf{A}(\xi_1, \xi_2, \dots, \xi_n) = (\eta_1, \eta_2, \dots, \eta_n)$ szám- n -est rendeljük, amelynek i -edik elemét az $\eta_i = \sum_{k=1}^n a_{ik} \xi_k$ képlet alapján számíthatjuk, ahol a_{ik} ($i, k = 1, 2, \dots, n$) adott komplex számok.

* * *

10. Példa. Tekintsük a legfeljebb n -edfokú polinomok $(n+1)$ -dimenziós terét és legyen $\mathbf{AP}(t) = P'(t)$, ahol $P'(t)$ a $P(t)$ polinom deriváltja.

* * *

11. Példa. Ha a $[0, 1]$ zárt intervallumban értelmezett $f(t)$ folytonos függvények terét tekintjük, akkor az $\mathbf{Af}(t) = \int_0^t f(\tau) d\tau$ transzformáció is lineáris.

* * *

Megjegyzés. Tekintettel arra, hogy az e könyv keretei között tárgyalt transzformációk kizárólag lineáris transzformációk, ezért a továbbiakban transzformáción mindig lineáris transzformációt értünk akkor is, ha – egyszerűség kedvéért – a „lineáris” jelzőt elhagyjuk.

2.4.1 tétel. Ha \mathbf{A} az \mathbf{R} tér lineáris transzformációja, akkor az \mathbf{Ax} alakú vektorok $\mathbf{R}^{(1)}$ halmaza, ahol \mathbf{x} befutja az egész \mathbf{R} teret, az \mathbf{R} tér egy alterét alkotja.

Bizonyítás. Legyen $\mathbf{y}_1 \in \mathbf{R}^{(1)}$ és $\mathbf{y}_2 \in \mathbf{R}^{(1)}$, vagyis $\mathbf{y}_1 = \mathbf{Ax}_1$, és $\mathbf{y}_2 = \mathbf{Ax}_2$. Ekkor $\mathbf{y}_1 + \mathbf{y}_2 = \mathbf{Ax}_1 + \mathbf{Ax}_2 = \mathbf{A}(\mathbf{x}_1 + \mathbf{x}_2)$, és így $\mathbf{y}_1 + \mathbf{y}_2 \in \mathbf{R}^{(1)}$. Hasonlóképpen, ha $\mathbf{y} = \mathbf{Ax}$, akkor $\lambda \mathbf{y} = \lambda \mathbf{Ax} = \mathbf{A}(\lambda \mathbf{x})$, tehát $\lambda \mathbf{y} \in \mathbf{R}^{(1)}$. Vagyis $\mathbf{R}^{(1)}$ valóban az \mathbf{R} tér altere. ■

2.4.2 definíció. Az \mathbf{R}_n lineáris tér $\mathbf{y} = \mathbf{Ax}$ alakú vektorai által alkotott alteret az \mathbf{A} transzformáció **képterének** nevezzük és $\mathbf{R}(\mathbf{A})$ -val jelöljük.**

*Lineáris transzformáción tágabb értelemben egy n -dimenziós lineáris tér leképezését értjük egy m -dimenziós lineáris térre (n és m nem szükségképpen egyenlő). E könyv keretei között csupán az n -dimenziós lineáris tér önmagába való leképezéseit vizsgáljuk.

**Az \mathbf{A} transzformáció képterére használatos az $\text{Im } \mathbf{A}$ jelölés is.

Az $\mathbf{R}(A)$ képtér tehát nem más, mint az $\mathbf{y} = A(\mathbf{x})$ függvény értékkészlete.

Ugyancsak könnyen belátható, hogy az \mathbf{R}_n lineáris tér azon $\mathbf{x} \in \mathbf{R}_n$ vektorainak a halmaza, amelyeket az A transzformáció a zérusvektorba képez le, vagyis amelyekre $A\mathbf{x} = \mathbf{0}$, az \mathbf{R}_n egy alterét alkotják.

2.4.3 definíció. Az \mathbf{R}_n lineáris tér azon $\mathbf{x} \in \mathbf{R}_n$ vektorai által alkotott alteret, amelyekre $A\mathbf{x} = \mathbf{0}$, az A transzformáció **nullterének (magterének)** nevezzük, és $N(A)$ -val jelöljük.*

Az alábbiakban megmutatjuk, hogy az n -dimenziós lineáris vektortérben értelmezett lineáris transzformációk mindig reprezentálhatók mátrixokkal, más szóval a mátrixok alkotják azt az apparátust, amellyel a lineáris transzformációk az n -dimenziós terekben egészen általánosan vizsgálhatók. Bebizonyítjuk e célból, hogy adott bázisban minden transzformációnak megfelel egy és csakis egy mátrix.

Legyen $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ az n -dimenziós tér egy bázisa és legyen A ennek a térnek egy lineáris transzformációja. Mindenekelőtt azt mutatjuk meg, hogy az A transzformációt egyértelműen meghatározzák a bázisvektorok transzformáltjai, legyenek ezek

$$(2.4.2) \quad \mathbf{g}_1 = A\mathbf{e}_1, \quad \mathbf{g}_2 = A\mathbf{e}_2, \quad \dots, \quad \mathbf{g}_n = A\mathbf{e}_n.$$

Ekkor bármely $\mathbf{x} = \xi_1 \mathbf{e}_1 + \xi_2 \mathbf{e}_2 + \dots + \xi_n \mathbf{e}_n$ vektor transzformáltja

$$\begin{aligned} A\mathbf{x} &= A(\xi_1 \mathbf{e}_1 + \xi_2 \mathbf{e}_2 + \dots + \xi_n \mathbf{e}_n) = \\ &= \xi_1 A\mathbf{e}_1 + \xi_2 A\mathbf{e}_2 + \dots + \xi_n A\mathbf{e}_n = \xi_1 \mathbf{g}_1 + \xi_2 \mathbf{g}_2 + \dots + \xi_n \mathbf{g}_n, \end{aligned}$$

vagyis $A\mathbf{x}$ valóban egyértelműen meghatározott a $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_n$ képvektorok ismeretében.

Ezen túlmenően még az is igaz, hogy *tetszőleges* $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_n$ vektorokhoz és $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázishoz létezik olyan A lineáris transzformáció, amelyre $A\mathbf{e}_1 = \mathbf{g}_1, A\mathbf{e}_2 = \mathbf{g}_2, \dots, A\mathbf{e}_n = \mathbf{g}_n$. Ha ugyanis az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ vektoroknak megfeleltetjük a $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_n$ vektorokat, és a tetszőleges $\mathbf{x} = \xi_1 \mathbf{e}_1 + \xi_2 \mathbf{e}_2 + \dots + \xi_n \mathbf{e}_n$ vektornak az $A\mathbf{x} = \xi_1 \mathbf{g}_1 + \xi_2 \mathbf{g}_2 + \dots + \xi_n \mathbf{g}_n$ vektort, akkor az \mathbf{x} vektor egyértelmű előállításá miatt $A\mathbf{x}$ is egyértelműen van meghatározva, és a transzformáció lineáris.

A transzformációt egyértelműen meghatározó \mathbf{g}_k vektoroknak az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisra vonatkozó koordinátáit jelölje $a_{1k}, a_{2k}, \dots, a_{nk}$, azaz

$$(2.4.3) \quad \mathbf{g}_k = A\mathbf{e}_k = \sum_{i=1}^n \mathbf{e}_i a_{ik} \quad (k = 1, 2, \dots, n);$$

* Az A transzformáció magterére használatos a $\text{Ker } A$ jelölés.

tekintsük ezeket a koordinátákat az $\mathbf{A} = [a_{ik}]$ mátrix elemeinek. Ezzel bebizonyítottuk a következő tételt.

2.4.2 tétel. *Bármely \mathbf{A} lineáris transzformációhoz egy adott bázisban a (2.4.3) összefüggés alapján egyértelműen hozzárendelhető egy mátrix és megfordítva, tetszőleges $[a_{ij}]$ mátrix egy adott bázisra vonatkozóan egyértelműen meghatároz egy lineáris transzformációt.*

A mondottakat szemléletesebbé tehetjük, ha a mátrixalgebrában definiált szorzásra bevezetett sor-oszlop konvenciót formálisan kiterjesztjük a bázisvektorokkal való szorzás esetére is. Ekkor ugyanis a (2.4.3) összefüggésben szereplő $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisvektorokat és ezek $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_n$ transzformáltjait egy-egy sorba, a lineáris kombinációk a_{ik} együtthatóit pedig értelemszerűen megfelelő oszlopokba elrendezve, a (2.4.3) összefüggés az alábbi alakban írható fel:

$$(2.4.4) \quad [\mathbf{g}_1 \quad \mathbf{g}_2 \dots \mathbf{g}_n] = [\mathbf{e}_1 \quad \mathbf{e}_2 \dots \mathbf{e}_n] \begin{bmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \vdots & \vdots \quad \vdots \\ a_{n1} & a_{n2} \dots a_{nn} \end{bmatrix}.$$

Ismételten hangsúlyozzuk, hogy itt nem „sorvektorokról” van szó, hanem a megfelelő vektorok „sorba” rendezéséről; ez a formális írásmód azonban meg fogja könnyíteni a további összefüggések felismerését.

Példaként írjuk fel néhány egyszerű lineáris transzformáció mátrixát egy adott bázisban.

12. Példa. Legyen \mathbf{R} a geometriai tér, \mathbf{A} pedig az a lineáris transzformáció, amely a tér vektorait egy derékszögű x_1, x_2, x_3 koordináta-rendszer x_1, x_2 koordinátasíkjára vetíti. Bázisként válasszuk a koordinátengelyek irányába mutató $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ egységvektorokat, ezek tehát a lineáris tér egy ortonormált bázisát alkotják. A vetítés során \mathbf{e}_1 és \mathbf{e}_2 nem változik meg, az \mathbf{e}_3 vektort viszont az adott transzformáció a zérusvektorba viszi át:

$$\mathbf{g}_1 = \mathbf{e}_1; \quad \mathbf{g}_2 = \mathbf{e}_2; \quad \mathbf{g}_3 = \mathbf{0}.$$

A transzformációt tehát a (2.4.4) összefüggésnek megfelelően felírva:

$$[\mathbf{g}_1 \quad \mathbf{g}_2 \quad \mathbf{g}_3] = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_3] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix};$$

ezért a transzformáció mátrixa: $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, oszlopai ugyanis $\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3$

koordinátáit tartalmazzák az $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ bázisban.

* * *

13. Példa. Legyen az R lineáris tér ismét a geometriai tér, és írjuk fel annak a transzformációnak a mátrixát, amely a tér vektorait egy derékszögű koordináta-rendszer x_3 tengelyére tükrözi. Bázisként válasszuk ismét a koordinátatengelyek irányába eső $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ egységvektorokat. A tükrözés itt azt jelenti, hogy \mathbf{e}_1 és \mathbf{e}_2 a transzformáció során előjelet vált, \mathbf{e}_3 pedig változatlan marad:

$$\mathbf{g}_1 = -\mathbf{e}_1; \quad \mathbf{g}_2 = -\mathbf{e}_2; \quad \mathbf{g}_3 = \mathbf{e}_3.$$

A transzformáció tehát a következő összefüggéssel jellemezhető:

$$[\mathbf{g}_1 \quad \mathbf{g}_2 \quad \mathbf{g}_3] = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_3] \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

ahol a transzformáció mátrixa

$$\mathbf{A} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

mivel oszlopai $\mathbf{g}_1, \mathbf{g}_2$ és \mathbf{g}_3 koordinátáit tartalmazzák az $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ bázisban.

* * *

14. Példa. Azt az E transzformációt, amely a lineáris tér minden vektorát önmagába transzformálja, egységtranszformációnak nevezzük. Mivel tetszőleges $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázis esetén ekkor

$$E\mathbf{e}_1 = \mathbf{e}_1, \quad E\mathbf{e}_2 = \mathbf{e}_2, \quad \dots \quad E\mathbf{e}_n = \mathbf{e}_n,$$

ez azt jelenti, hogy az E egységtranszformáció mátrixa *minden* bázisban és bármely lineáris tér esetén az egységmátrix:

$$\mathbf{E} = [\delta_{ik}].$$

* * *

15. Példa. Legyen R a legfeljebb $(n-1)$ -edfokú polinomok tere és a D transzformáció a differenciálás, vagyis $DP(t) = P'(t)$. Írjuk fel ennek a transzformációnak a mátrixát az

$$\mathbf{e}_1 = 1, \quad \mathbf{e}_2 = t, \quad \mathbf{e}_3 = \frac{t^2}{2!}, \quad \dots, \quad \mathbf{e}_n = \frac{t^{n-1}}{(n-1)!}$$

bázisban. A transzformációt alkalmazva a bázisvektorokra – vagyis deriválva a fenti polinomokat –, a következőket kapjuk a bázisvektorok transzformáltaként:

$$\begin{aligned} g_1 &= D e_1 = 1' = 0, \\ g_2 &= D e_2 = t' = 1 = e_1, \\ g_3 &= D e_3 = \left(\frac{t^2}{2!} \right)' = t = e_2, \\ &\dots\dots\dots \\ g_n &= D e_n = \left(\frac{t^{n-1}}{(n-1)!} \right)' = \frac{t^{n-2}}{(n-2)!} = e_{n-1}. \end{aligned}$$

A transzformáció mátrixa tehát

$$\mathbf{D} = \begin{bmatrix} 0 & 1 & 0 & 0 & . & 0 \\ 0 & 0 & 1 & 0 & . & 0 \\ 0 & 0 & 0 & 1 & . & 0 \\ \\ 0 & . & . & . & . & 1 \\ 0 & 0 & 0 & 0 & . & 0 \end{bmatrix}.$$

A következőkben definiáljuk a lineáris transzformációk közötti műveleteket, és megmutatjuk, hogy adott bázisban miként végezhetők el ezek a műveletek a lineáris transzformációk mátrixával.

2.4.4 definíció. Ha a lineáris tér bármely \mathbf{x} vektorára alkalmazzuk az A és B transzformációt, akkor azt a C transzformációt, amelyet az \mathbf{x} vektorra alkalmazni kell, hogy az $A\mathbf{x}$ és a $B\mathbf{x}$ vektorok összegét kapjuk, az A és B transzformációk összegének nevezzük. Tehát $C = A + B$ azt jelenti, hogy

$$Cx = Ax + Bx.$$

A transzformációk összegének mátrixát a következő tétel alapján kaphatjuk meg.

2.4.3 tétel. Ha A és B az R_n lineáris tér két transzformációja, akkor ezek összegének adott bázisra vonatkozó mátrixa egyenlő az A és B transzformációk ugyanazon bázisra vonatkozó mátrixának az összegével.

Bizonyítás. A (2.4.3) összefüggés felhasználásával írható, hogy

$$Ae_k = \sum_{i=1}^n e_i a_{ik}; \quad Be_k = \sum_{i=1}^n e_i b_{ik}; \quad Ce_k = \sum_{i=1}^n e_i c_{ik}.$$

Így a $C = A + B$ transzformáció definíciója alapján

$$C\mathbf{e}_k = A\mathbf{e}_k + B\mathbf{e}_k = \sum_{i=1}^n \mathbf{e}_i(a_{ik} + b_{ik}) = \sum_{i=1}^n \mathbf{e}_i c_{ik},$$

amiből az egyértelmű előállítás miatt következik

$$c_{ik} = a_{ik} + b_{ik}, \quad i, k = 1, \dots, n, \quad \text{azaz} \quad C = A + B. \quad \blacksquare$$

2.4.5 definíció. Az A lineáris transzformációnak λ számmal való szorzatán azt a λA lineáris transzformációt értjük, amely a tér bármely \mathbf{x} vektorához a $\lambda(A\mathbf{x})$ vektort rendeli.

A definícióból következik, hogy ha az A transzformáció mátrixa $\mathbf{A} = [a_{ik}]$, akkor a λA transzformáció mátrixa $\lambda \mathbf{A} = [\lambda a_{ik}]$.

2.4.6 definíció. Ha a lineáris tér bármely \mathbf{x} vektorára alkalmazzuk a B transzformációt: $\mathbf{y} = B\mathbf{x}$, majd a transzformációval kapott \mathbf{y} vektorra az A transzformációt: $\mathbf{z} = A\mathbf{y}$, akkor azt a C transzformációt, amelyet az \mathbf{x} vektorra alkalmazni kell, hogy az egymás után végrehajtott transzformációkkal nyert \mathbf{z} vektort kapjuk, az A és B transzformációk szorzatának nevezzük.

Tehát $C = AB$ azt jelenti, hogy minden \mathbf{x} vektorra fennáll

$$C\mathbf{x} = A(B\mathbf{x}), \quad \text{azaz} \quad AB\mathbf{x} = A(B\mathbf{x}).$$

Lineáris transzformációk szorzata is lineáris transzformáció, mert teljesülnek rá a (2.4.1) összefüggések. Ugyanis kihasználva azt, hogy az A és a B transzformáció is lineáris, a következőt kapjuk:

$$\begin{aligned} C(\mathbf{x}_1 + \mathbf{x}_2) &= A[B(\mathbf{x}_1 + \mathbf{x}_2)] = \\ &= A(B\mathbf{x}_1 + B\mathbf{x}_2) = A(B\mathbf{x}_1) + A(B\mathbf{x}_2) = C\mathbf{x}_1 + C\mathbf{x}_2, \end{aligned}$$

és hasonlóan, $C(\lambda\mathbf{x}) = A[B(\lambda\mathbf{x})] = A(\lambda B\mathbf{x}) = \lambda A(B\mathbf{x}) = \lambda C\mathbf{x}$. Ha E az egységtranszformáció, A pedig tetszőleges transzformáció, akkor nyilvánvalóan fennáll, hogy

$$AE = EA = A.$$

Értelmezhetők a lineáris transzformáció pozitív egész kitevőjű hatványai is:

$$A^2 = AA, \quad A^3 = AA^2, \quad \dots$$

Definíció szerint $A^0 = E$, továbbá nyilvánvaló, hogy $A^{p+q} = A^p A^q$.

16. Példa. Tekintsük a legfeljebb $(n-1)$ -edfokú $P(t)$ polinomok terét, és legyen a tér egy D lineáris transzformációja a deriválás: $DP(t) = P'(t)$. A D

transzformáció ismételt alkalmazásával kapjuk a transzformáció hatványait, ezek esetünkben a magasabbrendű deriválásnak felelnek meg:

$$D^2 P(t) = D(DP(t)) = (P'(t))' = P''(t).$$

* * *

Megjegyzés. A legfeljebb $(n - 1)$ -edfokú polinomok terében mindig fennáll

$$D^n P(t) = P^{(n)}(t) = 0.$$

Vizsgáljuk most meg, hogy adott e_1, e_2, \dots, e_n bázisban hogyan határozható meg az A és B transzformáció mátrixának ismeretében a $C = AB$ szorzattranszformáció mátrixa.

2.4.4 tétel. *Ha A és B az R_n tér két lineáris transzformációja, akkor ezek szorzatának adott bázisra vonatkozó mátrixa egyenlő az A és B transzformáció ugyanezen bázisra vonatkozó mátrixának – ugyanolyan sorrendben vett – szorzatával.*

Bizonyítás. Legyen az A transzformáció mátrixa az adott bázisban $A = [a_{ik}]$, a B transzformációé pedig $B = [b_{ik}]$.

A bázisvektorok és ezeknek a B transzformációval nyert leképezése közötti összefüggés (2.4.4) szerint:

$$[Be_1 \quad Be_2 \dots Be_n] = [e_1 \quad e_2 \dots e_n] \begin{bmatrix} b_{11} & b_{12} \dots b_{1n} \\ b_{21} & b_{22} \dots b_{2n} \\ \dots & \dots \\ b_{n1} & b_{n2} \dots b_{nn} \end{bmatrix}.$$

Alkalmazzuk az A transzformációt a kapott azonosság mindkét oldalán:

$$(2.4.5) \quad [A(Be_1) \quad A(Be_2) \dots A(Be_n)] = [Ae_1 \quad Ae_2 \dots Ae_n] \begin{bmatrix} b_{11} & b_{12} \dots b_{1n} \\ b_{21} & b_{22} \dots b_{2n} \\ \dots & \dots \\ b_{n1} & b_{n2} \dots b_{nn} \end{bmatrix}.$$

Vegyük most figyelembe, hogy a bázisvektorok és ezeknek az A transzformációval adódó leképezése közötti összefüggés

$$(2.4.6) \quad [Ae_1 \quad Ae_2 \dots Ae_n] = [e_1 \quad e_2 \dots e_n] \begin{bmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \dots & \dots \\ a_{n1} & a_{n2} \dots a_{nn} \end{bmatrix}.$$

Jelölje továbbá a $C = AB$ szorzattranszformáció mátrixát $C = [c_{ik}]$, így a (2.4.4) szerint

$$(2.4.7) \quad [Ce_1 \quad Ce_2 \dots Ce_n] = [e_1 \quad e_2 \dots e_n] \begin{bmatrix} c_{11} & c_{12} \dots c_{1n} \\ c_{21} & c_{22} \dots c_{2n} \\ \dots & \dots \\ c_{n1} & c_{n2} \dots c_{nn} \end{bmatrix}.$$

Behelyettesítve a (2.4.6) kifejezést a (2.4.5) összefüggésbe, (2.4.5) és (2.4.7) azonossága miatt fennáll, hogy

$$\begin{aligned}
 & [e_1 \quad e_2 \dots e_n] \begin{bmatrix} c_{11} & c_{12} \dots c_{1n} \\ c_{21} & c_{22} \dots c_{2n} \\ \dots & \dots \\ c_{n1} & c_{n2} \dots c_{nn} \end{bmatrix} = \\
 & = [e_1 \quad e_2 \dots e_n] \begin{bmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \dots & \dots \\ a_{n1} & a_{n2} \dots a_{nn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \dots b_{1n} \\ b_{21} & b_{22} \dots b_{2n} \\ \dots & \dots \\ b_{n1} & b_{n2} \dots b_{nn} \end{bmatrix}.
 \end{aligned}$$

Mivel a tér bármely vektora a bázisvektorok segítségével egyértelműen állítható elő, a kapott azonosság csak akkor állhat fenn, ha a bázisvektorok szorzói az azonosság két oldalán megegyeznek, tehát az A , B és $C = AB$ transzformációk mátrixára fenn áll az alábbi összefüggés:

$$(2.4.8) \quad C = AB.$$

Ezzel bebizonyítottuk a tételt. ■

A 2.4.2 tétel szerint rögzített bázis mellett az R_n tér lineáris transzformációi és komplex elemű n -edrendű mátrixok között kölcsönösen egyértelmű összefüggés áll fenn. Ezért a mátrixok alaptulajdonságaiból következnek a lineáris transzformációkra vonatkozó műveletek tulajdonságai: az összeadás kommutatív és asszociatív, a szorzás asszociatív, de nem kommutatív (azaz $AB \neq BA$), az összeadás és szorzás disztributív.

A 2.4.4–2.4.6 definíciók alapján most már értelmezhető a lineáris transzformáció polinomja is. Ha ugyanis adott a

$$P(t) = a_0 + a_1 t + a_2 t^2 + \dots + a_{m-1} t^{m-1}$$

valós (vagy komplex) együtthatós, skalár változójú polinom, akkor a $P(A)$ polinom a

$$P(A) = a_0 E + a_1 A + a_2 A^2 + \dots + a_{m-1} A^{m-1}$$

lineáris transzformációt értjük, ahol E az egységtranszformáció. Ennek mátrixa – a 2.4.4–2.4.6 definíciók szerint – az A transzformáció A mátrixának polinomja:

$$P(A) = a_0 E + a_1 A + a_2 A^2 + \dots + a_{m-1} A^{m-1},$$

ahol E az egységmátrix.

Végül definiáljuk az inverz transzformáció fogalmát, és megadjuk az inverz transzformáció létezésének szükséges és elégséges feltételét.

2.4.7 definíció. Ha az \mathbf{R}_n tér adott \mathbf{A} lineáris transzformációja olyan, hogy az $\mathbf{A}\mathbf{x}_1 = \mathbf{A}\mathbf{x}_2$ egyenlőség csak $\mathbf{x}_1 = \mathbf{x}_2$ esetén teljesül, akkor létezik az \mathbf{R}_n tér egy \mathbf{B} transzformációja a következő tulajdonsággal: $\mathbf{B}(\mathbf{x})$ az \mathbf{R}_n tér egyetlen vektora, amelyhez az \mathbf{A} lineáris transzformáció az \mathbf{x} vektort rendeli, azaz amelyre

$$\mathbf{A}\mathbf{B}(\mathbf{x}) = \mathbf{x}.$$

Ezt a $\mathbf{B}(\mathbf{x})$ transzformációt az \mathbf{A} **inverzének** nevezzük és \mathbf{A}^{-1} -gyel jelöljük. A fenti tulajdonságú \mathbf{A} lineáris transzformációt **nemszingulárisnak** mondjuk, ha nem ilyen tulajdonságú, akkor **szinguláris**.

Megmutatjuk, hogy az inverz transzformáció is lineáris.

2.4.5 tétel. Ha \mathbf{A} az \mathbf{R}_n tér nemszinguláris lineáris transzformációja, akkor inverze, az \mathbf{A}^{-1} transzformáció is lineáris.

Bizonyítás. Ha \mathbf{A} nemszinguláris, akkor a definíció alapján tetszőleges $\mathbf{y}_1, \mathbf{y}_2 \in \mathbf{R}_n$ vektorokhoz léteznek olyan $\mathbf{x}_1, \mathbf{x}_2 \in \mathbf{R}_n$ vektorok, amelyekre $\mathbf{A}\mathbf{x}_1 = \mathbf{y}_1$, $\mathbf{A}\mathbf{x}_2 = \mathbf{y}_2$. Ugyancsak az inverz definíciója szerint $\mathbf{A}^{-1}(\mathbf{y}_1)$ az egyetlen vektor, amelyhez az \mathbf{A} transzformáció az \mathbf{y}_1 vektort rendeli, ezért

$$(2.4.9) \quad \mathbf{x}_1 = \mathbf{A}^{-1}(\mathbf{y}_1), \quad \mathbf{x}_2 = \mathbf{A}^{-1}(\mathbf{y}_2).$$

Hasonlóan, az $\mathbf{A}(\mathbf{x}_1 + \mathbf{x}_2) = \mathbf{y}_1 + \mathbf{y}_2$ összefüggésből a (2.4.9) figyelembevételével nyerjük az

$$(2.4.10) \quad \mathbf{A}^{-1}(\mathbf{y}_1 + \mathbf{y}_2) = \mathbf{x}_1 + \mathbf{x}_2 = \mathbf{A}^{-1}(\mathbf{y}_1) + \mathbf{A}^{-1}(\mathbf{y}_2)$$

összefüggést. Mivel tetszőleges komplex λ esetén $\mathbf{A}(\lambda\mathbf{x}_1) = \lambda\mathbf{y}_1$, így ismét az inverz transzformáció definíciója alapján

$$(2.4.11) \quad \lambda\mathbf{x}_1 = \mathbf{A}^{-1}(\lambda\mathbf{y}_1).$$

A (2.4.10) és (2.4.11) szerint tehát \mathbf{A}^{-1} is lineáris. ■

2.4.6 tétel. Az \mathbf{R}_n tér \mathbf{A} lineáris transzformációja akkor és csak akkor nemszinguláris, ha az \mathbf{R}_n tér bármely $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$ bázisa esetén az $\mathbf{A}\mathbf{f}_1, \mathbf{A}\mathbf{f}_2, \dots, \mathbf{A}\mathbf{f}_n$ vektorok ismét bázist alkotnak.

Bizonyítás. Tegyük fel először, hogy \mathbf{A} az \mathbf{R}_n tér nemszinguláris transzformációja és legyen $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$ a tér egy bázisa. Az $\mathbf{A}\mathbf{f}_1, \mathbf{A}\mathbf{f}_2, \dots, \mathbf{A}\mathbf{f}_n$ transzformált vektorok lineárisan függetlenek – azaz bázist alkotnak –, ha a

$$(2.4.12) \quad \sum_{i=1}^n c_i \mathbf{A}\mathbf{f}_i = \mathbf{0}$$

egyenletből

$$(2.4.13) \quad c_1 = c_2 = \dots = c_n = 0$$

következik. Mivel a lineáris transzformációk definíciója értelmében $\sum_{i=1}^n c_i \mathbf{A} \mathbf{f}_i =$
 $= \mathbf{A} \left(\sum_{i=1}^n c_i \mathbf{f}_i \right)$, továbbá, ha \mathbf{A} nemszinguláris, akkor az $\mathbf{A} \left(\sum_{i=1}^n c_i \mathbf{f}_i \right) = \mathbf{0}$
 egyenlet fennállásából következik

$$(2.4.14) \quad \sum_{i=1}^n c_i \mathbf{f}_i = \mathbf{0},$$

ezért a (2.4.12) egyenletből (2.4.14), és az \mathbf{f}_i bázisvektorok lineáris függetlenségéből (2.4.13) adódik.

Most megfordítva, tegyük fel, hogy az $\mathbf{A} \mathbf{f}_1, \mathbf{A} \mathbf{f}_2, \dots, \mathbf{A} \mathbf{f}_n$ vektorok lineárisan függetlenek, hacsak $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$ az \mathbf{R}_n tér egy bázisa. Azt kell belátnunk, hogy \mathbf{A} nemszinguláris, vagyis hogy – a 2.4.7 definíció alapján – a tér bármely két különböző vektorát különböző vektorokba transzformálja. Ezt indirekt módszerrel látjuk be: legyen a tér egy-egy vektora

$$\mathbf{x} = \sum_{i=1}^n c_i \mathbf{f}_i \quad \text{és} \quad \mathbf{y} = \sum_{i=1}^n d_i \mathbf{f}_i,$$

amelyekre a transzformációt alkalmazva, fennáll:

$$\mathbf{A} \left(\sum_{i=1}^n c_i \mathbf{f}_i \right) = \mathbf{A} \left(\sum_{i=1}^n d_i \mathbf{f}_i \right).$$

Ekkor ismét a lineáris transzformációk definíciója alapján kapjuk, hogy
 $\sum_{i=1}^n (c_i - d_i) \mathbf{A} \mathbf{f}_i = \mathbf{0}$; innen az $\mathbf{A} \mathbf{f}_i$ vektorok lineáris függetlensége miatt $c_i = d_i$
 $(i = 1, 2, \dots, n)$, és ezzel a tételt bebizonyítottuk. ■

A 2.4.2 tétel szerint az \mathbf{R}_n térben a lineáris transzformáció és tetszőleges adott bázisra vonatkozó mátrixa között kölcsönösen egyértelmű kapcsolat áll fenn. Ezért az inverz mátrix létezésére vonatkozó 1.2.4 tételből közvetlenül adódik az inverz transzformáció létezésére vonatkozó következő tétel.

2.4.7 tétel. *Az \mathbf{A} lineáris transzformációnak akkor és csak akkor létezik inverze, ha mátrixa valamilyen bázisban nemszinguláris mátrix.*

Bizonyítás. Az inverz transzformáció 2.4.7 definíciójából következik, hogy ha az \mathbf{A} invertálható transzformáció inverzét \mathbf{B} -vel és az egységtranszformációt

E-vel jelöljük, akkor $BA = E$. Legyen egy tetszőleges felvett bázisban a B , ill. az A transzformáció mátrixa B , ill. A (az egységtranszformáció mátrixa, mint tudjuk, minden bázisban E). Ekkor a 2.4.4 tétel felhasználásával

$$(2.4.15) \quad BA = E$$

írható. Az 1.5.3 tétel szerint a (2.4.15) egyenlet akkor és csak akkor oldható meg B -re, ha A nonszinguláris, és ekkor $B = A^{-1}$, ami egyértelműen meghatározza a $B = A^{-1}$ inverz transzformációt. Ezzel a tételt bebizonyítottuk. ■

A bizonyításból, mint látjuk, az is következik, hogy ha az A transzformáció mátrixa egy adott bázisban a nonszinguláris A mátrix, akkor az A^{-1} inverz transzformációnak a mátrixa ugyanebben a bázisban A^{-1} . Az inverz mátrix tulajdonságaiból az is következik, hogy ha létezik az A^{-1} inverz transzformáció, akkor az egyértelmű, továbbá, az adott A transzformáció és az inverze a szorzásra nézve kommutatív, vagyis $AA^{-1} = A^{-1}A = E$.

A következő tétel a transzformáció képterének dimenziójára vonatkozik.

2.4.8 tétel. *Ha A az R_n tér egy lineáris transzformációja, akkor az Ax vektorok által meghatározott $R^{(1)}$ altér dimenziója egyenlő az A transzformáció tetszőleges bázishoz tartozó mátrixának rangjával.*

Bizonyítás. A 2.4.1 tételben bebizonyítottuk, hogy ha x az egész R_n teret befutja, akkor az Ax vektorok halmaza az R_n tér egy $R^{(1)}$ altérét alkotja. Legyen e_1, e_2, \dots, e_n az R_n tér egy tetszőleges bázisa. Bármely x vektor a bázis vektorainak lineáris kombinációja, ezért az $R^{(1)}$ altér bármely Ax vektora az Ae_1, Ae_2, \dots, Ae_n vektorok lineáris kombinációja. Legyen az Ae_1, Ae_2, \dots, Ae_n vektorokból kiválasztható lineárisan független vektorok maximális száma r . Ez azt jelenti, hogy kiválasztva az Ae_1, Ae_2, \dots, Ae_n vektorok közül r lineárisan független vektort, az összes többi Ae_i vektor – és ennek következtében az $R^{(1)}$ altér minden vektora is – kifejezhető ezek lineáris kombinációjaként. Tehát az $R^{(1)}$ altér r -dimenziós. Ha most az Ae_1, Ae_2, \dots, Ae_n vektorokat az e_1, e_2, \dots, e_n bázisra vonatkozó koordinátáikkal írjuk fel, akkor ezek éppen az A transzformáció e_1, e_2, \dots, e_n bázisra vonatkozó A mátrixának az oszlopai lesznek. Így azt kapjuk, hogy az A transzformációhoz tartozó A mátrix lineárisan független oszlopainak maximális száma, vagyis a mátrix rangja egyenlő az $R^{(1)}$ altér dimenziójával. Mivel ez utóbbi nem függ a bázistól, ezért az A transzformáció mátrixának a rangja valóban invariáns a bázis megválasztásával szemben; ezzel a tételt bebizonyítottuk. ■

Ez a tétel abban az esetben, ha az A transzformáció az R_n teret önmagára képezi le, azt jelenti, hogy a transzformációnak bármely bázishoz tartozó mátrixa nonszinguláris. Ha pedig az $R^{(1)}$ altér dimenziója $r < n$, akkor bármely bázishoz tartozó mátrixa r -edrangu szinguláris mátrix.

amit formálisan sor-oszlop kompozícióként is felírhatunk:

$$(2.5.3) \quad \mathbf{x} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{e}}_1 & \tilde{\mathbf{e}}_2 & \dots & \tilde{\mathbf{e}}_n \end{bmatrix} \begin{bmatrix} \tilde{\xi}_1 \\ \tilde{\xi}_2 \\ \vdots \\ \tilde{\xi}_n \end{bmatrix}.$$

Feladatunk a $\tilde{\xi}_k$ és a ξ_i koordináták közötti összefüggés meghatározása. Helyettesítsük e célból a (2.5.2) kifejezést a (2.5.3) összefüggés jobb oldalán álló $[\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \dots, \tilde{\mathbf{e}}_n]$ helyébe:

$$(2.5.4) \quad \mathbf{x} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} \dots c_{1n} \\ c_{21} & c_{22} \dots c_{2n} \\ \dots & \dots \dots \dots \\ c_{n1} & c_{n2} \dots c_{nn} \end{bmatrix} \begin{bmatrix} \tilde{\xi}_1 \\ \tilde{\xi}_2 \\ \vdots \\ \tilde{\xi}_n \end{bmatrix}.$$

Most már a kapott egyenlet mindkét oldalán az \mathbf{x} vektornak *ugyanabban* a bázisban való kifejezése áll, és mivel a tér bármely vektora a bázis vektorainak lineáris kombinációjaként *egyértelműen* állítható elő (lásd 2.1.2 tétel), ezért

$$(2.5.5) \quad \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \dots c_{1n} \\ c_{21} & c_{22} \dots c_{2n} \\ \dots & \dots \dots \dots \\ c_{n1} & c_{n2} \dots c_{nn} \end{bmatrix} \begin{bmatrix} \tilde{\xi}_1 \\ \tilde{\xi}_2 \\ \vdots \\ \tilde{\xi}_n \end{bmatrix}.$$

Ezzel az \mathbf{x} vektornak az eredeti és az új bázisra vonatkozó koordinátái közötti összefüggést lényegében megkaptuk. Mi azonban a $\tilde{\xi}_k$ új koordinátákat szeretnénk a ξ_i eredeti koordináták segítségével kifejezni. Az új bázis vektorainak lineáris függetlenségéből következik, hogy a bázistranszformáció \mathbf{C} mátrixa nemszinguláris, tehát invertálható; így a (2.5.5) összefüggést balról a \mathbf{C}^{-1} mátrixszal szorozhatjuk és a következőt kapjuk:

$$(2.5.6) \quad \begin{bmatrix} \tilde{\xi}_1 \\ \tilde{\xi}_2 \\ \vdots \\ \tilde{\xi}_n \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \dots c_{1n} \\ c_{21} & c_{22} \dots c_{2n} \\ \dots & \dots \dots \dots \\ c_{n1} & c_{n2} \dots c_{nn} \end{bmatrix}^{-1} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}.$$

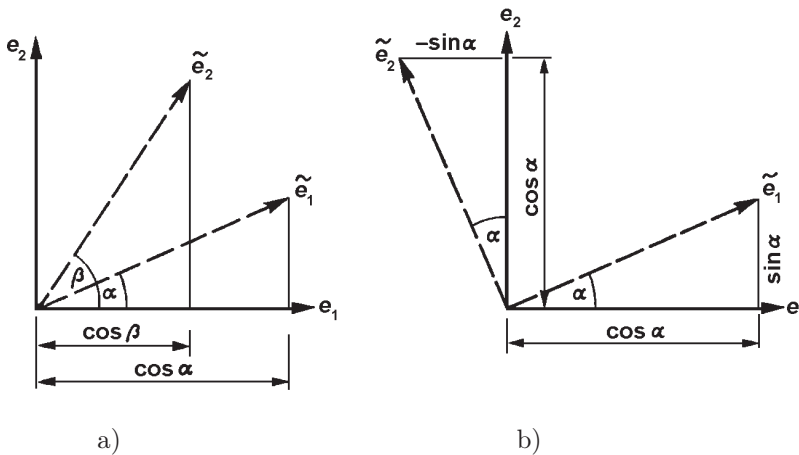
Érvényes tehát a következő tétel.

2.5.1 tétel. *Adott vektornak új bázisra vonatkozó koordinátáit az eredeti bázisra vonatkozó koordinátáiból úgy kapjuk meg, hogy a bázistranszformáció mátrixának inverzével szorozzuk az eredeti koordináták oszlopvektorát.*

Hangsúlyozni kell, hogy ebben a feladatban a tér \mathbf{x} vektora tetszőleges, de *rögzített*, az új bázisra való áttérés ennek a rögzített vektornak a különböző bázisokra vonatkozó koordinátái közötti összefüggést adja meg. Éppen ezért a bázis transzformációját másképpen – az analitikus geometria szóhasználatát átvéve – a *koordináta-rendszer* transzformációjának is nevezhetjük.

Megjegyzés. Az elmondottakat a következőkben egy egyszerű síkgeometriai koordináta-transzformációs példán világítjuk meg. Előrebocsátjuk azonban, hogy ennek az egyszerű példának alapvető szerepe lesz a későbbiekben, a többdimenziós koordinátatranszformációkkal kapcsolatos bonyolultabb feladatok tárgyalásában.

17. Példa. Tekintsük a sík vektorait, és eredeti $\mathbf{e}_1, \mathbf{e}_2$ bázisként válasszuk a derékszögű koordináta-rendszer egységvektorait. Az új $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2$ bázisvektorok olyan egységvektorok legyenek, amelyek az \mathbf{e}_1 vektorral α , ill. β szöget zárnak be (lásd az ábrát).



Határozzuk meg a transzformáció mátrixát (a) az általános esetre; (b) $\beta = \alpha + \frac{\pi}{2}$ esetre.

Megoldás. (a) Az eredeti és az új bázis vektorai közötti összefüggés a következő:

$$\begin{bmatrix} \tilde{\mathbf{e}}_1 & \tilde{\mathbf{e}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 \end{bmatrix} \begin{bmatrix} \cos \alpha & \cos \beta \\ \sin \alpha & \sin \beta \end{bmatrix},$$

vagyis a bázistranszformáció (koordinátatranszformáció) mátrixa

$$(2.5.7) \quad \mathbf{C} = \begin{bmatrix} \cos \alpha & \cos \beta \\ \sin \alpha & \sin \beta \end{bmatrix}.$$

Mivel ennek determinánsa $|\mathbf{C}| = \sin(\beta - \alpha)$, az inverze ($\beta \neq \alpha$ esetén)

$$(2.5.8) \quad \mathbf{C}^{-1} = \frac{1}{\sin(\beta - \alpha)} \begin{bmatrix} \sin \beta & -\cos \beta \\ -\sin \alpha & \cos \alpha \end{bmatrix}.$$

(Ha $\beta = \alpha$, akkor $\tilde{\mathbf{e}}_1 = \tilde{\mathbf{e}}_2$, tehát nem alkotnak bázist.) A sík bármely \mathbf{x} vektorának az új bázisra vonatkozó koordinátái tehát a következőképpen számíthatók ki:

$$\begin{bmatrix} \tilde{\xi}_1 \\ \tilde{\xi}_2 \end{bmatrix} = \frac{1}{\sin(\beta - \alpha)} \begin{bmatrix} \sin \beta & -\cos \beta \\ -\sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix};$$

vagyis

$$(2.5.9) \quad \begin{aligned} \tilde{\xi}_1 &= \frac{\sin \beta}{\sin(\beta - \alpha)} \xi_1 - \frac{\cos \beta}{\sin(\beta - \alpha)} \xi_2; \\ \tilde{\xi}_2 &= -\frac{\sin \alpha}{\sin(\beta - \alpha)} \xi_1 + \frac{\cos \alpha}{\sin(\beta - \alpha)} \xi_2. \end{aligned}$$

(b) Abban a speciális esetben, amikor $\beta = \alpha + \frac{\pi}{2}$, tehát amikor a koordináta-rendszert pozitív irányban α a szöggel *elforgatjuk*, a bázistranszformáció mátrixa

$$(2.5.10) \quad \mathbf{Q} = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix},$$

ennek inverze pedig

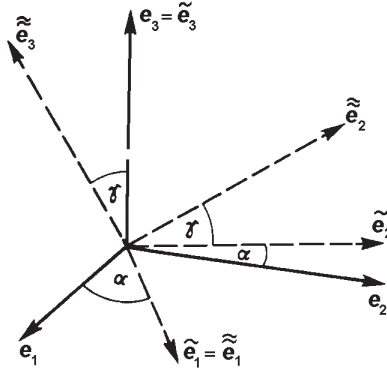
$$(2.5.11) \quad \mathbf{Q}^{-1} = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix};$$

tehát $\mathbf{Q}^{-1} = \mathbf{Q}^T$, azaz \mathbf{Q} ortogonális mátrix (lásd az 1.4.10 definíciót).

* * *

18. Példa. Tekintsük most a geometriai tér vektorainak lineáris terét és bázisként válasszuk egy térbeli derékszögű koordináta-rendszer egymásra merőleges $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ egységvektorait. Most transzformáljuk a koordináta-rendszert a következőképpen: először forgassuk el a koordináta-rendszert az \mathbf{e}_3 tengely körül α szöggel, az így kapott bázisvektorok legyenek $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3$; ezután pedig forgassuk el ezt a koordináta-rendszert $\tilde{\mathbf{e}}_1$ körül γ szöggel (lásd

az ábrát). Írjuk fel a transzformáció mátrixát!



Megoldás. Az első lépésben az \mathbf{e}_3 bázisvektor a forgatás során nem változott, az \mathbf{e}_1 és \mathbf{e}_2 bázisvektorokat pedig pozitív irányban α szöggel elforgattuk az $[\mathbf{e}_1, \mathbf{e}_2]$ síkban. A (2.5.10) alatti mátrix elemeinek a felhasználásával tehát

$$(2.5.12) \quad [\tilde{\mathbf{e}}_1 \quad \tilde{\mathbf{e}}_2 \quad \tilde{\mathbf{e}}_3] = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_3] \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Jelölje $\mathbf{Q}^{(3)}$ ennek a transzformációnak a mátrixát:

$$(2.5.13) \quad \mathbf{Q}^{(3)} = \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

A következő lépésben az $\tilde{\mathbf{e}}_1$ bázisvektor nem változik a forgatás során, és az $\tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3$ bázisvektorokat forgatjuk el pozitív irányban γ szöggel az $\tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3$ vektorok síkjában. A transzformáció az alábbi alakban írható:

$$(2.5.14) \quad [\tilde{\mathbf{e}}_1 \quad \tilde{\mathbf{e}}_2 \quad \tilde{\mathbf{e}}_3] = [\tilde{\mathbf{e}}_1 \quad \tilde{\mathbf{e}}_2 \quad \tilde{\mathbf{e}}_3] \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \gamma & -\sin \gamma \\ 0 & \sin \gamma & \cos \gamma \end{bmatrix}.$$

Jelölje $\mathbf{Q}^{(1)}$ ennek a transzformációnak a mátrixát:

$$(2.5.15) \quad \mathbf{Q}^{(1)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \gamma & -\sin \gamma \\ 0 & \sin \gamma & \cos \gamma \end{bmatrix}.$$

Ha most a (2.5.12) összefüggést behelyettesítjük a (2.5.14) jobb oldalába, akkor megkapjuk az eredeti bázisvektorok és a két elforgatás után adódó bázisvektorok közötti összefüggést:

$$\begin{bmatrix} \tilde{\mathbf{e}}_1 & \tilde{\mathbf{e}}_2 & \tilde{\mathbf{e}}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \end{bmatrix} \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \gamma & -\sin \gamma \\ 0 & \sin \gamma & \cos \gamma \end{bmatrix}.$$

Innen kiolvasható, hogy a koordináta-rendszer kétszeri elforgatása után nyert mátrix az egyes elforgatások mátrixának a szorzata. Jelölje \mathbf{Q} a keresett mátrixot, ekkor

$$(2.5.16) \quad \mathbf{Q} = \mathbf{Q}^{(3)} \mathbf{Q}^{(1)}.$$

A 1.4.2 tétel értelmében ez is ortogonális mátrix.

* * *

Megjegyzés. Rövid számolás után meggyőződhetünk arról, hogy

$$\mathbf{Q}^{(1)} \mathbf{Q}^{(3)} \neq \mathbf{Q}^{(3)} \mathbf{Q}^{(1)},$$

azaz a fenti két bázistranszformáció sorrendje nem felcserélhető!

2.5.2 Bilineáris alak mátrixának transzformációja új bázisra való áttérés esetén (kongruens transzformáció)

Az előzőekben láttuk, hogyan lehet új bázisra való áttérés esetén vektorok új koordinátáit a régi koordináták segítségével kifejezni. Most megmutatjuk, hogyan alkalmazható ez arra a feladatra, hogy az új bázisban meghatározzuk bilineáris alakok mátrixát, az eredeti bázisra vonatkozó mátrixuk segítségével.

Tekintsük az \mathbf{R}_n komplex lineáris térben értelmezett $A(\mathbf{x}; \mathbf{y})$ bilineáris alakot és legyen $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ az n -dimenziós tér egy bázisa. Az \mathbf{x} vektornak a bázisra vonatkozó koordinátái legyenek $\xi_1, \xi_2, \dots, \xi_n$, az \mathbf{y} vektoré pedig $\eta_1, \eta_2, \dots, \eta_n$. A bilineáris alak 2.3.2 definíciója értelmében

$$(2.5.17) \quad A(\mathbf{x}; \mathbf{y}) = A \left(\sum_{j=1}^n \xi_j \mathbf{e}_j; \sum_{i=1}^n \eta_i \mathbf{e}_i \right) = \sum_{j=1}^n \sum_{i=1}^n \bar{\eta}_i \xi_j A(\mathbf{e}_j; \mathbf{e}_i);$$

az $a_{ij} = A(\mathbf{e}_j; \mathbf{e}_i)$ jelöléssel a bilineáris alak:

$$(2.5.18) \quad A(\mathbf{x}; \mathbf{y}) = \begin{bmatrix} \bar{\eta}_1 & \bar{\eta}_2 & \dots & \bar{\eta}_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} = \mathbf{y}^H \mathbf{A} \mathbf{x}.$$

Az itt fellépő \mathbf{A} mátrix az $A(\mathbf{x}; \mathbf{y})$ bilineáris alak $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisra vonatkozó mátrixa, \mathbf{x} és \mathbf{y} pedig a ξ_j , ill. η_i koordinátákból alkotott koordinátavektor. Ha most a (2.5.2) összefüggés szerint egy új, $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \dots, \tilde{\mathbf{e}}_n$ bázisra térünk át, akkor a bilineáris alak mátrixának transzformációja a következőképpen nyerhető. Az \mathbf{x} és \mathbf{y} vektoroknak az új bázisra vonatkozó koordinátái legyenek $\tilde{\xi}_1, \tilde{\xi}_2, \dots, \tilde{\xi}_n$ és $\tilde{\eta}_1, \tilde{\eta}_2, \dots, \tilde{\eta}_n$, ezek kapcsolata az eredeti bázisra vonatkozó koordinátákkal a (2.5.5) összefüggés szerint

$$\begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{bmatrix} \begin{bmatrix} \tilde{\xi}_1 \\ \tilde{\xi}_2 \\ \vdots \\ \tilde{\xi}_n \end{bmatrix} = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}; \quad \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{bmatrix} \begin{bmatrix} \tilde{\eta}_1 \\ \tilde{\eta}_2 \\ \vdots \\ \tilde{\eta}_n \end{bmatrix} = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix};$$

azaz

$$(2.5.19) \quad \mathbf{C}\tilde{\mathbf{x}} = \mathbf{x}; \quad \mathbf{C}\tilde{\mathbf{y}} = \mathbf{y}.$$

Behelyettesítve ezeket az összefüggéseket a bilineáris alak (2.5.18) alatti kifejezésébe,

$$(2.5.20) \quad A(\mathbf{x}; \mathbf{y}) = \tilde{\mathbf{y}}^H \mathbf{C}^H \mathbf{A} \mathbf{C} \tilde{\mathbf{x}}$$

adódik. Jelölje $\tilde{\mathbf{A}} = [\tilde{a}_{ij}]$ az $A(\mathbf{x}; \mathbf{y})$ bilineáris alaknak az új, $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \dots, \tilde{\mathbf{e}}_n$ bázisra vonatkozó mátrixát:

$$(2.5.21) \quad A(\mathbf{x}; \mathbf{y}) = \tilde{\mathbf{y}}^H \tilde{\mathbf{A}} \tilde{\mathbf{x}}.$$

A (2.5.20) és a (2.5.21) kifejezések egybevetéséből látható, hogy $\tilde{\mathbf{A}} = \mathbf{C}^H \mathbf{A} \mathbf{C}$. Az eredmény a következő tételben fogalmazható meg.

2.5.2 tétel. Ha az \mathbf{R}_n komplex lineáris térben értelmezett $A(\mathbf{x}; \mathbf{y})$ bilineáris alak mátrixa az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisban \mathbf{A} , az $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \dots, \tilde{\mathbf{e}}_n$ bázisban $\tilde{\mathbf{A}}$, az új bázisra való áttérést pedig az

$$\tilde{\mathbf{e}}_k = \sum_{i=1}^n \mathbf{e}_i c_{ik}$$

összefüggéssel definiált $\mathbf{C} = [c_{ik}]$ mátrix közvetíti, akkor az \mathbf{A} , $\tilde{\mathbf{A}}$ és \mathbf{C} mátrixok között fennáll az

$$(2.5.22) \quad \tilde{\mathbf{A}} = \mathbf{C}^H \mathbf{A} \mathbf{C}$$

összefüggés.

2.5.1 definíció. Ha az \mathbf{A} mátrixot egy nemszinguláris \mathbf{C} mátrixszal jobbról, \mathbf{C} transzponált konjugáltjával pedig balról szorozzuk, akkor azt mondjuk, hogy az \mathbf{A} mátrixot **kongruens transzformációnak** vetjük alá. A kapott $\tilde{\mathbf{A}}$ mátrixot és az eredeti \mathbf{A} mátrixot egymással **kongruenseknek** nevezzük.

2.5.3 Az x és Ax vektorok koordinátái közötti összefüggés

Most megmutatjuk, hogy ha A egy lineáris tér adott transzformációja, akkor hogyan határozhatók meg az Ax vektor (adott bázisra vonatkozó) koordinátái az x vektor koordinátáinak segítségével. Legyen e_1, e_2, \dots, e_n az n -dimenziós tér egy bázisa; a bázisvektorokat az A transzformáció az

$$(2.5.23) \quad Ae_k = \sum_{j=1}^n e_j a_{jk} \quad (k = 1, \dots, n)$$

vektorokba viszi, ahol $A = [a_{jk}]$ most is a transzformációnak az e_1, e_2, \dots, e_n bázisra vonatkozó mátrixát jelöli. A (2.5.23) szemléletesen az

$$(2.5.24) \quad A[e_1 \ e_2 \ \dots \ e_n] = [Ae_1 \ Ae_2 \ \dots \ Ae_n] = [e_1 \ e_2 \ \dots \ e_n]A$$

alakban írható. Az x és az $y = Ax$ vektor előállítására az e_1, e_2, \dots, e_n bázisban legyen

$$x = \sum_{j=1}^n e_j \xi_j, \quad y = Ax = \sum_{j=1}^n e_j \eta_j,$$

ami az

$$(2.5.25) \quad x = [e_1 \ e_2 \ \dots \ e_n] \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix},$$

$$(2.5.26) \quad y = Ax = [e_1 \ e_2 \ \dots \ e_n] \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix}$$

alakban írható. Alkalmazzuk az A lineáris transzformációt a (2.5.25) alatti x vektorra, és vegyük figyelembe a (2.5.24) összefüggést. Ekkor

$$y = Ax = [e_1 \ e_2 \ \dots \ e_n] A \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}.$$

Összevetve ezt a (2.5.26) kifejezéssel, az \mathbf{x} és \mathbf{Ax} vektorok koordinátái között a következő összefüggést kapjuk:

$$(2.5.27) \quad \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}.$$

Érvényes tehát a következő tétel.

2.5.3 tétel. *Ha az \mathbf{A} lineáris transzformáció mátrixa az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisban \mathbf{A} , akkor az \mathbf{x} vektor $\mathbf{y} = \mathbf{Ax}$ képének az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisra vonatkozó koordinátáit az \mathbf{x} vektor ugyanezen bázisra vonatkozó koordinátáiból úgy kapjuk meg, hogy a transzformáció \mathbf{A} mátrixával szorozzuk az \mathbf{x} vektor koordinátáiból alkotott oszlopvektort.*

Megjegyzés. A (2.5.24) és (2.5.27) összefüggések egybevetésével megállapítható, hogy a lineáris tér \mathbf{A} transzformációjának a bázisvektorai a transzformáció (ezen bázisra vonatkozó) mátrixának az oszlopai szerint, míg egy tetszőleges vektor (ezen bázisra vonatkozó) koordinátái a mátrix sorai szerint transzformálódnak.

2.5.4 A lineáris transzformáció mátrixának transzformációja új bázisra való áttérés esetén (hasonlósági transzformáció)

Jelölje $\mathbf{A} = [a_{ij}]$, illetve $\tilde{\mathbf{A}} = [\tilde{a}_{ij}]$ az \mathbf{A} lineáris transzformáció mátrixát az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$, illetve az $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \dots, \tilde{\mathbf{e}}_n$ bázisra vonatkozóan, azaz

$$(2.5.28) \quad \mathbf{A}\mathbf{e}_k = \sum_{j=1}^n \mathbf{e}_j a_{jk} \quad (k = 1, \dots, n),$$

illetve

$$(2.5.29) \quad \mathbf{A}\tilde{\mathbf{e}}_k = \sum_{j=1}^n \tilde{\mathbf{e}}_j \tilde{a}_{jk} \quad (k = 1, \dots, n).$$

A (2.5.28) és (2.5.29) összefüggések szemléletesen és tömören az

$$(2.5.30) \quad \mathbf{A}[\mathbf{e}_1 \mathbf{e}_2 \dots \mathbf{e}_n] = [\mathbf{A}\mathbf{e}_1 \dots \mathbf{A}\mathbf{e}_n] = [\mathbf{e}_1 \dots \mathbf{e}_n] \mathbf{A},$$

illetve az

$$(2.5.31) \quad \mathbf{A}[\tilde{\mathbf{e}}_1 \tilde{\mathbf{e}}_2 \dots \tilde{\mathbf{e}}_n] = [\mathbf{A}\tilde{\mathbf{e}}_1 \dots \mathbf{A}\tilde{\mathbf{e}}_n] = [\tilde{\mathbf{e}}_1 \dots \tilde{\mathbf{e}}_n] \tilde{\mathbf{A}}$$

alakban írhatók. Jelölje $\mathbf{C} = [c_{ij}]$ a két bázis közötti kapcsolatot közvetítő mátrixot:

$$(2.5.32) \quad [\tilde{\mathbf{e}}_1 \tilde{\mathbf{e}}_2 \dots \tilde{\mathbf{e}}_n] = [\mathbf{e}_1 \mathbf{e}_2 \dots \mathbf{e}_n] \mathbf{C},$$

és alkalmazzuk az A lineáris transzformációt (2.5.32) mindkét oldalán. A bal oldalon (2.5.31) és (2.5.32) figyelembevételével

$$(2.5.33) \quad A[\tilde{e}_1 \tilde{e}_2 \dots \tilde{e}_n] = [\tilde{e}_1 \tilde{e}_2 \dots \tilde{e}_n] \tilde{A} = [e_1 e_2 \dots e_n] C \tilde{A},$$

míg a jobb oldalon (2.5.30) figyelembevételével

$$(2.5.34) \quad A[e_1 e_2 \dots e_n] C = [e_1 e_2 \dots e_n] AC.$$

Végül, (2.5.33) és (2.5.34) egyenlőségéből $C \tilde{A} = AC$ következik.

Az eredményt a következő tételben fogalmazzuk meg.

2.5.4 a) tétel. *Ha az R_n tér A lineáris transzformációjának a mátrixa az e_1, e_2, \dots, e_n bázisban A , az $\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_n$ bázisban \tilde{A} , és a két bázis közötti kapcsolatot az*

$$\tilde{e}_k = \sum_{i=1}^n e_i c_{ik}$$

összefüggéssel definiált $C = [c_{ik}]$ mátrix közvetíti, akkor az A , \tilde{A} és C mátrix között fennáll az

$$(2.5.35) \quad \tilde{A} = C^{-1}AC$$

összefüggés.

2.5.2 definíció. *Ha egy A mátrixot egy nemszinguláris C mátrixszal jobbról, ennek inverzével pedig balról szorzunk, akkor azt mondjuk, hogy az A mátrixon **hasonlósági transzformációt** végzünk, az így kapott \tilde{A} mátrixot és az adott A mátrixot pedig **hasonló mátrixoknak** nevezzük.*

A 2.5.4 a) tétel jelentősége abban áll, hogy a (2.3.35) képlettel megadja azt az általános érvényű összefüggést, amely az R_n tér egy A lineáris transzformációjának két, tetszőlegesen felvett bázishoz tartozó mátrixa között fennáll. Mivel pedig a bázis megválasztása – az analitikus geometria kifejezőmódjával – a koordináta-rendszer megválasztását jelenti, a 2.5.4 a) tétel a következőképpen is fogalmazható.

2.5.4 b) tétel. *Bármely lineáris transzformációnak adott koordináta-rendszerhez tartozó mátrixa a koordináta-rendszer transzformációja esetén hasonlósági transzformáción megy át.*

Ez más szóval azt jelenti, hogy adott lineáris transzformáció mátrixa hasonlósági transzformáció erejéig egyértelműen meghatározott. A következő szakaszokban majd éppen az fogja vizsgálataink központi feladatát képezni, miként célszerű a koordináta-rendszert úgy felvenni, hogy az adott lineáris

transzformáció mátrixa lehetőleg egyszerű legyen, és a transzformáció jellemző tulajdonságait tükrözze. Azt is látni fogjuk, hogy a kapott eredmények a lineáris transzformációk osztályozását is lehetővé teszik.

Összehasonlítva a 2.5.4 a), ill. 2.5.2 tételben is szereplő (2.5.35), ill. a (2.5.22) összefüggéseket, látható, hogy az új bázisra való áttérés során általában másként transzformálódik a lineáris transzformációk mátrixa és a bilineáris alakok mátrixa. E tételek az n -dimenziós komplex lineáris térre vonatkoztak, nem volt szükségünk sem vektorok skaláris szorzatára, sem az ortogonalitás fogalmára. Abban az esetben, ha euklideszi térben értelmezett bilineáris alakot, ill. lineáris transzformációt tekintünk, ahol definiálható a vektorok ortogonalitása, célszerű ortonormált bázisokat felvenni. A következő szakaszban be fogjuk bizonyítani, hogy az n -dimenziós \mathbf{E}_n komplex euklideszi (unitér) tér két ortonormált bázisa közötti kapcsolatot mindig unitér mátrix közvetíti, amelyet a

$$(2.5.36) \quad \mathbf{C}^{-1} = \mathbf{C}^H$$

összefüggéssel definiáltunk. Ez azt jelenti, hogy ebben az esetben a (2.5.35) és a (2.5.22) képletek ugyanazt a transzformációt jelentik, vagyis a hasonlósági transzformáció megegyezik a kongruens transzformációval.

2.5.3 definíció. *Ha az \mathbf{A} mátrixot jobbról egy unitér mátrixszal, balról pedig annak inverzével (vagyis transzponált konjugáltjával) szorozzuk, akkor azt mondjuk, hogy az \mathbf{A} mátrixot **unitér transzformációnak** vetjük alá.*

A 2.8.2 pontban bebizonyítjuk majd a következő tételt (lásd a 2.8.5 tételt):

2.5.5 tétel. *Ha az n -dimenziós \mathbf{E}_n unitér tér egy ortonormált bázisáról egy másik ortonormált bázisra való áttérést a \mathbf{C} unitér mátrix közvetíti, akkor az \mathbf{E}_n tér \mathbf{A} lineáris transzformációjának és $A(\mathbf{x}; \mathbf{y})$ bilineáris alakjának az adott bázisra vonatkozó mátrixa ugyanazzal az unitér transzformációval vihető át az új bázisra vonatkozó mátrixba.*

Valós euklideszi tér esetén ennek a transzformációnak a mátrixa speciálisan ortogonális mátrix és ekkor a megfelelő transzformációt is ortogonálisnak nevezzük. Néhány egyszerű lineáris transzformáció ortonormált bázisra vonatkozó mátrixának meghatározására a következő néhány feladatban a koordináta-rendszer transzformációjával kapcsolatos tételeket a geometriai térben – tehát háromdimenziós valós euklideszi térben – alkalmazzuk. Ez az oka annak, hogy a részletes tárgyalást megelőzve az unitér transzformáció fogalmát és említett tulajdonságait előrebocsátottuk. E mátrixok meghatározásának alapgondolata a következő. Ha a szóban forgó transzformáció mátrixa egy speciálisan felvett koordináta-rendszerben ismert vagy könnyen meghatározható, akkor a feladatot úgy oldjuk meg, hogy az adott ortonormált koordináta-rendszert először forgatásokkal – azaz ortogonális transz-

formációkkal – ebbe a lineáris transzformáció szempontjából kedvező speciális helyzetbe hozzuk; ezután felírjuk a lineáris transzformáció mátrixát a speciális helyzetű koordináta-rendszerben, majd a koordináta-rendszert vizszostranszformálva, megkapjuk a lineáris transzformáció keresett mátrixát az adott ortonormált bázisban.

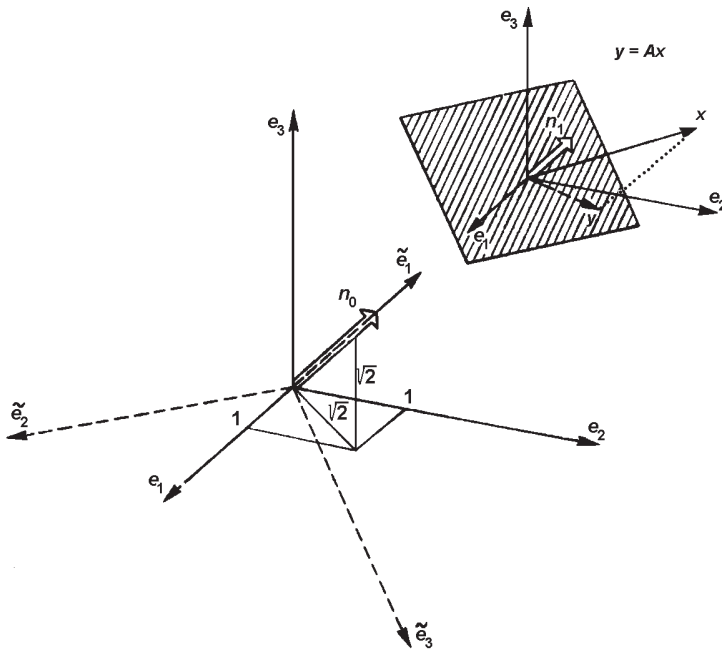
Tekintsük tehát a geometriai teret, és bázisként válasszuk egy derékszögű koordináta-rendszer tengelyei irányába mutató \mathbf{e}_1 , \mathbf{e}_2 , \mathbf{e}_3 egységvektorokat. Ha valamelyik koordinátatengelyre vagy koordinátasíkra vetítjük, ill. tükrözzük, vagy valamelyik koordinátatengely körül elforgatjuk az adott lineáris tér vektorait, akkor az ennek megfelelő lineáris transzformáció mátrixa igen egyszerűen felírható. Ezeket a mátrixokat táblázatban foglaltuk össze.

	Vetítés	Tükrözés	Forgatás
	mátrixa		
	V	T	F
$\mathbf{e}_1\mathbf{e}_2$ síkra	$\begin{bmatrix} 1 & & \\ & 1 & \\ & & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & & \\ & 1 & \\ & & -1 \end{bmatrix}$	\mathbf{e}_1 tengely körül $\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \vartheta & -\sin \vartheta \\ 0 & \sin \vartheta & \cos \vartheta \end{bmatrix}$
$\mathbf{e}_1\mathbf{e}_3$ síkra	$\begin{bmatrix} 1 & & \\ & 0 & \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & & \\ & -1 & \\ & & 1 \end{bmatrix}$	
$\mathbf{e}_2\mathbf{e}_3$ síkra	$\begin{bmatrix} 0 & & \\ & 1 & \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & & \\ & 1 & \\ & & 1 \end{bmatrix}$	\mathbf{e}_2 tengely körül $\begin{bmatrix} \cos \vartheta & 0 & \sin \vartheta \\ 0 & 1 & 0 \\ -\sin \vartheta & 0 & \cos \vartheta \end{bmatrix}$
\mathbf{e}_1 tengelyre	$\begin{bmatrix} 1 & & \\ & 0 & \\ & & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & & \\ & -1 & \\ & & -1 \end{bmatrix}$	
\mathbf{e}_2 tengelyre	$\begin{bmatrix} 0 & & \\ & 1 & \\ & & 0 \end{bmatrix}$	$\begin{bmatrix} -1 & & \\ & 1 & \\ & & -1 \end{bmatrix}$	\mathbf{e}_3 tengely körül $\begin{bmatrix} \cos \vartheta & -\sin \vartheta & 0 \\ \sin \vartheta & \cos \vartheta & 0 \\ 0 & 0 & 1 \end{bmatrix}$
\mathbf{e}_3 tengelyre	$\begin{bmatrix} 0 & & \\ & 0 & \\ & & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & & \\ & -1 & \\ & & 1 \end{bmatrix}$	

Ha most egy olyan lineáris transzformáció mátrixát kívánjuk felírni, amely a tér minden vektorát az origóra illeszkedő általános helyzetű síkra vagy egyenesre vetíti, ill. tükrözi, akkor először a koordináta-rendszert úgy forgatjuk el, hogy az adott általános helyzetű sík, ill. egyenes az elforgatott koordináta-

rendszer valamelyik koordinátságijával, ill. tengelyével essen egybe. A vetítésnek vagy tükrözésnek e koordináta-rendszerre vonatkozó mátrixa az előbbi táblázatból kiolvasható. Ezután már csak az elforgatott koordináta-rendszert kell *visszaforatnunk* az eredeti koordináta-rendszerbe, ami annyit jelent, hogy a táblázatban található valamelyik egyszerű mátrixon egy ortogonális transzformációt hajtunk végre. Az eljárást néhány példán bemutatjuk.

19. Példa. Írjuk fel annak a lineáris transzformációnak a mátrixát, amely a geometriai tér minden vektorát az $x + y + \sqrt{2}z = 0$ egyenletű síkba vetíti (lásd az ábrát), ahol az x, y, z koordináta-rendszert az $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ páronként merőleges egységvektorok határozzák meg.



Megoldás. A sík egyenletéből kiolvasható, hogy a normális irányú egységvektor:

$$\mathbf{n}_0 = \left[\frac{1}{2}, \frac{1}{2}, \frac{\sqrt{2}}{2} \right]^T.$$

Legyen az új bázis egyik vektora $\tilde{\mathbf{e}}_1 = \mathbf{n}_0$. Vegyünk fel a síkban egy tetszőleges egységvektort:

$$\tilde{\mathbf{e}}_2 = \left[\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 0 \right]^T.$$

Az $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2$ vektorrendszert egészítsük ki bázissá:

$$\tilde{\mathbf{e}}_3 = \tilde{\mathbf{e}}_1 \times \tilde{\mathbf{e}}_2 = \left[\frac{1}{2}, \frac{1}{2}, -\frac{\sqrt{2}}{2} \right]^T.$$

Megjegyezzük, hogy az új $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3$ bázist az $\mathbf{n}_0, \mathbf{e}_2, \mathbf{e}_3$ bázisból a Schmidt-féle ortogonalizálási eljárással is meghatározhattuk volna.

Az $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3$ bázisban a lineáris transzformáció mátrixa:

$$\tilde{\mathbf{V}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

a bázistranszformáció mátrixa:

$$\mathbf{C} = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{2}}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{\sqrt{2}}{2} & \frac{1}{2} \\ \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \end{bmatrix}.$$

Ha figyelembe vesszük, hogy a \mathbf{C} mátrix ortogonális, azaz $\mathbf{C}^T = \mathbf{C}^{-1}$, akkor a keresett \mathbf{V} mátrixot a $\tilde{\mathbf{V}}$ mátrixból a \mathbf{C} mátrixszal végzett ortogonális transzformációval kapjuk:

$$\begin{aligned} \mathbf{V} = \mathbf{C}\tilde{\mathbf{V}}\mathbf{C}^T &= \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{2}}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{\sqrt{2}}{2} & \frac{1}{2} \\ \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix} = \\ &= \begin{bmatrix} \frac{3}{4} & -\frac{1}{4} & -\frac{\sqrt{2}}{4} \\ -\frac{1}{4} & \frac{3}{4} & -\frac{\sqrt{2}}{4} \\ -\frac{\sqrt{2}}{4} & -\frac{\sqrt{2}}{4} & \frac{1}{2} \end{bmatrix}. \end{aligned}$$

* * *

Vizsgáljuk most meg, mi jellemzi a vetítő mátrixokat. Az előbbieken található legegyszerűbb \mathbf{V} vetítőmátrixok szerkezetéből azonnal látható, hogy ezek hatványozás során nem változnak, vagyis *projektorok* (megjegyezzük, hogy a projektorok elnevezését éppen az itt említett geometriai tulajdonságuk indokolja): $\mathbf{V}^2 = \mathbf{V}$.

Be kell még látnunk, hogy ez a tulajdonság a koordináta-rendszer transzformációjával szemben invariáns:

$$(\tilde{\mathbf{V}})^2 = \mathbf{C}^{-1}\mathbf{V}\mathbf{C}\mathbf{C}^{-1}\mathbf{V}\mathbf{C} = \mathbf{C}^{-1}\mathbf{V}^2\mathbf{C} = \mathbf{C}^{-1}\mathbf{V}\mathbf{C} = \tilde{\mathbf{V}}.$$

tehát ha V projektor volt, akkor \tilde{V} is az.

20. Példa. Írjuk fel most annak a transzformációnak a mátrixát, amely a geometriai tér vektorait az \mathbf{e}_2 tengely körül forgatja el β szöggel pozitív irányban.

Megoldás. A (2.5.10) mátrix segítségével az $\mathbf{e}_3, \mathbf{e}_1$ vektorok síkjában a következő transzformációt végezzük (lásd az ábrát):

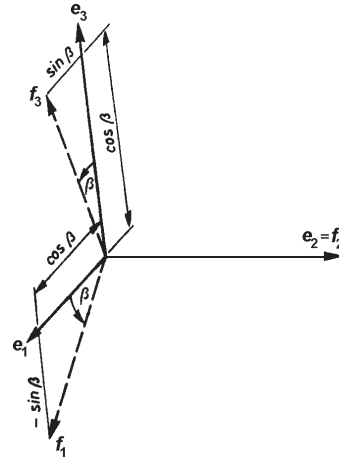
Jelölje az $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ képét $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$.

$$(2.5.37) \quad [\mathbf{f}_3 \quad \mathbf{f}_1] = [\mathbf{e}_3 \quad \mathbf{e}_1] \begin{bmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{bmatrix},$$

azaz

$$\mathbf{f}_3 = \mathbf{e}_3 \cos \beta + \mathbf{e}_1 \sin \beta,$$

$$\mathbf{f}_1 = -\mathbf{e}_3 \sin \beta + \mathbf{e}_1 \cos \beta.$$



Mivel az \mathbf{e}_2 bázisvektor a forgatás során helyben marad, azaz $\mathbf{f}_2 = \mathbf{e}_2$, a kapott összefüggéseket a szokásos sorrendbe írva,

$$(2.5.38) \quad [\mathbf{f}_1 \quad \mathbf{f}_2 \quad \mathbf{f}_3] = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_3] \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix}.$$

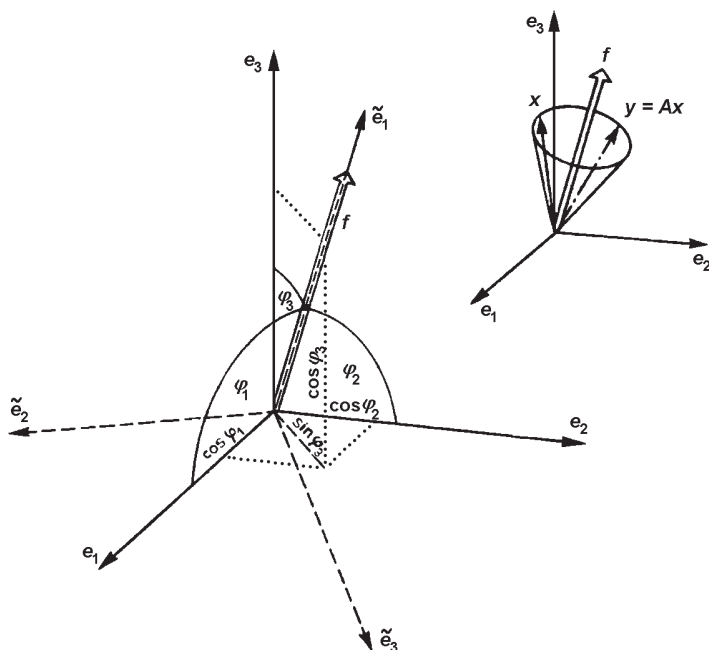
Jelölje $\mathbf{F}^{(2)}$ az \mathbf{e}_2 bázisvektor körüli forgatás mátrixát; ekkor

$$(2.5.39) \quad \mathbf{F}^{(2)} = \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix}.$$

(Ügyeljünk a negatív előjel helyére!)

* * *

21. Példa. Határozzuk meg annak a transzformációnak a mátrixát, amely a geometriai tér valamennyi vektorát az $\mathbf{f} = (\cos \varphi_1, \cos \varphi_2, \cos \varphi_3)$ egységvektor körül θ szöggel elforgatja.



Megoldás. Az 1. példa megoldásához hasonlóan, vegyünk fel egy új, ortonormált bázist, és ebben írjuk fel először a lineáris transzformáció mátrixát. Az új bázis vektorai:

$$\tilde{\mathbf{e}}_1 = \mathbf{f},$$

$$\tilde{\mathbf{e}}_2 = \left[\frac{\cos \varphi_2}{\sin \varphi_3}, -\frac{\cos \varphi_1}{\sin \varphi_3}, 0 \right]^T,$$

$$\tilde{\mathbf{e}}_3 = \left[\frac{\cos \varphi_1 \cos \varphi_3}{\sin \varphi_3}, \frac{\cos \varphi_2 \cos \varphi_3}{\sin \varphi_3}, -\sin \varphi_3 \right]^T.$$

A bázistranszformáció mátrixa:

$$\mathbf{C} = \begin{bmatrix} \cos \varphi_1 & \frac{\cos \varphi_2}{\sin \varphi_3} & \frac{\cos \varphi_1 \cos \varphi_3}{\sin \varphi_3} \\ \cos \varphi_2 & -\frac{\cos \varphi_1}{\sin \varphi_3} & \frac{\cos \varphi_2 \cos \varphi_3}{\sin \varphi_3} \\ \cos \varphi_3 & 0 & -\sin \varphi_3 \end{bmatrix}.$$

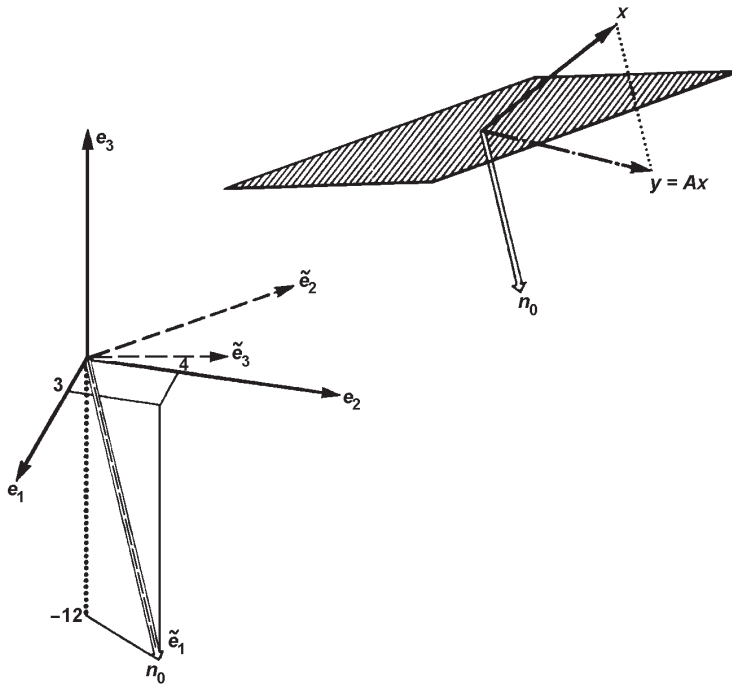
A lineáris transzformáció mátrixa az $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3$ bázisban:

$$\tilde{\mathbf{F}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix}.$$

Figyelembe véve, hogy a \mathbf{C} mátrix ortogonális, a keresett \mathbf{F} mátrix az $\mathbf{F} = \mathbf{C}\tilde{\mathbf{F}}\mathbf{C}^T$ összefüggés alapján számítható.

* * *

22. Példa. Határozzuk meg annak a transzformációnak a mátrixát, amely a geometriai tér vektorait a $3x + 4y - 12z = 0$ egyenletű síkra tükrözi.



Megoldás. A sík normális irányú egységvektora

$$\mathbf{n}_0 = \left[\frac{3}{13}, \frac{4}{13}, -\frac{12}{13} \right]^T.$$

Válasszuk meg az új koordináta-rendszert a következőképpen. Az $\tilde{\mathbf{e}}_1$ vektor essen egybe az \mathbf{n}_0 vektorral. Az $\tilde{\mathbf{e}}_2$ vektor legyen a sík egy tetszőleges egységvektora, például

$$\tilde{\mathbf{e}}_2 = \left[-\frac{4}{5}, \frac{3}{5}, 0 \right]^T.$$

A harmadik bázisvektor:

$$\tilde{\mathbf{e}}_3 = \tilde{\mathbf{e}}_1 \times \tilde{\mathbf{e}}_2 = \left[\frac{36}{65}, \frac{48}{65}, \frac{5}{13} \right]^T,$$

így az $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3$ vektorok jobbsodrású koordináta-rendszert alkotnak.

A koordinátatranszformáció \mathbf{C} mátrixa

$$(2.5.40) \quad \mathbf{C} = \begin{bmatrix} \frac{3}{13} & -\frac{4}{5} & \frac{36}{65} \\ \frac{4}{13} & \frac{3}{5} & \frac{48}{65} \\ -\frac{12}{13} & 0 & \frac{5}{13} \end{bmatrix}.$$

Jelölje \mathbf{T} az adott síkra tükrözés mátrixát az eredeti $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ bázisban, $\tilde{\mathbf{T}}$ pedig a transzformált $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3$ bázisban. Mivel a sík normálvektora az $\tilde{\mathbf{e}}_1$ irányba mutat, $\tilde{\mathbf{T}}$ az $\tilde{\mathbf{e}}_2 \tilde{\mathbf{e}}_3$ koordinátasíkra végzett tükrözés mátrixa:

$$(2.5.41) \quad \mathbf{T} = \begin{bmatrix} -1 & & \\ & 1 & \\ & & 1 \end{bmatrix}.$$

Az eredeti és a transzformált bázisra vonatkozó mátrixok közötti (2.5.35) összefüggés szerint

$$(2.5.42) \quad \tilde{\mathbf{T}} = \mathbf{C}^{-1} \mathbf{T} \mathbf{C},$$

ahol \mathbf{C} a koordináta-rendszer elforgatását jellemző (2.5.40) alatti ortogonális mátrix: $\mathbf{C}^{-1} = \mathbf{C}^T$. A (2.5.42) összefüggésből tehát meghatározható az adott síkra végrehajtott tükrözés keresett \mathbf{T} mátrixa:

$$\mathbf{T} = \mathbf{C} \tilde{\mathbf{T}} \mathbf{C}^T.$$

Behelyettesítve \mathbf{C} és $\tilde{\mathbf{T}}$ helyére a (2.5.40), ill. (2.5.41) mátrixokat, a beszorzás elvégzése után megkapjuk a keresett tükrözőmátrixot:

$$\mathbf{T} = \frac{1}{169} \begin{bmatrix} 151 & 145 & 241 \\ 145 & 137 & 65 \\ 241 & 65 & -119 \end{bmatrix}.$$

* * *

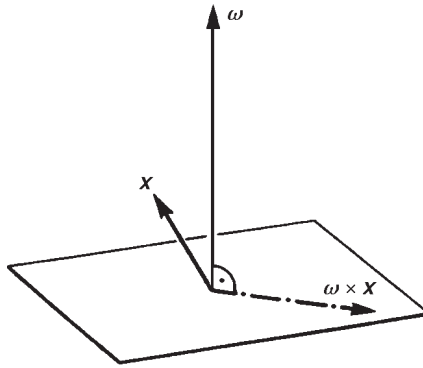
Vizsgáljuk most meg, milyen tulajdonságaik vannak a tükrözőmátrixoknak. Azok a mátrixok, amelyek csupán koordinátasíkokra vagy koordináta-tengelyekre végzett tükrözésnek felelnek meg, igen egyszerű szerkezetűek. Nem nehéz észrevenni, hogy mindegyiknek a négyzete az egységmátrixot adja, ezek tehát *involutórius* mátrixok. Megmutatjuk most, hogy ez a tulajdonság

a koordináta-rendszer transzformációjával szemben invariáns. Ha ugyanis a $\tilde{\mathbf{T}} = \mathbf{C}^{-1}\mathbf{T}\mathbf{C}$ mátrixot négyzetre emeljük, akkor a következő adódik:

$$\tilde{\mathbf{T}}^2 = \mathbf{C}^{-1}\mathbf{T}\mathbf{C}\mathbf{C}^{-1}\mathbf{T}\mathbf{C} = \mathbf{C}^{-1}\mathbf{T}^2\mathbf{C},$$

ami azt jelenti, hogy $\mathbf{T}^2 = \mathbf{E}$ esetén $\tilde{\mathbf{T}}^2 = \mathbf{E}$; vagyis minden tükrözőmátrix involutórius.

23. Példa. Határozzuk meg annak a transzformációnak a mátrixát, amely a geometriai tér minden \mathbf{x} vektorához az $\boldsymbol{\omega} \times \mathbf{x}$ vektoriális szorzatot rendeli hozzá, ahol $\boldsymbol{\omega}$ a tér egy rögzített vektora.



Megoldás. Bázisként válasszunk egy derékszögű koordináta-rendszer tengelyei irányába mutató $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ egységvektorokat; az $\boldsymbol{\omega}$ vektor koordinátái legyenek $\omega_1, \omega_2, \omega_3$, tehát

$$\boldsymbol{\omega} = \omega_1\mathbf{e}_1 + \omega_2\mathbf{e}_2 + \omega_3\mathbf{e}_3.$$

Keressük azt az \mathbf{A} lineáris transzformációt, amelyre

$$(2.5.43) \quad \mathbf{A}\mathbf{x} = \boldsymbol{\omega} \times \mathbf{x}.$$

A transzformáció mátrixának egyes oszlopait megkapjuk, ha meghatározzuk az $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ bázisvektorok $\mathbf{A}\mathbf{e}_1, \mathbf{A}\mathbf{e}_2, \mathbf{A}\mathbf{e}_3$ képének az $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ bázisra vonatkozó koordinátáit. Ki kell tehát számítani az

$$\mathbf{A}\mathbf{e}_1 = \boldsymbol{\omega} \times \mathbf{e}_1, \quad \mathbf{A}\mathbf{e}_2 = \boldsymbol{\omega} \times \mathbf{e}_2, \quad \mathbf{A}\mathbf{e}_3 = \boldsymbol{\omega} \times \mathbf{e}_3$$

vektoriális szorzatokat. Ezek rendre a következők:

$$\begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \omega_1 & \omega_2 & \omega_3 \\ 1 & 0 & 0 \end{bmatrix} = \begin{matrix} \omega_3\mathbf{e}_2 - \omega_2\mathbf{e}_3, \\ \omega_1\mathbf{e}_3 - \omega_3\mathbf{e}_1, \\ \omega_2\mathbf{e}_1 - \omega_1\mathbf{e}_2 \end{matrix}$$

$$\begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \omega_1 & \omega_2 & \omega_3 \\ 0 & 1 & 0 \end{bmatrix} = -\omega_3 \mathbf{e}_1 + \omega_1 \mathbf{e}_3,$$

$$\begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \omega_1 & \omega_2 & \omega_3 \\ 0 & 0 & 1 \end{bmatrix} = \omega_2 \mathbf{e}_1 - \omega_1 \mathbf{e}_2.$$

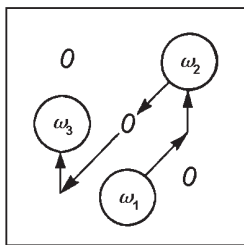
A transzformáció mátrixa tehát

$$(2.5.44) \quad \mathbf{A} = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix},$$

ami ferdén szimmetrikus mátrix.

* * *

Mivel a transzformációt az adott bázisban a mátrixa egyértelműen meghatározza, és mivel a fenti példában az $\boldsymbol{\omega}$ vektor koordinátáira semmilyen kikötést nem tettünk, ezért a példából következik, hogy minden harmadrendű (valós) ferdén szimmetrikus mátrix egy olyan lineáris transzformációt határoz meg, amely a tér minden \mathbf{x} vektorához egy rögzített vektornak az \mathbf{x} vektorral alkotott vektoriális szorzatát rendeli hozzá. A rögzített vektor koordinátái a ferdén szimmetrikus mátrix elemeiből az alábbi séma szerint olvashatók ki:



Megjegyzés. A vektoriális szorzat definíciójából következik, hogy adott $\boldsymbol{\omega}$ vektorral való vektoriális szorzás a tér valamennyi vektorát az $\boldsymbol{\omega}$ vektorra merőleges síkra képezi le. Az $\boldsymbol{\omega}$ vektorral végzett vektoriális szorzat olyan elfajuló (nem invertálható) lineáris transzformáció, melynek mátrixa bármely ortonormált bázisban ferdén szimmetrikus mátrix, amelyből az $\boldsymbol{\omega}$ vektor adott bázisra vonatkozó koordinátái egyértelműen kiolvashatók. Tehát minden harmadrendű, ferdén szimmetrikus mátrix jellemezhető egy invariáns vektorral, amely egyben annak a síknak a normálvektora, amely síkba a transzformáció a geometriai tér összes vektorait leképezi.

2.6 LINEÁRIS TRANSZFORMÁCIÓ SAJÁTVEKTORAI ÉS SAJÁTÉRTÉKEI

Ebben a szakaszban bevezetjük a lineáris algebra legfontosabb feladattípusának, az ún. *sajátérték-feladatnak* az alapfogalmait. A műszaki és a természettudományok számos elméleti és gyakorlati problémájának matematikai megfogalmazása vezet sajátérték-feladatra. E problémák egy része közvetlenül mint lineáris transzformációs probléma fogalmazható meg (pl. a szilárdságtanban a főfeszültségek és az ezeknek megfelelő főirányok meghatározása); egy másik jelentős része állandó együtthatós lineáris differenciálegyenlet-rendszerekre vezet – amelyekről a 3.6 szakaszban lesz szó – és amelyek megoldása a sajátérték-feladat alkalmazásának talán legfontosabb területe (pl. többszabadságfokú rendszerek kis rezgéseinek vizsgálata).

Mindenekelőtt be kell vezetnünk a lineáris transzformáció *invariáns altereinek* fogalmát.

2.6.1 definíció. Az \mathbf{R} lineáris tér $\mathbf{R}^{(1)}$ lineáris alterét a tér \mathbf{A} transzformációjára nézve **invariáns altérnek** nevezzük, ha $\mathbf{R}^{(1)}$ minden \mathbf{x} vektorára \mathbf{Ax} is benne van az $\mathbf{R}^{(1)}$ altérben.

Tekintsük példaként a geometriai tér egy olyan transzformációját, amely a tér vektorait egy ponton áthaladó tengely körül elforgatja. Ennek a transzformációnak invariáns altere pl. egyrészt a forgástengely, másrészt a forgástengelyre merőleges és az adott ponton áthaladó sík. Az előbbi egydimenziós invariáns altér, az utóbbi pedig kétdimenziós invariáns altér.

Tekintsük másik példaként a legfeljebb $(n - 1)$ -ed fokú polinomok terében a differenciálást mint lineáris transzformációt. Ekkor a legfeljebb k -ad fokú $(k \leq n - 1)$ polinomok halmaza invariáns alteret alkot.

A lineáris transzformációk elméletében különleges szerepük van az *egydimenziós* invariáns altereknek. Legyen $\mathbf{R}^{(1)}$ a zérustól különböző \mathbf{x} vektor segítségével képzett $\alpha\mathbf{x}$ vektorok halmaza, ahol α minden (komplex) számértéket felvehet, vagyis az $\mathbf{x} \neq \mathbf{0}$ vektor által generált egydimenziós altér. Az $\mathbf{R}^{(1)}$ altér akkor és csak akkor invariáns altér, ha \mathbf{Ax} is benne van $\mathbf{R}^{(1)}$ -ben, vagyis ha \mathbf{Ax} az \mathbf{x} vektornak többszöröse, mondjuk λ -szorosa.

2.6.2 definíció. Ha az

$$(2.6.1) \quad \mathbf{Ax} = \lambda\mathbf{x}$$

egyenletnek valamely λ számérték mellett van megoldása, vagyis létezik olyan $\mathbf{x} \neq \mathbf{0}$ vektor, amelyet az \mathbf{A} transzformáció λ -szorosába visz át, akkor a λ számértéket az \mathbf{A} lineáris transzformáció **sajátértékének** nevezzük, az egyenletet kielégítő \mathbf{x} vektort pedig a lineáris transzformáció (λ sajátértékhez tartozó) **sajátvektorának**.

2.6 LINEÁRIS TRANSZFORMÁCIÓ SAJÁTVEKTORAI ÉS SAJÁTÉRTÉKEI 207

Megjegyzés. Nyilvánvaló, hogy ha valamely \mathbf{x} vektor kielégíti a (2.6.1) egyenletet, akkor minden $\alpha \mathbf{x}$ is kielégíti (ahol α zérustól különböző tetszőleges szám lehet). Ezért ezeket nem tekintjük különböző sajátvektoroknak.

Tehát egy lineáris transzformáció egydimenziós invariáns alterének összes, zérustól különböző vektorai sajátvektorok. Az alábbiakban bebizonyítjuk a sajátvektorok létezésére vonatkozó alapvető tételt.

2.6.1 tétel. *Az \mathbf{R}_n (komplex) lineáris térben minden lineáris transzformációnak létezik legalább egy sajátvektora.*

Bizonyítás. Legyen az \mathbf{R}_n tér egy bázisa $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$; legyen továbbá az \mathbf{A} lineáris transzformáció mátrixa ebben a bázisban $\mathbf{A} = [a_{ij}]$. Ha az \mathbf{x} vektornak az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisra vonatkozó koordinátáit $\xi_1, \xi_2, \dots, \xi_n$ jelöli, akkor a sajátvektorokat definiáló $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ összefüggésnek eleget tevő \mathbf{x} vektor $\xi_1, \xi_2, \dots, \xi_n$ koordinátáinak meghatározására az alábbi egyenletrendszer adódik:

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} = \lambda \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}.$$

Ezt rendezés után a következő alakban írhatjuk fel:

$$(2.6.2) \quad \begin{bmatrix} \lambda - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \dots & -a_{2n} \\ \dots & \dots & \dots & \dots \\ -a_{n1} & -a_{n2} & \dots & \lambda - a_{nn} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} = \mathbf{0}.$$

Ennek a homogén lineáris egyenletrendszernek akkor és csak akkor van a triviálisól különböző megoldása, ha az együtthatókból alkotott determináns zérus:

$$(2.6.3) \quad |\lambda \mathbf{E} - \mathbf{A}| = 0.$$

A kapott determinánst kifejtve, λ -ra n -edfokú egyenletet kapunk, amelynek az algebra alaptétele szerint van legalább egy – általában komplex – λ_0 gyöke. Ez a λ_0 szám a transzformáció egy sajátértéke. Ha ezt a számot behelyettesítjük a (2.6.2) egyenletrendszerbe, akkor ennek van nemtriviális $\xi_1^{(0)}, \xi_2^{(0)}, \dots, \xi_n^{(0)}$ megoldása. Az így kapott értékek a transzformáció egy sajátvektorának az adott bázisra vonatkozó koordinátái. Ezzel a tételt bebizonyítottuk. ■

Megjegyzés. A bizonyítás ugyanígy végezhető akkor is, ha az \mathbf{A} transzformációt nem az egész térben, hanem tetszőleges invariáns altérben tekintjük; ezért a tétel a következőképpen is megfogalmazható: *minden invariáns altérben létezik az \mathbf{A} transzformációnak legalább egy sajátvektora.*

Az A lineáris transzformáció sajátértékeinek meghatározására szolgáló (2.6.3) egyenletnek fontos szerepe van a lineáris transzformációk elméletében. Célszerű ezért néhány új fogalmat bevezetni.

2.6.3 definíció. *Bármely (n -edrendű) A mátrixszal képzett*

$$\lambda E - A$$

*mátrixot, ahol λ skalár paraméter, az A mátrixhoz tartozó **karakterisztikus mátrixnak** nevezzük; a karakterisztikus mátrix*

$$(2.6.4) \quad D_n(\lambda) = |\lambda E - A|$$

*determinánsát – amely a λ paraméter (n -edfokú) polinomja – az A mátrix **karakterisztikus polinomjának**, a karakterisztikus polinomból nyert*

$$(2.6.5) \quad D_n(\lambda) \equiv |\lambda E - A| = 0$$

*egyenletet pedig az A mátrix **karakterisztikus egyenletének** nevezzük.*

Eszerint tehát a 2.6.1 tétel bizonyításából az tűnik ki, hogy egy A lineáris transzformáció sajátértékeit bármely bázisra vonatkozó mátrixából megkaphatjuk, hiszen a sajátértékek éppen e mátrix karakterisztikus egyenletének gyökei. Mivel pedig a transzformáció sajátértékeit a bázis megválasztásától függetlenül értelmeztük, ezért ebből az következik, hogy a transzformációhoz tartozó mátrix karakterisztikus egyenletének gyökei szintén függetlenek a bázis megválasztásától. Az alábbi tételben bebizonyítjuk, hogy nemcsak a gyökök, hanem maga a karakterisztikus polinom is független a bázis megválasztásától.

2.6.2 tétel. *Egy A lineáris transzformáció mátrixának karakterisztikus polinomja a bázis transzformációjával szemben invariáns.*

Bizonyítás. Legyen az A lineáris transzformáció mátrixa az e_1, e_2, \dots, e_n bázisban A . Áttérve az

$$[\tilde{e}_1 \tilde{e}_2 \dots \tilde{e}_n] = [e_1 e_2 \dots e_n] \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ \dots & \dots & \dots & \dots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{bmatrix}$$

transzformációval az $\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_n$ új bázisra, az A lineáris transzformációnak ezen új bázisra vonatkozó \tilde{A} mátrixát (2.5.35) alapján az alábbi hasonlósági transzformációval kapjuk meg:

$$(2.6.6) \quad \tilde{A} = C^{-1}AC.$$

Az \tilde{A} mátrix karakterisztikus polinomja $|\lambda E - \tilde{A}|$. Ebbe behelyettesítve a (2.6.6) összefüggést,

$$|\lambda E - \tilde{A}| = |\lambda E - C^{-1}AC|$$

adódik, ahonnan az $\mathbf{E} = \mathbf{C}^{-1}\mathbf{C}$ összefüggés felhasználásával, majd \mathbf{C}^{-1} és \mathbf{C} kiemelésével következik, hogy

$$|\lambda \mathbf{E} - \tilde{\mathbf{A}}| = |\lambda \mathbf{C}^{-1}\mathbf{C} - \mathbf{C}^{-1}\mathbf{A}\mathbf{C}| = |\mathbf{C}^{-1}||\lambda \mathbf{E} - \mathbf{A}||\mathbf{C}| = |\lambda \mathbf{E} - \mathbf{A}|,$$

és ezzel a tételt bebizonyítottuk. ■

E tétel alapján a továbbiakban beszélhetünk a *lineáris transzformáció karakterisztikus polinomjáról*, hiszen megmutattuk, hogy ez nem függ a bázis megválasztásától.

A 2.6.1 tételben beláttuk, hogy minden lineáris transzformációnak van legalább egy sajátvektora. A következőkben feltételeket adunk meg, amelyeknek segítségével a transzformációkat osztályozhatjuk; majd azzal foglalkozunk, hogy a különböző osztályokhoz tartozó sajátérték-feladatokat milyen tulajdonságok jellemzik, és ezek alapján hogyan lehet a transzformáció mátrixát legáttekinthetőbb és legegyszerűbben kezelhető alakban megadni.

Már most rámutatunk arra, hogy az n -dimenziós lineáris tér lineáris transzformációit aszerint oszthatjuk két nagy csoportra, létezik-e n lineárisan független sajátvektoruk vagy sem. Ugyanis, ha a lineáris transzformációnak van n lineárisan független sajátvektora, akkor a transzformáció mátrixa igen egyszerű alakú lehet. Erre vonatkozik az alábbi tétel.

2.6.3 tétel. *Ha az n -dimenziós lineáris tér egy \mathbf{A} lineáris transzformációjának n lineárisan független sajátvektora van, akkor ezeket választva bázisnak, a transzformáció mátrixa diagonálmátrix.*

Bizonyítás. Legyen $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ a lineáris transzformáció n lineárisan független sajátvektora, azaz

$$\begin{aligned} \mathbf{A}\mathbf{u}_1 &= \lambda_1 \mathbf{u}_1, \\ \mathbf{A}\mathbf{u}_2 &= \lambda_2 \mathbf{u}_2, \\ &\dots\dots\dots \\ \mathbf{A}\mathbf{u}_n &= \lambda_n \mathbf{u}_n. \end{aligned}$$

Ha most az $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ sajátvektorokat választjuk bázisnak, akkor az \mathbf{A} lineáris transzformáció mátrixa ebben a bázisban – a 2.4 szakasz alapján –

$$\mathbf{A} = \begin{bmatrix} \lambda_1 & 0 \dots 0 \\ 0 & \lambda_2 \dots 0 \\ & & \ddots \\ 0 & 0 \dots \lambda_n \end{bmatrix}.$$

Ezzel a tételt bebizonyítottuk. ■

Megjegyzés. A bizonyításból egyébként következik, hogy a tétel megfordítása is igaz: ha egy bázisban a transzformáció mátrixa diagonálmátrix, akkor a bázis összes vektorai e transzformáció lineárisan független sajátvektora.

A kapott tételből látható, miért olyan nagy jelentőségű a sajátvektorok vizsgálata. Ha ugyanis egy lineáris transzformációnak van n lineárisan független sajátvektora, akkor a lineáris transzformáció mátrixa – bármilyen bázisban legyen is adott – a 2.5.4 tétel értelmében a koordináta-rendszer alkalmas transzformációjával, vagyis hasonlósági transzformációval diagonálalakra hozható.

A következőkben lineárisan független sajátvektorok létezésére vonatkozó tételeket bizonyítunk be.

2.6.4 tétel. *A lineáris transzformáció egy sajátértékéhez tartozó sajátvektorok alteret alkotnak.*

Bizonyítás. Legyen \mathbf{x}_1 és \mathbf{x}_2 ugyanahhoz a λ_0 sajátértékhez tartozó egy-egy sajátvektor, azaz $A\mathbf{x}_1 = \lambda_0\mathbf{x}_1$ és $A\mathbf{x}_2 = \lambda_0\mathbf{x}_2$. Ekkor \mathbf{x}_1 és \mathbf{x}_2 tetszőleges lineáris kombinációja szintén a λ_0 sajátértékhez tartozó sajátvektor. Ugyanis

$$A(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2) = \alpha_1A\mathbf{x}_1 + \alpha_2A\mathbf{x}_2 = \lambda_0(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2).$$

Vagyis A ugyanahhoz a λ_0 sajátértékhez tartozó sajátvektorai valóban alteret alkotnak. Ennek az alternek a dimenziója egyenlő a λ_0 sajátértékhez tartozó lineárisan független sajátvektorok számával. ■

2.6.5 tétel. *A lineáris transzformáció különböző sajátértékeihez tartozó sajátvektorok lineárisan függetlenek.*

Bizonyítás. Legyenek $\lambda_1, \lambda_2, \dots, \lambda_k$ az A lineáris transzformáció különböző sajátértékei, $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ pedig hozzájuk tartozó sajátvektorok. A tételt teljes indukcióval bizonyítjuk. Egyetlen sajátvektorra az állítás nyilvánvaló. Tegyük most fel, hogy $k-1$ különböző sajátértékhez tartozó sajátvektorok lineárisan függetlenek, és ebből következtetünk arra, hogy k különböző sajátértékhez tartozó sajátvektorok is azok. Ezt indirekt módszerrel láthatjuk be: feltesszük, hogy az $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ sajátvektorok lineárisan összefüggőek, vagyis

$$(2.6.7) \quad \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \dots + \alpha_k\mathbf{u}_k = \mathbf{0},$$

és valamelyik együttható – pl. α_1 – nem zérus. Alkalmazzuk az A transzformációt a (2.6.7) összefüggésre. Mivel a zérusvektort bármely lineáris transzformáció a zérusvektorba viszi át, ezért

$$A(\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \dots + \alpha_k\mathbf{u}_k) = \mathbf{0}.$$

A transzformációt tagonként hajtjuk végre:

$$(2.6.8) \quad \alpha_1\lambda_1\mathbf{u}_1 + \alpha_2\lambda_2\mathbf{u}_2 + \dots + \alpha_k\lambda_k\mathbf{u}_k = \mathbf{0}.$$

Ha most a (2.6.7.) egyenlet λ_k -szorosát levonjuk a (2.6.8) egyenletből, akkor az alábbi összefüggésre jutunk:

$$\alpha_1(\lambda_1 - \lambda_k)\mathbf{u}_1 + \alpha_2(\lambda_2 - \lambda_k)\mathbf{u}_2 + \dots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k)\mathbf{u}_{k-1} = \mathbf{0},$$

2.6 LINEÁRIS TRANSZFORMÁCIÓ SAJÁTVEKTORAI ÉS SAJÁTÉRTÉKEI 211

ahol a fentiek szerint $\alpha_1 \neq 0$ és $\lambda_1 - \lambda_k \neq 0$. Ezek szerint az $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}$ vektorok lineárisan összefüggőek volnának, és így ellentmondásra jutottunk. Ezzel a tételt bebizonyítottuk. ■

Ebből a tételből következik, hogy egy lineáris transzformáció n lineárisan független sajátvektorának létezésére könnyen adható egy elégséges feltétel. Erre vonatkozik a következő tétel.

2.6.6 tétel. *Ha egy lineáris transzformáció sajátértékei mind különbözőek, akkor a hozzájuk tartozó sajátvektorok lineárisan függetlenek.*

Eszerint tehát, ha egy lineáris transzformáció sajátértékei mind különbözőek, akkor a 2.6.3 tétel szerint van olyan bázis, amelyben a transzformáció mátrixa diagonálmátrix. Hangsúlyozni kell azonban, hogy az n különböző sajátérték létezése elégséges, de nem szükséges feltétele annak, hogy a lineáris transzformáció mátrixa diagonalizálható legyen. Más szavakkal: ha egy lineáris transzformáció karakterisztikus egyenletének csupa egyszeres gyöke van, akkor a lineáris transzformáció bármely bázisban megadott mátrixa hasonlósági transzformációval biztosan diagonalizálható; abban az esetben, ha a karakterisztikus egyenletnek van többszörös gyöke is, általában még nem lehet tudni, van-e a transzformációnak n lineárisan független sajátvektora. A 2.6.4 tétel alapján annyi mindenestre megállapítható, hogy amennyiben a transzformáció valamennyi többszörös sajátértékének multiplicitása megegyezik a hozzá tartozó sajátvektorok által generált altér dimenziójával, akkor van a transzformációnak n lineárisan független sajátvektora, tehát a transzformáció mátrixa hasonlósági transzformációval diagonalizálható.

2.6.4 definíció. *Az n -dimenziós lineáris tér egy lineáris transzformációjának n lineárisan független sajátvektorát – feltéve, hogy létezik – egy teljes sajátvektor-rendszernek nevezzük.*

2.6.5 definíció. *Azokat a lineáris transzformációkat, amelyeknek létezik teljes sajátvektor-rendszerük, diagonalizálható lineáris transzformációknak nevezzük.*

2.6.6 definíció. *A hasonlósági transzformációval diagonalizálható mátrixokat egyszerű struktúrájú mátrixoknak nevezzük.*

Tehát a diagonalizálható lineáris transzformációk bármely bázisra vonatkozó mátrixa egyszerű struktúrájú.

24. Példa. Határozzuk meg az $\mathbf{A} = \begin{bmatrix} 6 & 2 & -3 \\ 4 & 6 & -4 \\ 6 & 4 & -3 \end{bmatrix}$ mátrix sajátvektorait és

hasonlósági transzformációval diagonalizáljuk az adott mátrixot!

Megoldás. Először meghatározzuk a mátrix sajátértékeit. A karakterisztikus polinomra

$$D(\lambda) \equiv |\lambda \mathbf{E} - \mathbf{A}| = \begin{vmatrix} \lambda - 6 & -2 & 3 \\ -4 & \lambda - 6 & 4 \\ -6 & -4 & \lambda + 3 \end{vmatrix} = \lambda^3 - 9\lambda^2 + 26\lambda - 24.$$

Könnyen igazolható, hogy a karakterisztikus egyenlet gyökei 2, 3 és 4, tehát a mátrixnak három különböző sajátértéke

$$\lambda_1 = 2, \quad \lambda_2 = 3, \quad \lambda_3 = 4.$$

Ebből következik, hogy a mátrixnak van teljes sajátvektor-rendszere, tehát hasonlósági transzformációval diagonalizálható. Rendre behelyettesítve a λ_k sajátértékeket a

$$(\lambda_k \mathbf{E} - \mathbf{A})\mathbf{u}_k = 0$$

egyenletrendszerbe, a sajátvektorokat az alábbi homogén lineáris egyenlet-rendszerek megoldása szolgáltatja. $\lambda_1 = 2$ esetén

$$\begin{bmatrix} -4 & -2 & 3 \\ -4 & -4 & 4 \\ -6 & -4 & 5 \end{bmatrix} \begin{bmatrix} u_{11} \\ u_{21} \\ u_{31} \end{bmatrix} = 0.$$

Mivel az együtthatómátrix rangja 2, az ismeretlenek arányát az (1.7.33) képlet alapján például a harmadik sor elemeihez tartozó előjeles aldeterminánsokból számíthatjuk ki:

$$u_{11} : u_{21} : u_{31} = A_{31} : A_{32} : A_{33} = 4 : 4 : 8.$$

Ugyanígy $\lambda_2 = 3$ esetén

$$\begin{bmatrix} -3 & -2 & 3 \\ -4 & -3 & 4 \\ -6 & -4 & 6 \end{bmatrix} \begin{bmatrix} u_{12} \\ u_{22} \\ u_{32} \end{bmatrix} = 0.$$

Ismét a harmadik sor elemeihez tartozó előjeles aldeterminánsok segítségével a sajátvektor elemeinek arányára

$$u_{12} : u_{22} : u_{32} = A_{31} : A_{32} : A_{33} = 1 : 0 : 1$$

adódik.

Megjegyzés. Mivel az együtthatómátrix harmadik sora az első sor kétszerese, ezért ebben az esetben *nem* választhatnánk a második sor elemeihez tartozó aldeterminánsok arányát, mert az $0 : 0 : 0$ határozatlan alakot eredményezne.

Végül $\lambda_3 = 4$ esetén

$$\begin{bmatrix} -2 & -2 & 3 \\ -4 & -2 & 4 \\ -6 & -4 & 7 \end{bmatrix} \begin{bmatrix} u_{13} \\ u_{23} \\ u_{33} \end{bmatrix} = 0.$$

A sajátvektor elemeire

$$u_{13} : u_{23} : u_{33} = A_{31} : A_{32} : A_{33} = -2 : -4 : -4$$

adódik. A sajátvektorokat definiáló egyenleteket egyetlen mátrixegyenletbe foglalva

$$\begin{bmatrix} 6 & 2 & -3 \\ 4 & 6 & -4 \\ 6 & 4 & -3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 2 & 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & & \\ & 3 & \\ & & 4 \end{bmatrix}$$

írható. Ebből azonnal látható, hogy a mátrix hasonlósági transzformációval diagonalizálható. Ugyanerre az eredményre jutunk, ha a sajátvektorokat választjuk új bázisnak. A bázistranszformáció mátrixa ugyanis az

$$[\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3] = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 2 & 1 & 2 \end{bmatrix}$$

összefüggésből

$$\mathbf{U} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 2 & 1 & 2 \end{bmatrix},$$

tehát az \mathbf{U} mátrixszal végzett hasonlósági transzformációval $\mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \mathbf{\Lambda}$ adódik. Figyelembe véve, hogy

$$\mathbf{U}^{-1} = \begin{bmatrix} -2 & -1 & 2 \\ 2 & 0 & -1 \\ 1 & 1 & -1 \end{bmatrix},$$

a műveletek elvégzésével a fenti összefüggés helyessége közvetlenül igazolható.

A következőkben megismerkedünk a diagonalizálható transzformációk néhány igen nevezetes osztályával, és megmutatjuk, hogy ezek mátrixával az 1.4 szakaszban vizsgált speciális tulajdonságú mátrixok között már találkoztunk.

2.7 ADJUNGÁLT LINEÁRIS TRANSZFORMÁCIÓK

A 2.3 és 2.4 szakaszban bevezettük a bilineáris alak fogalmát, és vizsgáltuk az n -dimenziós tér transzformációit. Vizsgálataink során azonban nem tételeztük fel, hogy a térben skaláris szorzat van értelmezve. Ha feltesszük, hogy skaláris szorzat is értelmezve van, vagyis euklideszi terekben vizsgáljuk a lineáris transzformációkat és a bilineáris alakokat, akkor a kettő között igen szoros kapcsolatot állapíthatunk meg. Ebben a szakaszban ezt a kapcsolatot mutatjuk meg, és ezen keresztül bevezetjük az adjungált transzformáció fogalmát, aminek segítségével a diagonalizálható transzformációk néhány igen fontos osztálya definiálható.

Legyen \mathbf{E}_n egy n -dimenziós komplex euklideszi tér, és jelöljön $A(\mathbf{x}; \mathbf{y})$ ebben a térben egy bilineáris alakot. Legyen továbbá $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ az \mathbf{E}_n tér egy ortonormált bázisa, és jelölje erre a bázisra nézve a tér \mathbf{x} , ill. \mathbf{y} vektorának koordinátáit $\xi_1, \xi_2, \dots, \xi_n$, ill. $\eta_1, \eta_2, \dots, \eta_n$, az ezekből alkotott koordinátavektorokat \mathbf{x} , ill. \mathbf{y} , az $A(\mathbf{x}; \mathbf{y})$ bilineáris alak \mathbf{A} mátrixának elemeit pedig $a_{ij} = A(\mathbf{e}_j; \mathbf{e}_i)$.

Ekkor az $A(\mathbf{x}; \mathbf{y})$ bilineáris alak a következőképpen írható fel:

$$\begin{aligned} A(\mathbf{x}; \mathbf{y}) &= \sum_{j=1}^n \sum_{i=1}^n \bar{\eta}_i \xi_j A(\mathbf{e}_j; \mathbf{e}_i) = \sum_{j=1}^n \sum_{i=1}^n \bar{\eta}_i a_{ij} \xi_j = \\ &= [\bar{\eta}_1 \quad \bar{\eta}_2 \dots \bar{\eta}_n] \begin{bmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \dots & \dots \dots \dots \\ a_{n1} & a_{n2} \dots a_{nn} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} = \mathbf{y}^H \mathbf{A} \mathbf{x}. \end{aligned}$$

Az $\mathbf{y}^H \mathbf{A} \mathbf{x}$ a mátrixszorzat asszociativitása miatt az

$$(2.7.1) \quad \mathbf{y}^H \mathbf{A} \mathbf{x} = \mathbf{y}^H (\mathbf{A} \mathbf{x}),$$

illetve az

$$(2.7.2) \quad \mathbf{y}^H \mathbf{A} \mathbf{x} = (\mathbf{A}^H \mathbf{y})^H \mathbf{x}$$

alakban írható. S minthogy az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázis ortonormált, e bázisra vonatkozóan a skaláris szorzat standard, így a (2.7.1), ill. (2.7.2) alaknak megfelelően az $A(\mathbf{x}; \mathbf{y})$ bilineáris alak kétféleképpen is felfogható két vektor skaláris szorzataként: (2.7.1) szerint

$$(2.7.3) \quad A(\mathbf{x}; \mathbf{y}) = (\mathbf{A} \mathbf{x}, \mathbf{y}),$$

ahol A az a lineáris transzformáció, melynek az e_1, e_2, \dots, e_n bázisra vonatkozó mátrixa \mathbf{A} , (2.7.2) alapján pedig

$$(2.7.4) \quad A(\mathbf{x}; \mathbf{y}) = (\mathbf{x}, \mathbf{A}^H \mathbf{y}),$$

ahol \mathbf{A}^H az a lineáris transzformáció, amelynek az e_1, e_2, \dots, e_n bázisra vonatkozó mátrixa az \mathbf{A} mátrix transzponált konjugáltja.

Minthogy egy lineáris transzformációt egy adott bázisra vonatkozó mátrixa egyértelműen meghatároz, így a (2.7.3), ill. (2.7.4) jobb oldalán szereplő \mathbf{A} , illetve \mathbf{A}^H transzformációt is egyértelműen meghatározza a (2.7.1), illetve (2.7.2) jobb oldalán álló \mathbf{A} , ill. \mathbf{A}^H mátrix. Az \mathbf{A} mátrix viszont az $A(\mathbf{x}; \mathbf{y})$ bilineáris alaknak az e_1, e_2, \dots, e_n bázisra vonatkozó mátrixa. Ilyen módon az E_n euklideszi térben az $A(\mathbf{x}; \mathbf{y})$ bilineáris alak és az A lineáris transzformáció, az $A(\mathbf{x}; \mathbf{y})$ bilineáris alak és az \mathbf{A}^H lineáris transzformáció, illetve az A és az \mathbf{A}^H lineáris transzformáció kölcsönösen egyértelműen meghatározzák egymást az alábbi összefüggések alapján:

$$(2.7.5) \quad A(\mathbf{x}; \mathbf{y}) = (\mathbf{A}\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{A}^H \mathbf{y}).$$

Eredményeinket a 2.7.1 és 2.7.2 tételben fogalmazzuk meg.

2.7.1 tétel. *Az E_n (komplex) euklideszi térben minden $A(\mathbf{x}; \mathbf{y})$ bilineáris alakhoz egyértelműen hozzárendelhető egy A lineáris transzformáció, amelynek segítségével a bilineáris alak az $\mathbf{A}\mathbf{x}$ és az \mathbf{y} vektor skaláris szorzataként írható fel:*

$$A(\mathbf{x}; \mathbf{y}) = (\mathbf{A}\mathbf{x}, \mathbf{y}),$$

és megfordítva, minden A lineáris transzformációhoz az

$$(\mathbf{A}\mathbf{x}, \mathbf{y}) = A(\mathbf{x}; \mathbf{y})$$

összefüggéssel egyértelműen tartozik egy bilineáris alak.

Mielőtt a 2.7.2 tételt megfogalmazzuk, bevezetjük az adjungált transzformáció fogalmát.

2.7.1 definíció. *Ha A a komplex n -dimenziós euklideszi tér egy lineáris transzformációja, akkor az*

$$(2.7.6) \quad (\mathbf{A}\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{A}^H \mathbf{y})$$

összefüggéssel meghatározott \mathbf{A}^H transzformációt az A transzformáció **adjungáltjának** nevezzük.

Megjegyzés. Az előzőek alapján nyilvánvaló, hogy bármely ortonormált bázisban az adjungált transzformáció mátrixa az eredeti transzformáció mátrixának transzponált konjugáltja:

$$(2.7.7) \quad \mathbf{A}^H = \overline{\mathbf{A}}^T.$$

Ez indokolja, hogy a mátrix transzponált konjugáltját egyszerűen H betűvel jelöltük. Sajnos nem nevezhetjük „adjungált” mátrixnak, hiszen ez az elnevezés a mátrixalgebrában már más értelemben (adj $A = [A_{ji}]$) honosodott meg.

2.7.2 tétel. *Az euklideszi térben minden lineáris transzformációnak egy és csak egy adjungált transzformáció felel meg.*

Az áttérés az adjungált transzformációra formálisan a következőképpen fogalmazható: az (Ax, y) skaláris szorzatban az A transzformációt úgy „vihetjük át” a második helyre, hogy oda az adjungáltját tesszük (H betűvel látjuk el).

A lineáris transzformációkra megismert műveleti szabályok alkalmazásával és az adjungált transzformáció definíciójával könnyen beláthatók az adjungálásra vonatkozó következő azonosságok:

- (a) $(AB)^H = B^H A^H$,
- (b) $(A^H)^H = A$,
- (c) $(A + B)^H = A^H + B^H$,
- (d) $(\lambda A)^H = \lambda A^H$,
- (e) $E^H = E$.

2.8 DIAGONALIZÁLHATÓ TRANSZFORMÁCIÓK, TRANSZFORMÁCIÓPÁROK, ÁLTALÁNOSÍTOTT SAJÁTÉRTÉK-FELADAT

Az adjungált transzformáció fogalma lehetővé teszi néhány fontos, speciális transzformációtípus bevezetését. Ebben a szakaszban először az elmélet és az alkalmazások szempontjából egyaránt jelentős önadjungált és unitér transzformációkat, majd a normális transzformációkat vizsgáljuk. E transzformációk közös tulajdonsága, hogy teljes sajátvektor-rendszerük van, tehát valamennyien diagonalizálhatók. Az önadjungált transzformációkra vonatkozóan bebizonyítjuk az alapvető fontosságai főtengetyételt, majd általánosítjuk transzformációpárokra, és így eljutunk az általánosított sajátérték-feladathoz. Végül megmutatjuk az önadjungált transzformációk, ill. transzformációpárok sajátértékeinek egy nevezetes extrémális tulajdonságát.

2.8.1 Önadjungált transzformációk

Az adjungált transzformáció értelmezése alapján kézenfekvő az önadjungált transzformáció fogalmának bevezetése a következő módon.

2.8.1 definíció. *Az E euklideszi tér A lineáris transzformációját önadjungáltnak nevezzük, ha*

$$(2.8.1) \quad A^H = A,$$

vagyis ha A megegyezik az adjungáltjával.

A lineáris transzformációk és a bilineáris alakok közötti összefüggés alapján közvetlenül belátható az alábbi tétel.

2.8.1 tétel. Az E euklideszi tér A lineáris transzformációja akkor és csak akkor önadjungált, ha az $A(\mathbf{x}; \mathbf{y}) = (A\mathbf{x}, \mathbf{y})$ bilineáris alak hermitikus.

Bizonyítás. Ha $A(\mathbf{x}; \mathbf{y})$ hermitikus alak, vagyis

$$A(\mathbf{x}; \mathbf{y}) = \overline{A(\mathbf{y}; \mathbf{x})}$$

akkor ebből a megfelelő lineáris transzformációra áttérve,

$$(A\mathbf{x}, \mathbf{y}) = \overline{(A\mathbf{y}, \mathbf{x})} = (\mathbf{x}, A\mathbf{y}),$$

másrészt – az adjungált transzformáció (2.7.6) definiáló összefüggése szerint –

$$(A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, A^H\mathbf{y}),$$

vagyis A valóban önadjungált transzformáció.

Fordítva, ha az A lineáris transzformáció önadjungált, akkor

$$(A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, A\mathbf{y}) = \overline{(A\mathbf{y}, \mathbf{x})}$$

miatt a megfelelő bilineáris alakra fennáll, hogy

$$A(\mathbf{x}; \mathbf{y}) = \overline{A(\mathbf{y}; \mathbf{x})},$$

tehát az valóban hermitikus. ■

A tételből következik az is, hogy az A önadjungált lineáris transzformáció mátrixa bármely ortonormált bázisban hermitikus mátrix.

2.8.2 tétel. Az euklideszi tér bármely A lineáris transzformációja felírható

$$(2.8.2) \quad A = A_1 + iA_2$$

alakban, ahol A_1 és A_2 önadjungált transzformációk és i az imaginárius egység.

Bizonyítás. Az euklideszi tér A lineáris transzformációjára felírható az alábbi azonosság:

$$A = \frac{1}{2}(A + A^H) + \frac{i}{2i}(A - A^H).$$

Bevezetjük a következő jelöléseket:

$$A_1 = \frac{1}{2}(A + A^H), \quad A_2 = \frac{1}{2i}(A - A^H).$$

Azonnal látható, hogy A_1 és A_2 önadjungált, hiszen

$$A_1^H = \left(\frac{1}{2}(A + A^H) \right)^H = \frac{1}{2}(A^H + A) = \frac{1}{2}(A + A^H) = A_1,$$

és

$$A_2^H = \left(\frac{1}{2i}(A - A^H) \right)^H = -\frac{1}{2i}(A^H - (A^H)^H) = -\frac{1}{2i}(A^H - A) = A_2,$$

és ezzel a tételt bebizonyítottuk. ■

A kapott tétel azt fejezi ki, hogy az önadjungált transzformációk hasonló szerepet töltenek be a lineáris transzformációk között, mint a valós számok a komplex számok körében. (A mátrixokra vonatkozó megfelelő tételt megismertük az 1. fejezetben, lásd az 1.2.1 tételt.) Ez az analógia még élesebben megmutatkozik az önadjungált transzformációknak abban az érdekes tulajdonságában, hogy sajátértékeik mindig a valós tengelyen helyezkednek el. Az erre vonatkozó alábbi tétel, amely az alkalmazások szempontjából rendkívül jelentőséggel, könnyen bizonyítható a lineáris algebra imént megismert eszközeivel.

2.8.3 tétel. *Önadjungált transzformáció sajátértékei mindig valósak.*

Bizonyítás. Legyen u az A önadjungált transzformáció sajátvektora, és λ a hozzá tartozó sajátérték, vagyis teljesüljön

$$Au = \lambda u, \quad \text{ahol} \quad u \neq 0.$$

Mivel A önadjungált, vagyis $A^H = A$, ezért $(Au, u) = (u, Au)$, azaz $(\lambda u, u) = (u, \lambda u)$. Kiemelve a λ értéket:

$$\lambda(u, u) = \overline{\lambda}(u, u),$$

és mivel $(u, u) \neq 0$, ezért $\lambda = \overline{\lambda}$, amit bizonyítani kellett. ■

2.8.2 Unitér transzformációk

A lineáris transzformációk között – az önadjungált transzformációk mellett – sok szempontból különleges helyet foglalnak el az unitér transzformációk (lásd a 2.5.3 definíciót).

2.8.2 definíció. *Az euklideszi tér U lineáris transzformációját unitér transzformációnak nevezzük, ha*

$$(2.8.3) \quad UU^H = UU^{-1} = E,$$

vagyis ha

$$(2.8.4) \quad U^{-1} = U^H,$$

azaz ha az inverz transzformáció megegyezik az adjungált transzformációval.

A definíció alapján belátható az unitér transzformáció következő nevezetes tulajdonsága.

2.8.4 tétel. Az n -dimenziós euklideszi térben a tér egy lineáris transzformációjával szemben a skaláris szorzat akkor és csak akkor invariáns, ha a transzformáció unitér.

Bizonyítás. Legyen U unitér transzformáció. Ekkor az adjungált transzformáció (2.7.6) és az unitér transzformáció (2.8.3) definiáló egyenlete alapján

$$(Ux, Uy) = (x, U^H U y) = (x, y),$$

vagyis a skaláris szorzat valóban invariáns.

Megfordítva, ha valamely U transzformációval szemben bármely x, y vektorpárra a skaláris szorzat invariáns, vagyis

$$(2.8.5) \quad (Ux, Uy) = (x, y),$$

akkor

$$(U^H U x, y) = (x, y), \text{ azaz } ((U^H U - E)x, y) = 0.$$

Mínthogy ez tetszőleges x és y mellett fennáll, így szükségképpen

$$U^H U - E = 0.$$

Ezzel a tételt bebizonyítottuk. ■

Az $x = y$ esetben (2.8.5) az alábbi alakot veszi fel:

$$(2.8.6) \quad (Ux, Ux) = (x, x),$$

vagyis az unitér transzformáció nem változtatja meg a vektorok hosszát.

Vizsgáljuk meg, mi jellemzi az unitér transzformációk mátrixát. Legyen e_1, e_2, \dots, e_n az n -dimenziós euklideszi tér egy ortonormált bázisa és $U = [u_{ik}]$ az U unitér transzformációnak erre a bázisra vonatkozó mátrixa. Mivel (2.7.7) szerint az adjungált transzformáció mátrixa ortonormált bázisban egyenlő az eredeti transzformáció mátrixának transzponált konjugáltjával, továbbá az egységtranszformáció mátrixa az egységmátrixszal egyenlő, az unitér transzformáció 2.8.2 definíciójából következik, hogy az U mátrix kielégíti az alábbi azonosságot:

$$U^H U = U U^H = E.$$

Tehát unitér transzformáció mátrixa ortonormált bázisban unitér mátrix.

A 2.8.4 tétel alapján, ha e_1, e_2, \dots, e_n az n -dimenziós euklideszi tér egy ortonormált bázisa és U a tér egy unitér transzformációja, akkor

$$(Ue_i, Ue_k) = (e_i, e_k) = \delta_{ik},$$

vagyis unitér transzformáció a tér bármely ortonormált bázisát ismét ortonormált bázisra transzformálja. Ugyanebből a tételből az is következik, hogy ha egy transzformáció tetszőleges ortonormált bázist ismét ortonormált bázisba visz át, akkor a transzformáció unitér. Vagyis érvényes a következő tétel.

2.8.5 tétel. *Az euklideszi tér valamely lineáris transzformációja akkor és csak akkor unitér, ha bármely ortonormált bázist ugyancsak ortonormált bázisba visz át.*

A következő tétel unitér transzformációk sajátértékeinek elhelyezkedésére vonatkozik.

2.8.6 tétel. *Unitér transzformáció sajátértékeinek abszolút értéke 1.*

Bizonyítás. Legyen u az U unitér transzformáció λ sajátértékéhez tartozó sajátvektor, vagyis

$$(2.8.7) \quad Uu = \lambda u \quad (u \neq 0).$$

A (2.8.6) azonosságból (2.8.7) felhasználásával a következő adódik:

$$(u, u) = (Uu, Uu) = (\lambda u, \lambda u).$$

Mindkét tényezőtől megfelelően kiemelve λ értékét, azt kapjuk, hogy

$$(u, u) = \lambda \bar{\lambda} (u, u),$$

amiből következik $\lambda \bar{\lambda} = 1$, azaz $|\lambda| = 1$, amit bizonyítani kellett. ■

2.8.3 Felcserélhető és normális transzformációk

Az alábbiakban először a felcserélhető transzformációkkal* foglalkozunk, majd az euklideszi térben definiált transzformációk egy fontos osztályát, a normális transzformációkat vizsgáljuk. Ennek jelentőségét többek között az adja meg, hogy mind az önadjungált, mind pedig az unitér transzformációk a normális transzformációk egy-egy speciális esetét alkotják.

*Két transzformációt akkor nevezünk felcserélhetőnek, ha a szorzásra nézve kommutatívak.

2.8.7 tétel. Ha A és B az \mathbf{R} lineáris térben definiált két, felcserélhető lineáris transzformáció, azaz

$$(2.8.8) \quad AB = BA,$$

akkor az A transzformáció egy λ sajátértékéhez tartozó összes sajátvektorok a B transzformációval szemben invariáns alteret alkotnak.

Részletesebben kifejtve ez azt jelenti, hogy ha \mathbf{x} az A transzformáció λ sajátértékéhez tartozó valamely sajátvektor, akkor egyúttal $B\mathbf{x}$ is a λ sajátértékhez tartozó sajátvektor, vagyis a B transzformáció nem vezet ki az A transzformáció λ sajátértékhez tartozó lineáris altérből:

$$(2.8.9) \quad AB\mathbf{x} = \lambda B\mathbf{x}.$$

Bizonyítás. Mivel A és B felcserélhető, ezért $AB\mathbf{x} = BA\mathbf{x}$, ahonnan $A\mathbf{x} = \lambda\mathbf{x}$ miatt $AB\mathbf{x} = B\lambda\mathbf{x} = \lambda B\mathbf{x}$, és ezzel a tételt bebizonyítottuk. ■

2.8.8 tétel. Bármely két felcserélhető transzformációnak van közös sajátvektora.

Bizonyítás. Legyen $AB = BA$ és legyen $\mathbf{R}^{(1)}$ az A transzformáció λ sajátértékéhez tartozó valamennyi \mathbf{x} sajátvektorának altere. A 2.8.7 tétel értelmében az $\mathbf{R}^{(1)}$ altér invariáns a B transzformációra nézve, és ezért a 2.6.1 tétel utáni megjegyzés értelmében található benne olyan \mathbf{x} vektor, amely sajátvektora a B transzformációnak. Ez a vektor sajátvektora az A transzformációnak is, tehát ezzel a tételt bebizonyítottuk. Meg kell azonban jegyezni, hogy az A transzformáció nem minden sajátvektora sajátvektora B -nek is! ■

A következő tételben a felcserélhetőség mint szükséges és elégséges feltétel szerepel.

Az 1. fejezetben már rámutattunk arra, hogy két hermitikus mátrix szorzata általában nem hermitikus. Ennek megfelelően általában az euklideszi térben két önadjungált transzformáció szorzata sem önadjungált transzformáció. Könnyen belátható azonban a következő tétel:

2.8.9 tétel. Az euklideszi térben két önadjungált transzformáció szorzata akkor és csak akkor önadjungált transzformáció, ha a két transzformáció felcserélhető.

Bizonyítás. Legyen A és B a két önadjungált transzformáció. Az AB szorzat adjungáltja

$$(AB)^H = B^H A^H = BA;$$

az AB szorzatra tehát akkor és csak akkor teljesül az

$$(AB)^H = AB$$

összefüggés, azaz AB akkor és csak akkor önadjungált, ha

$$AB = BA,$$

vagyis ha A és B felcserélhető. ■

A következő tétel arra vonatkozóan mond ki szükséges és elégséges feltételt, hogy az euklideszi térben adott lineáris transzformációhoz létezzék olyan ortonormált bázis, amelyben a transzformáció mátrixa diagonálmátrix. Ehhez szükségünk van a normális transzformáció fogalmára.

2.8.3 definíció. Az euklideszi tér A lineáris transzformációját **normális transzformációnak** nevezzük, ha

$$(2.8.10) \quad AA^H = A^H A,$$

azaz ha a transzformáció felcserélhető az adjungáltjával.

2.8.10 tétel. Az euklideszi tér A lineáris transzformációjához akkor és csak akkor található olyan ortonormált bázis, amelyben a transzformáció mátrixa diagonálmátrix, ha a transzformáció normális.

Bizonyítás. Először azt látjuk be, hogy a (2.8.10) feltétel szükséges. Legyen egy ortonormált bázisban az A transzformáció mátrixa diagonálmátrix:

$$\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

Ortonormált bázisban az A^H adjungált transzformáció mátrixa az eredetinek transzponált konjugáltja, ezért most az A^H transzformáció mátrixa:

$$\begin{bmatrix} \bar{\lambda}_1 & & & \\ & \bar{\lambda}_2 & & \\ & & \ddots & \\ & & & \bar{\lambda}_n \end{bmatrix}.$$

Mivel az A és A^H transzformáció mátrixai diagonálmátrixok, ezért felcserélhetők, következésképpen felcserélhetők maguk a transzformációk is.

Azt, hogy a feltétel elegendő is, a következőképpen látjuk be. Feltesszük az A és A^H transzformációról, hogy felcserélhető. Ekkor a 2.8.8 tétel értelmében az A és az A^H transzformációnak van közös e_1 sajátvektora, amelyre tehát

$$Ae_1 = \lambda_1 e_1, \quad A^H e_1 = \mu_1 e_1.$$

Az \mathbf{e}_1 sajátvektorra ortogonális vektorokból álló $(n-1)$ -dimenziós $\mathbf{E}^{(1)}$ altér invariáns mind az \mathbf{A} , mind pedig az \mathbf{A}^H transzformációra nézve. Ugyanis legyen $\mathbf{x} \in \mathbf{E}^{(1)}$, vagyis $(\mathbf{x}, \mathbf{e}_1) = 0$; ekkor

$$(\mathbf{A}\mathbf{x}; \mathbf{e}_1) = (\mathbf{x}, \mathbf{A}^H \mathbf{e}_1) = (\mathbf{x}, \mu_1 \mathbf{e}_1) = \bar{\mu}_1 (\mathbf{x}, \mathbf{e}_1) = 0,$$

tehát valóban $\mathbf{A}\mathbf{x} \in \mathbf{E}^{(1)}$. Ezzel beláttuk, hogy $\mathbf{E}^{(1)}$ invariáns az \mathbf{A} transzformációra nézve, és ugyanígy belátható, hogy $\mathbf{E}^{(1)}$ invariáns az \mathbf{A}^H transzformációra nézve is.

Ha ismét alkalmazzuk a 2.8.8 tételt, de most az $\mathbf{E}^{(1)}$ euklideszi térre, akkor látható, hogy $\mathbf{E}^{(1)}$ -ben van olyan \mathbf{e}_2 vektor, amely sajátvektora mind az \mathbf{A} , mind pedig az \mathbf{A}^H transzformációnak. Jelölje $\mathbf{E}^{(2)}$ azt az $(n-2)$ -dimenziós alteret, amelyet az $\mathbf{E}^{(1)}$ altér \mathbf{e}_2 vektorra merőleges vektorai alkotnak. Az eljárást folytatva, páronként ortogonális $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ vektorokhoz jutunk, amelyek mindegyike sajátvektora \mathbf{A} -nak is és \mathbf{A}^H -nak is. Ezért az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ vektorok olyan ortogonális bázist alkotnak, amelyben mind az \mathbf{A} , mint pedig az \mathbf{A}^H transzformáció mátrixa diagonálmátrix, és ezzel a tételt bebizonyítottuk. ■

Ha \mathbf{A} önadjungált transzformáció, vagyis $\mathbf{A}^H = \mathbf{A}$, akkor $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$, vagyis \mathbf{A} normális transzformáció.

Ha \mathbf{U} unitér transzformáció, vagyis $\mathbf{U}^H = \mathbf{U}^{-1}$, akkor – mivel bármely transzformáció az inverzével felcserélhető – fennáll, hogy

$$\mathbf{U}\mathbf{U}^H = \mathbf{U}^H\mathbf{U},$$

vagyis az unitér transzformáció szintén normális transzformáció.

Azt látjuk tehát, hogy az önadjungált és az unitér transzformáció a normális transzformációknak egy-egy speciális esete; az imént bebizonyított 2.8.10 tétel alapján megállapíthatjuk, hogy az n -dimenziós euklideszi térben mind az önadjungált, mind pedig az unitér transzformációk esetén mindig található n darab, páronként ortogonális sajátvektor, és ezeket választva bázisnak, a transzformáció mátrixa diagonálmátrix.

A 2.8.5 tételben kimondtuk, hogy az a transzformáció, amely minden ortonormált bázist (másik) ortonormált bázisba visz át, unitér transzformáció; tehát az ennek megfelelő koordinátatranszformáció mátrixa unitér mátrix. Ebből következik, hogy ha egy normális transzformáció mátrixát ismerjük egy ortonormált bázisban, akkor – mivel a sajátvektorokból alkotott bázisra unitér transzformációval térhetünk át – a normális transzformáció mátrixa unitér transzformációval diagonalizálható:

$$(2.8.11) \quad \mathbf{\Lambda} = \mathbf{U}^H \mathbf{A} \mathbf{U},$$

ahol $\mathbf{\Lambda} = \langle \lambda_k \rangle$ az adott normális transzformáció sajátértékeiből alkotott diagonálmátrix, \mathbf{A} a normális transzformáció mátrixa tetszőleges $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$

ortonormált bázisban, \mathbf{U} pedig az a mátrix, amelynek oszlopai az adott normális transzformáció sajátvektorainak a koordinátái az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisban.

2.8.4 definíció. Ha az euklideszi tér \mathbf{A} lineáris transzformációjának van teljes sajátvektor-rendszere, akkor azt a mátrixot, amelynek oszlopait a transzformáció sajátvektorainak adott bázisra vonatkozó koordinátái alkotják, a transzformáció \mathbf{e} bázisra vonatkozó **modálmátrixának** nevezzük.

25. Példa. Mutassuk meg, hogy az

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & 1 \end{bmatrix}$$

mátrix normális és diagonalizáljuk unitér transzformációval.

Megoldás. Tekintettel arra, hogy

$$\mathbf{A}\mathbf{A}^\top = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix}$$

és

$$\mathbf{A}^\top \mathbf{A} = \begin{bmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix},$$

azaz $\mathbf{A}\mathbf{A}^\top = \mathbf{A}^\top \mathbf{A}$, tehát \mathbf{A} normális mátrix.

Az \mathbf{A} mátrix karakterisztikus polinomja

$$\begin{aligned} D(\lambda) \equiv |\lambda \mathbf{E} - \mathbf{A}| &= \begin{vmatrix} \lambda - 1 & 1 & -1 \\ -1 & \lambda - 1 & 1 \\ 1 & -1 & \lambda - 1 \end{vmatrix} = (\lambda - 1)^3 + 3(\lambda - 1) = \\ &= (\lambda - 1)[(\lambda - 1)^2 + 3], \end{aligned}$$

tehát a sajátértékeire

$$\lambda_1 = 1, \quad \lambda_2 = 1 + i\sqrt{3}, \quad \lambda_3 = 1 - i\sqrt{3}$$

adódik. A sajátvektorok az alábbi egyenletrendszerekből számíthatók:

$$\lambda_1 = 1$$

$$\begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} u_{11} \\ u_{21} \\ u_{31} \end{bmatrix} = 0.$$

Innen $u_{11} = u_{21} = u_{31} = 1 : 1 : 1$.

$$\lambda_2 = 1 + i\sqrt{3}$$

$$\begin{bmatrix} i\sqrt{3} & 1 & -1 \\ -1 & i\sqrt{3} & 1 \\ 1 & -1 & i\sqrt{3} \end{bmatrix} \begin{bmatrix} u_{12} \\ u_{22} \\ u_{32} \end{bmatrix} = 0.$$

Innen $u_{12} : u_{22} : u_{32} = (1 + i\sqrt{3}) : (1 - i\sqrt{3}) : (-2)$.

$$\lambda_3 = 1 - i\sqrt{3}$$

$$\begin{bmatrix} -i\sqrt{3} & 1 & -1 \\ -1 & -i\sqrt{3} & 1 \\ 1 & -1 & -i\sqrt{3} \end{bmatrix} \begin{bmatrix} u_{13} \\ u_{23} \\ u_{33} \end{bmatrix} = 0.$$

Innen $u_{13} : u_{23} : u_{33} = (1 - i\sqrt{3}) : (1 + i\sqrt{3}) : (-2)$. Mivel

$$|\mathbf{u}_1|^2 = 3, \quad |\mathbf{u}_2|^2 = 12, \quad |\mathbf{u}_3|^2 = 12,$$

a normált sajátvektorokból alkotott modálmátrix

$$\mathbf{U} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & \frac{1+i\sqrt{3}}{2} & \frac{1-i\sqrt{3}}{2} \\ 1 & \frac{1-i\sqrt{3}}{2} & \frac{1+i\sqrt{3}}{2} \\ 1 & -1 & -1 \end{bmatrix}$$

és mivel

$$\mathbf{U}^H \mathbf{U} = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ \frac{1-i\sqrt{3}}{2} & \frac{1+i\sqrt{3}}{2} & -1 \\ \frac{1+i\sqrt{3}}{2} & \frac{1-i\sqrt{3}}{2} & -1 \end{bmatrix} \begin{bmatrix} 1 & \frac{1+i\sqrt{3}}{2} & \frac{1-i\sqrt{3}}{2} \\ 1 & \frac{1-i\sqrt{3}}{2} & \frac{1+i\sqrt{3}}{2} \\ 1 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix},$$

tehát az \mathbf{U} modálmátrix unitér mátrix, ennek segítségével végzett unitér transzformációval az \mathbf{A} mátrix diagonalizálható:

$$\begin{aligned} & \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 1 \\ \frac{1-i\sqrt{3}}{2} & \frac{1+i\sqrt{3}}{2} & -1 \\ \frac{1+i\sqrt{3}}{2} & \frac{1-i\sqrt{3}}{2} & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & \frac{1+i\sqrt{3}}{2} & \frac{1-i\sqrt{3}}{2} \\ 1 & \frac{1-i\sqrt{3}}{2} & \frac{1+i\sqrt{3}}{2} \\ 1 & -1 & -1 \end{bmatrix} \frac{1}{\sqrt{3}} = \\ & = \begin{bmatrix} 1 & & \\ & 1 + i\sqrt{3} & \\ & & 1 - i\sqrt{3} \end{bmatrix}. \\ & \qquad \qquad \qquad * \quad * \quad * \end{aligned}$$

2.8.4 Pozitív definit transzformációk

Most bevezetjük a pozitív definit transzformációk fogalmát, majd ezek fontosabb tulajdonságait vizsgáljuk.

2.8.5 definíció. Az euklideszi tér A lineáris transzformációját **pozitív definitnek** nevezzük, ha A önadjungált és $(Ax, x) > 0$ minden $x \neq 0$ vektorra, és **pozitív szemidefinitnek** nevezzük, ha $(Ax, x) \geq 0$ minden $x \neq 0$ -ra.

A 2.8.2 tételben láttuk, hogy a komplex számok algebrai alakjának analógiájára minden lineáris transzformáció felírható $A_1 + iA_2$ alakban, ahol A_1 és A_2 önadjungált transzformációk. A következőkben a komplex számok és a lineáris transzformációk közötti további analógiára mutatunk rá. Emlékeztünk arra, hogy egy (zérustól különböző) komplex szám exponenciális alakja $(\varrho e^{i\varphi})$ egy pozitív valós számnak (ϱ) és egy egységnyi abszolút értékű komplex számnak $(e^{i\varphi})$ a szorzata. Mielőtt a lineáris transzformációkra vonatkozó analóg tételt kimondanánk, bebizonyítunk a pozitív definit lineáris transzformációkra vonatkozó néhány tételt.

2.8.11 tétel. Ha A az euklideszi tér nemszinguláris lineáris transzformációja, akkor az AA^H transzformáció pozitív definit.

Bizonyítás. Mivel

$$(AA^H)^H = (A^H)^H A^H = AA^H,$$

tehát AA^H önadjungált; továbbá mivel az A^H transzformáció is nemszinguláris lineáris transzformáció és minden $x \neq 0$ vektorra $A^H x \neq 0$ (ha ugyanis $A^H x = 0$ állna fenn, akkor az inverz transzformáció alkalmazásával $x = 0$ adódna), ezért $(A^H x, A^H x) > 0$; ennek felhasználásával

$$(AA^H x, x) = (A^H x, A^H x) > 0,$$

azaz AA^H a 2.8.5 definíció értelmében valóban pozitív definit transzformáció. ■

2.8.12 tétel. Az euklideszi tér A önadjungált transzformációja akkor és csak akkor pozitív definit, ha minden sajátértéke pozitív.

Bizonyítás. Legyen az A transzformáció λ sajátértékéhez tartozó sajátvektor u ; ekkor $Au = \lambda u$, és így $(Au, u) = (\lambda u, u) = \lambda(u, u)$. Ha A pozitív definit, akkor a 2.8.5 definíció értelmében $(u, u) > 0$ és $(Au, u) > 0$; ebből következik, hogy $\lambda > 0$, vagyis a feltétel szükséges.

Tegyük most fel, hogy az A önadjungált transzformáció mindegyik sajátértéke pozitív, és legyen u_1, u_2, \dots, u_n a transzformáció sajátvektoraiból alkotott ortonormált bázis; legyen továbbá x a tér tetszőleges vektora, amelynek

koordinátái az $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ bázisra nézve $\xi_1, \xi_2, \dots, \xi_n$, azaz

$$\boldsymbol{x} = \sum_{i=1}^n \xi_i \boldsymbol{u}_i.$$

Képezzük az (Ax, x) skaláris szorzatot:

$$(Ax, x) = \left(\sum_{j=1}^n \xi_j A u_j, \sum_{i=1}^n \xi_i u_i \right) = \left(\sum_{j=1}^n \xi_j \lambda_j u_j, \sum_{i=1}^n \xi_i u_i \right) = \sum_{i=1}^n \lambda_i |\xi_i|^2.$$

Mivel az összes λ_i sajátérték pozitív, ezért minden $\mathbf{x} \neq 0$ vektorra $(\mathbf{A}\mathbf{x}, \mathbf{x}) > 0$, tehát \mathbf{A} pozitív definit, vagyis a feltétel elégséges is. ■

Most térjünk vissza a 2.2.1 tétel bizonyítására. E tétel szerint, ha $\mathbf{x}_1, \dots, \mathbf{x}_k$ az n -dimenziós \mathbf{E}_n euklideszi tér vektorai, akkor a

$$G(\mathbf{x}_1, \dots, \mathbf{x}_k) = \frac{\begin{vmatrix} (\mathbf{x}_1, \mathbf{x}_1) & \dots & (\mathbf{x}_1, \mathbf{x}_k) \\ \vdots & \ddots & \vdots \\ (\mathbf{x}_k, \mathbf{x}_1) & \dots & (\mathbf{x}_k, \mathbf{x}_k) \end{vmatrix}}{\begin{vmatrix} (\mathbf{x}_1, \mathbf{x}_1) & \dots & (\mathbf{x}_1, \mathbf{x}_k) \\ \vdots & \ddots & \vdots \\ (\mathbf{x}_k, \mathbf{x}_1) & \dots & (\mathbf{x}_k, \mathbf{x}_k) \end{vmatrix}}.$$

Gram-féle determináns nemnegatív, és zérussal akkor és csak akkor egyenlő, ha az $\mathbf{x}_1, \dots, \mathbf{x}_k$ vektorok lineárisan összefüggőek.

Bizonyítás. Az $\mathbf{x}_1, \dots, \mathbf{x}_k$ vektorok lineárisan függetlenek, ha a

$$(2.8.12) \quad c_1 \mathbf{x}_1 + \cdots + c_k \mathbf{x}_k = \mathbf{0}$$

összefüggés csak $c_1 = c_2 = \dots = c_k = 0$ esetén áll fenn. Képezve a (2.8.12) egyenlet skaláris szorzatát rendre az $\mathbf{x}_1, \dots, \mathbf{x}_k$ vektorokkal, a c_1, \dots, c_k ismeretlenekre egy homogén lineáris egyenletrendszert kapunk:

$$(2.8.13) \quad \begin{array}{rcl} c_1(\mathbf{x}_1, \mathbf{x}_1) + \cdots + c_k(\mathbf{x}_1, \mathbf{x}_k) & = & 0 \\ \dots\dots\dots & & \\ c_1(\mathbf{x}_k, \mathbf{x}_1) + \cdots + c_k(\mathbf{x}_k, \mathbf{x}_k) & = & 0. \end{array}$$

Írjuk fel a (2.8.13) egyenletrendszert mátrixegyenletként:

$$(2.8.14) \quad \mathbf{G}\mathbf{c} = \mathbf{0}.$$

A (2.8.14) egyenletnek akkor és csak akkor létezik $\mathbf{c} \neq \mathbf{0}$ megoldása, ha az $\mathbf{x}_1, \dots, \mathbf{x}_k$ vektorok lineárisan összefüggőek. Ekkor a Gram-féle determináns értéke zérus: $|\mathbf{G}| = 0$. Ha a (2.8.14) egyenletnek csak a $\mathbf{c} = 0$ megoldása létezik, akkor az $\mathbf{x}_1, \dots, \mathbf{x}_k$ vektorok lineárisan függetlenek, azaz $|\mathbf{G}| \neq 0$ és értéke valós (\mathbf{G} hermitikus mátrix).

Az $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ lineárisan független vektorok kifeszítik az \mathbf{E}_n tér egy k -dimenziós \mathbf{E}_k alterét, ami szintén euklideszi tér és egy bázisa: $\mathbf{x}_1, \dots, \mathbf{x}_k$. Vegyük észre, hogy a

$$\mathbf{G}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k) = [(\mathbf{x}_i, \mathbf{x}_j)]$$

mátrix felfogható úgy, mint az \mathbf{E}_k euklideszi térben értelmezett

$$(2.8.15) \quad \mathbf{A}(\mathbf{u}; \mathbf{v}) = (\mathbf{u}; \mathbf{v})$$

hermitikus bilineáris alaknak az $\mathbf{x}_1, \dots, \mathbf{x}_k$ bázisra vonatkozó mátrixa. De a (2.8.15) bilineáris alaknak az \mathbf{E}_k tér bármely $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_k$ ortonormált bázisra vonatkozó mátrixa az \mathbf{E} egységmátrix. Legyen \mathbf{C} a bázistranszformáció mátrixa, azaz

$$[\mathbf{x}_1 \quad \mathbf{x}_2 \dots \mathbf{x}_k] = [\mathbf{f}_1 \quad \mathbf{f}_2 \dots \mathbf{f}_k] \mathbf{C}.$$

Ekkor a 2.5.2 tétel alapján

$$[(\mathbf{x}_i, \mathbf{x}_j)] = \mathbf{C}^H \mathbf{E}_k \mathbf{C} = \mathbf{C}^H \mathbf{C}.$$

Ezért a 2.8.11 tétel szerint a $\mathbf{G} = \mathbf{C}^H \mathbf{C}$ mátrix pozitív definit, és így $|\mathbf{G}| > 0$. ■

2.8.13 tétel. Minden pozitív definit \mathbf{A} transzformációhoz található olyan pozitív definit \mathbf{B} transzformáció, amelyre $\mathbf{B}^2 = \mathbf{A}$; a \mathbf{B} transzformációt az \mathbf{A} transzformáció négyzetgyökének nevezzük és rá a $\mathbf{B} = \sqrt{\mathbf{A}} = \mathbf{A}^{1/2}$ jelölést használjuk.

Bizonyítás. A 2.8.10 tétel értelmében minden önadjungált transzformációnak van teljes ortogonális sajátvektor-rendszere. Ezt választva bázisnak, az \mathbf{A} transzformáció \mathbf{A} mátrixa diagonálmátrix, és mivel a transzformáció pozitív definit, ezért a mátrix elemei a 2.8.12 tétel alapján a transzformáció pozitív λ_i sajátértékei:

$$\mathbf{A} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

Képezzük most a λ_i sajátértékek pozitív négyzetgyökeiből alkotott diagonálmátrixot:

$$\mathbf{B} = \begin{bmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n} \end{bmatrix}.$$

Nyilvánvaló, hogy $\mathbf{B}^2 = \mathbf{A}$, és mivel a \mathbf{B} diagonálmátrix elemei mind pozitívak, a hozzá tartozó (önadjungált) transzformáció is pozitív definit. ■

Ezek után könnyen bebizonyítható a lineáris transzformációk faktorizációjára vonatkozó alábbi tétel.

2.8.14 tétel. *Az euklideszi tér bármely nonszinguláris lineáris transzformációja előállítható egy pozitív definit (önadjungált) transzformáció és egy unitér transzformáció szorzataként.*

Bizonyítás. Legyen \mathbf{A} tetszőleges, nonszinguláris transzformáció. A 2.8.11 tétel szerint $\mathbf{A}\mathbf{A}^H$ pozitív definit transzformáció, amelynek a 2.8.13 tétel alapján létezik pozitív definit négyzetgyöke:

$$(2.8.16) \quad \mathbf{H} = \sqrt{\mathbf{A}\mathbf{A}^H}.$$

Bebizonyítjuk, hogy a (2.8.16) képlettel definiált \mathbf{H} transzformáció segítségével képzett

$$(2.8.17) \quad \mathbf{U} = \mathbf{H}^{-1}\mathbf{A}$$

transzformáció unitér. Ugyanis figyelembe véve, hogy

$$(\mathbf{H}^{-1}\mathbf{A})^H = \mathbf{A}^H(\mathbf{H}^{-1})^H = \mathbf{A}^H\mathbf{H}^{-1}$$

(mivel önadjungált transzformáció inverze szintén önadjungált),

$$\begin{aligned} \mathbf{U}\mathbf{U}^H &= \mathbf{H}^{-1}\mathbf{A}(\mathbf{H}^{-1}\mathbf{A})^H = \mathbf{H}^{-1}\mathbf{A}\mathbf{A}^H\mathbf{H}^{-1} = \mathbf{H}^{-1}\mathbf{H}^2\mathbf{H}^{-1} = \\ &= (\mathbf{H}^{-1}\mathbf{H})(\mathbf{H}\mathbf{H}^{-1}) = \mathbf{E}. \end{aligned}$$

Így (2.8.17)-ből felírható, hogy $\mathbf{A} = \mathbf{H}\mathbf{U}$, tehát a nonszinguláris \mathbf{A} transzformációt valóban sikerült felbontani a pozitív definit \mathbf{H} és az unitér \mathbf{U} transzformáció szorzatára. ■

A 2.8.3 tételből tudjuk, hogy az önadjungált transzformáció sajátértékei mindig valósak. Mivel két önadjungált transzformáció szorzata általában nem önadjungált, ezért a szorzat sajátértékeiről sem állítható általában, hogy valósak. A szorzat sajátértékeinek valós voltát biztosítja azonban az, ha legalább az egyik tényező pozitív definit. Erre vonatkozik a következő tétel, amelynek a bizonyításához felhasználjuk a pozitív definit transzformációk négyzetgyökének a fogalmát.

2.8.15 tétel. *Legyen \mathbf{A} és \mathbf{B} egy-egy önadjungált transzformáció. Ha \mathbf{A} pozitív definit, akkor az \mathbf{AB} transzformáció sajátértékei valósak.*

Bizonyítás. Mivel \mathbf{A} pozitív definit transzformáció, ezért a 2.8.13 tétel szerint létezik az ugyancsak pozitív definit $\mathbf{A}^{1/2}$ transzformáció. Legyen

e_1, e_2, \dots, e_n az E_n euklideszi tér egy bázisa és jelölje \mathbf{A} és \mathbf{B} az \mathbf{A} , ill. \mathbf{B} transzformáció e bázishoz tartozó mátrixát. Az \mathbf{AB} szorzattranszformációnak e bázishoz tartozó mátrixa \mathbf{AB} . Ha most az \mathbf{AB} mátrixon az $\mathbf{A}^{1/2}$ transzformációhoz tartozó $\mathbf{A}^{1/2}$ mátrixszal hasonlósági transzformációt végzünk – azaz új bázisra térünk át:

$$(\mathbf{A}^{1/2})^{-1}(\mathbf{AB})\mathbf{A}^{1/2},$$

akkor a sajátértékei nem változnak meg. A kapott mátrix hermitikus, ugyanis figyelembe véve, hogy

$$(\mathbf{A}^{1/2})^{-1}\mathbf{A} = (\mathbf{A}^{1/2})^{-1}\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}^{1/2},$$

ezért

$$(\mathbf{A}^{1/2})^{-1}(\mathbf{AB})\mathbf{A}^{1/2} = \mathbf{A}^{1/2}\mathbf{BA}^{1/2}$$

és

$$(\mathbf{A}^{1/2}\mathbf{BA}^{1/2})^H = (\mathbf{A}^{1/2})^H\mathbf{B}^H(\mathbf{A}^{1/2})^H = \mathbf{A}^{1/2}\mathbf{BA}^{1/2}.$$

A kapott mátrix sajátértékei – és így \mathbf{AB} sajátértékei is – valósak. Ezzel a tételt bebizonyítottuk. ■

2.8.5 Főtengelytétel és általánosítása

A 2.8.1 tételben bebizonyítottuk, hogy az euklideszi térben minden hermitikus bilineáris alakhoz egyértelműen hozzárendelhető egy önadjungált lineáris transzformáció. Az alábbi tételben azt mutatjuk meg, hogy minden hermitikus alakhoz található olyan bázis, amelyben az hermitikus alak négyzetösszegként írható fel. Azt a transzformációt, amellyel ezt megkapjuk, főtengety-transzformációnak nevezzük.

2.8.16 tétel (főtengelytétel). *Ha $A(\mathbf{x}; \mathbf{y})$ az n -dimenziós euklideszi térben egy hermitikus bilineáris alak, akkor mindig létezik olyan ortonormált bázis, amelyben a megfelelő $A(\mathbf{x}; \mathbf{y})$ hermitikus alak négyzetösszegként állítható elő:*

$$(2.8.18) \quad A(\mathbf{x}; \mathbf{x}) = \sum_{i=1}^n \lambda_i |\xi_i|^2,$$

ahol a λ_i számok valósak, a ξ_i értékek pedig az \mathbf{x} vektornak erre az ortonormált bázisra vonatkozó koordinátái.

Bizonyítás. Az $A(\mathbf{x}; \mathbf{y})$ hermitikus bilineáris alaknak a 2.8.1 tétel szerint az

$$A(\mathbf{x}; \mathbf{y}) = (A\mathbf{x}, \mathbf{y}) = \overline{A(\mathbf{y}; \mathbf{x})}$$

azonosság alapján megfelel egy önadjungált transzformáció; ennek a 2.8.10 tétel szerint viszont létezik n , páronként ortogonális sajátvektora, jelölje ezeket $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$. Válasszuk a sajátvektorokat bázisként és legyenek az \mathbf{x} , ill. \mathbf{y} vektor koordinátái erre a bázisra vonatkozóan $\xi_1, \xi_2, \dots, \xi_n$ ill. $\eta_1, \eta_2, \dots, \eta_n$:

$$\mathbf{x} = \sum_{i=1}^n \xi_i \mathbf{e}_i, \quad \mathbf{y} = \sum_{k=1}^n \eta_k \mathbf{e}_k.$$

Mivel az \mathbf{e}_i vektorok sajátvektorok, ezért $\mathbf{A}\mathbf{e}_i = \lambda_i \mathbf{e}_i$, és mivel a sajátvektorok ortogonálisak egymásra, ezért $(\mathbf{e}_i, \mathbf{e}_k) = \delta_{ik}$, ahol δ_{ik} a Kronecker-féle szimbólum. Behelyettesítve ezeket az $A(\mathbf{x}; \mathbf{y})$ hermitikus alakba,

$$\begin{aligned} A(\mathbf{x}; \mathbf{y}) &= (\mathbf{A}\mathbf{x}, \mathbf{y}) = \left(\sum_{i=1}^n \xi_i \mathbf{A}\mathbf{e}_i, \sum_{k=1}^n \eta_k \mathbf{e}_k \right) = \\ &= \left(\sum_{i=1}^n \lambda_i \xi_i \mathbf{e}_i, \sum_{k=1}^n \eta_k \mathbf{e}_k \right) = \sum_{i=1}^n \lambda_i \xi_i \bar{\eta}_i, \end{aligned}$$

ahonnan $\mathbf{x} = \mathbf{y}$ esetén azt kapjuk, hogy

$$(2.8.19) \quad A(\mathbf{x}; \mathbf{x}) = \sum_{i=1}^n \lambda_i |\xi_i|^2,$$

amit bizonyítani kellett. ■

Az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ sajátvektorok irányát az hermitikus alakhoz tartozó *főtengelyeknek* nevezzük.

A főtengelytétel általánosításaként most azt vizsgáljuk meg, hogy két adott hermitikus alakhoz található-e és ha igen, milyen feltételek mellett, olyan bázis, amelyben mindkét hermitikus alak négyzetösszegbe megy át. Erre vonatkozik a következő tétel.

2.8.17 tétel. *Legyen $A(\mathbf{x}; \mathbf{x})$ és $B(\mathbf{x}; \mathbf{x})$ az n -dimenziós lineáris tér egy-egy hermitikus alakja és legyen $B(\mathbf{x}; \mathbf{x})$ pozitív definit. Ekkor a térnek mindig van olyan bázisa, amelyben mindkét hermitikus alak négyzetösszegként állítható elő.*

Bizonyítás. Mivel $B(\mathbf{x}; \mathbf{x})$ pozitív definit hermitikus alak, az adott lineáris térben definiálható a skaláris szorzat az

$$(2.8.20) \quad (\mathbf{x}, \mathbf{y}) = B(\mathbf{x}; \mathbf{y})$$

előírással. Ezáltal az adott n -dimenziós lineáris teret euklideszi térré tettük. A 2.8.16 tétel szerint létezik tehát olyan – a (2.8.20) skalárszorzatra

nézve ortogonális – $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázis, amelyben az $A(\mathbf{x}; \mathbf{x})$ hermitikus alak négyzetösszege redukálódik:

$$(2.8.21) \quad A(\mathbf{x}; \mathbf{x}) = \lambda_1 |\xi_1|^2 + \lambda_2 |\xi_2|^2 + \dots + \lambda_n |\xi_n|^2.$$

Mivel ortonormált bázisban

$$(\mathbf{x}, \mathbf{x}) = |\xi_1|^2 + |\xi_2|^2 + \dots + |\xi_n|^2,$$

ezért esetünkben (2.8.20) alapján

$$(2.8.22) \quad B(\mathbf{x}; \mathbf{x}) = |\xi_1|^2 + \dots + |\xi_n|^2.$$

Mivel az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ vektorok az \mathbf{R}_n térben is bázist alkotnak, így belátuk, hogy az \mathbf{R}_n térnek van olyan bázisa, amelyben az $A(\mathbf{x}; \mathbf{x})$ és a $B(\mathbf{x}; \mathbf{x})$ hermitikus alak egyidejűleg redukálódik négyzetösszeggé; ezzel a tételt bebizonyítottuk. ■

Határozzuk meg a (2.8.21) négyzetösszegben szereplő λ_i együtthatókat. Mivel az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisban az $A(\mathbf{x}; \mathbf{x})$ és a $B(\mathbf{x}; \mathbf{x})$ hermitikus alak egyidejűleg négyzetösszegként írható fel, a (2.8.21) és a (2.8.22) kifejezések alapján ebben a bázisban a fenti hermitikus alakok mátrixa diagonálmátrix:

$$(2.8.23) \quad \mathbf{A} = \begin{bmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \lambda_3 & & \\ & & & \ddots & \\ & & & & \lambda_n \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}.$$

Tetszőleges $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \dots, \tilde{\mathbf{e}}_n$ bázisra áttérve, az hermitikus alakok mátrixa az

$$\tilde{\mathbf{e}}_k = \sum_{i=1}^n \mathbf{e}_i c_{ik}$$

bázistranszformáció $\mathbf{C} = [c_{ik}]$ mátrixával képzett kongruens transzformációval (lásd a 2.5.2 pontot) transzformálódik, tehát ha $\tilde{\mathbf{A}}$ és $\tilde{\mathbf{B}}$ jelöli a megfelelő hermitikus alakok mátrixát az új bázisban, akkor

$$\tilde{\mathbf{A}} = \mathbf{C}^H \mathbf{A} \mathbf{C}, \quad \tilde{\mathbf{B}} = \mathbf{C}^H \mathbf{B} \mathbf{C} = \mathbf{C}^H \mathbf{C}.$$

A keresett λ_i értékek a $|\lambda \mathbf{B} - \mathbf{A}|$ polinom gyökei, ugyanis \mathbf{A} és \mathbf{B} (2.8.23) alatti kifejezését figyelembe véve,

$$(2.8.24) \quad |\lambda \mathbf{B} - \mathbf{A}| = \prod_{k=1}^n (\lambda - \lambda_k).$$

Az új bázisra áttérve, a megfelelő $|\lambda\tilde{\mathbf{B}} - \tilde{\mathbf{A}}|$ determináns csak egy állandó szorzóban tér el a (2.8.24) determinánstól. Ugyanis

$$\begin{aligned} |\lambda\tilde{\mathbf{B}} - \tilde{\mathbf{A}}| &= |\lambda\mathbf{C}^H\mathbf{B}\mathbf{C} - \mathbf{C}^H\mathbf{A}\mathbf{C}| = |\mathbf{C}^H| \cdot |\lambda\mathbf{B} - \mathbf{A}| \cdot |\mathbf{C}| = \\ &= |\mathbf{C}^H| |\mathbf{C}| \prod_{k=1}^n (\lambda - \lambda_k) = |\tilde{\mathbf{B}}| \prod_{k=1}^n (\lambda - \lambda_k), \end{aligned}$$

tehát a keresett λ_i értékeket a

$$(2.8.25) \quad |\lambda\tilde{\mathbf{B}} - \tilde{\mathbf{A}}| = 0$$

egyenlet gyökei szolgáltatják, ahol $\tilde{\mathbf{A}}$ és $\tilde{\mathbf{B}}$ az $A(\mathbf{x}; \mathbf{x})$ és $B(\mathbf{x}; \mathbf{x})$ hermitikus alakok mátrixa tetszőleges bázisban. A 2.8.17 tételből következik, hogy ha $B(\mathbf{x}; \mathbf{x})$ pozitív definit hermitikus alak, akkor a (2.8.25) egyenletnek csupa valós gyöke van és mindig található olyan – a (2.8.20) skaláris szorzatra nézve ortogonális – bázis, amelyben az $A(\mathbf{x}; \mathbf{x})$ hermitikus alak mátrixa a λ_i valós számokból alkotott diagonálmátrix, a $B(\mathbf{x}; \mathbf{x})$ pozitív definit hermitikus alak mátrixa pedig egység mátrix.

Tekintettel arra, hogy az euklideszi térben minden hermitikus alakhoz egyértelműen hozzárendelhető egy önadjungált lineáris transzformáció, a két hermitikus alak egyidejű négyzetösszegre redukálása ekvivalens a lineáris transzformációk sajátérték-feladatának általánosításával.

2.8.6 definíció. Legyen A és B az n -dimenziós euklideszi tér két lineáris transzformációja. Azokat a λ értékeket, amelyekre létezik olyan $\mathbf{u} \neq 0$ vektor, hogy

$$(2.8.26) \quad A\mathbf{u} = \lambda B\mathbf{u}$$

teljesül, az A , B transzformációpár általánosított sajátértékeinek, az egyenletet kielégítő \mathbf{u} vektorokat pedig (a λ általánosított sajátértékhez tartozó) általánosított sajátvektorainak nevezzük. Legyen egy tetszőleges bázisban az A , ill. B transzformáció mátrixa \mathbf{A} , ill. \mathbf{B} ; ekkor a $\lambda\mathbf{B} - \mathbf{A}$ mátrixot az \mathbf{A} , \mathbf{B} mátrixpár általánosított karakterisztikus mátrixának, a $D_n(\lambda) = |\lambda\mathbf{B} - \mathbf{A}|$ polinomot általánosított karakterisztikus polinomnak, a

$$(2.8.27) \quad |\lambda\mathbf{B} - \mathbf{A}| = 0$$

egyenletet pedig az A , B transzformációpár (vagy az \mathbf{A} , \mathbf{B} mátrixpár) általánosított karakterisztikus egyenletének nevezzük.

Könnyen belátható, hogy az általánosított karakterisztikus polinom – a karakterisztikus polinomhoz hasonlóan (lásd a 2.6.3 definíciót) – a bázis transzformációjával szemben invariáns; ugyanis – tekintettel arra, hogy bármely \mathbf{C}

nemszinguláris mátrixra $\mathbf{C}\mathbf{C}^{-1} = \mathbf{E}$ és ezért determinánsukra $|\mathbf{C}| |\mathbf{C}^{-1}| = 1$,
– felírható, hogy

$$|\lambda \mathbf{B} - \mathbf{A}| = |\mathbf{C}^{-1}| \cdot |\lambda \mathbf{B} - \mathbf{A}| \cdot |\mathbf{C}| = |\lambda \mathbf{C}^{-1} \mathbf{B} \mathbf{C} - \mathbf{C}^{-1} \mathbf{A} \mathbf{C}|,$$

ahol a jobb oldal éppen a transzformált bázisban felírt általánosított karakterisztikus polinom.

2.8.7 definíció. Legyen \mathbf{B} az n -dimenziós euklideszi tér pozitív definit önadjungált transzformációja. Akkor a $(\mathbf{B}\mathbf{x}, \mathbf{x}) = B(\mathbf{x}; \mathbf{x})$ összefüggéssel definiált hermitikus alak pozitív definit s így van standard bázisa. Az $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ vektorrendszert **B-unitérnek** nevezzük, ha standard bázisa a $B(\mathbf{x}; \mathbf{x})$ pozitív definit hermitikus alaknak.

Legyen most a \mathbf{B} transzformációnak – s egyben a $B(\mathbf{x}; \mathbf{x})$ hermitikus alaknak, ill. a $B(\mathbf{x}; \mathbf{y})$ bilineáris alaknak – a tér valamely $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ ortonormált bázisára vonatkozó mátrixa az hermitikus \mathbf{B} mátrix, és a \mathbf{B} -unitér vektorrendszer \mathbf{u}_k vektorának koordinátavektora \mathbf{u}_k ($k = 1, 2, \dots, n$). Jelölje \mathbf{U} azt a mátrixot, amelynek k -adik oszlopa \mathbf{u}_k . Ha most az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisról áttérünk az $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ bázisra, akkor a $B(\mathbf{x}; \mathbf{y})$ bilineáris alak mátrixa – lévén az új bázis standard – az egységmátrix lesz, s ezt az \mathbf{U} mátrixszal végzett kongruens transzformációval nyerjük:

$$(2.8.28) \quad \mathbf{U}^H \mathbf{B} \mathbf{U} = \mathbf{E},$$

vagy elemenként felírva:

$$\mathbf{u}_j^H \mathbf{B} \mathbf{u}_i = \delta_{ij} \quad (i, j = 1, 2, \dots, n).$$

Ez pedig felfogható, mint a

$$(\mathbf{B} \mathbf{u}_i, \mathbf{u}_j) = \delta_{ij} \quad (i, j = 1, 2, \dots, n)$$

összefüggéseknek az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisban felírt alakja. (Ez indokolja az $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ rendszerre a *B-ortogonális* elnevezést is.)

A 2.8.5–2.8.7 definíciók segítségével megfogalmazzuk a 2.8.17 tételt lineáris transzformációkra.

2.8.18 tétel. Legyen \mathbf{A} és \mathbf{B} az n -dimenziós euklideszi tér egy-egy önadjungált transzformációja. Ha \mathbf{B} pozitív definit transzformáció, akkor az \mathbf{A}, \mathbf{B} transzformációpár λ_k általánosított sajátértékei valósak, a hozzájuk tartozó \mathbf{u}_k általánosított sajátvektorok pedig (teljes) \mathbf{B} -unitér vektorrendszert alkotnak.

Bizonyítás. Legyen az \mathbf{A} , ill. \mathbf{B} transzformációnak valamely $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ ortonormált bázisra vonatkozó mátrixa \mathbf{A} , ill. \mathbf{B} , és a sajátvektorok koor-

dinátavektoraiból alkotott mátrix \mathbf{U} . Ekkor az \mathbf{U} mátrixszal végzett kongruens transzformáció az \mathbf{A} mátrixot a sajátértékekből álló $\mathbf{\Lambda} = \langle \lambda_k \rangle$ diagonál-mátrixba viszi át:

$$(2.8.29) \quad \mathbf{U}^H \mathbf{A} \mathbf{U} = \mathbf{\Lambda}.$$

A (2.8.29) összefüggés közvetlenül adódik az általánosított sajátvektorokat definiáló (2.8.26) egyenletből, ha figyelembe vesszük a (2.8.28) összefüggést. A (2.8.26) egyenlet ugyanis az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisban az \mathbf{A} , \mathbf{B} , ill. \mathbf{U} mátrix segítségével

$$(2.8.30) \quad \mathbf{A} \mathbf{U} = \mathbf{B} \mathbf{U} \mathbf{\Lambda}$$

alakban írható. Balról megszorozva a (2.8.30) egyenlet az \mathbf{U}^H mátrixszal, (2.8.28) behelyettesítésével valóban a (2.8.29) összefüggést kapjuk. ■

2.8.6 Sajátértékek extrémális tulajdonsága

Most megmutatjuk az önadjungált transzformációk sajátértékeinek egy nevezetes extrémális tulajdonságát. Ezzel a sajátértékek sajátvektorok létezésének egy másik bizonyítását adjuk, ami egyúttal lehetővé teszi a sajátértékek kiszámítását egy minimumfeladat megoldása útján. Ehhez szükségünk lesz a következő tételre:

2.8.19 tétel. *Ha \mathbf{A} a (komplex) euklideszi tér egy pozitív szemidefinit önadjungált transzformációja, vagyis bármely $\mathbf{x} \neq 0$ vektorra*

$$(2.8.31) \quad (\mathbf{A} \mathbf{x}, \mathbf{x}) \geq 0,$$

akkor ebben az egyenlőtlenségben az egyenlőség csak olyan $\mathbf{x} = \mathbf{u}$ vektorra teljesülhet, amelyre

$$\mathbf{A} \mathbf{u} = 0.$$

Ez más szavakkal annyit jelent, hogy ha valamely \mathbf{u} vektorra a (2.8.31) összefüggésben az egyenlőségjel érvényes, akkor \mathbf{u} az \mathbf{A} transzformáció nullterében (magterében) van.

Bizonyítás. Tegyük fel, hogy $(\mathbf{A} \mathbf{u}, \mathbf{u}) = 0$ és legyen t tetszőleges valós szám, \mathbf{y} pedig a tér tetszőleges vektora. A feltétel értelmében

$$(\mathbf{A}(\mathbf{u} + t\mathbf{y}), \mathbf{u} + t\mathbf{y}) \geq 0,$$

vagy részletesen kifejezve:

$$(\mathbf{A} \mathbf{u}, \mathbf{u}) + t[(\mathbf{A} \mathbf{u}, \mathbf{y}) + (\mathbf{A} \mathbf{y}, \mathbf{u})] + t^2(\mathbf{A} \mathbf{y}, \mathbf{y}) \geq 0.$$

Mivel $(\mathbf{A}\mathbf{u}, \mathbf{u}) = 0$, a kapott kifejezés csak akkor lehet a t változó minden értékére nemnegatív, ha t együtthatója zérus, azaz

$$(2.8.32) \quad (\mathbf{A}\mathbf{u}, \mathbf{y}) + (\mathbf{A}\mathbf{y}, \mathbf{u}) = 0.$$

De $(\mathbf{A}\mathbf{y}, \mathbf{u}) = (\mathbf{y}, \mathbf{A}\mathbf{u}) = \overline{(\mathbf{A}\mathbf{u}, \mathbf{y})}$. Így a (2.8.32) feltétel ekvivalens a

$$(2.8.33) \quad \operatorname{Re}(\mathbf{A}\mathbf{u}, \mathbf{y}) = 0$$

feltétellel, és mivel ez az azonosság minden \mathbf{y} vektorra érvényes, ezért $\mathbf{A}\mathbf{u} = 0$, amit bizonyítani kellett. ■

Ezek után bebizonyítjuk az önadjungált transzformációk sajátértékeinek extrémális tulajdonságára vonatkozó következő fontos tételt.

2.8.20 tétel. *Bármely \mathbf{A} önadjungált transzformációhoz rendelt $(\mathbf{A}\mathbf{x}, \mathbf{x})$ kvadrátikus alak az $(\mathbf{x}, \mathbf{x}) = 1$ feltételt kielégítő \mathbf{x} vektorok halmazán felveszi minimumát. Ha \mathbf{u}_1 minimumhely, akkor az \mathbf{u}_1 vektor az \mathbf{A} transzformáció λ_1 legkisebb sajátértékéhez tartozó sajátvektor, a minimum értéke pedig éppen λ_1 .*

Megjegyzés. A geometriai térből vett szemlélet alapján az adott feltételt kielégítő vektorok halmazát „egységgömbnek” nevezhetjük.

Bizonyítás. Az analízis elemeiből következik, hogy az egységgömb az n -dimenziós tér korlátos, zárt halmaza és $(\mathbf{A}\mathbf{x}, \mathbf{x})$ ezen a halmazon folytonos függvény, ezért valamilyen \mathbf{u}_1 helyen felveszi minimumát. Jelölje λ_1 a minimum értékét; tehát

$$(2.8.34) \quad (\mathbf{A}\mathbf{x}, \mathbf{x}) \geq \lambda_1, \quad \text{ha} \quad (\mathbf{x}, \mathbf{x}) = 1,$$

és

$$(\mathbf{A}\mathbf{u}_1, \mathbf{u}_1) = \lambda_1, \quad \text{ahol} \quad (\mathbf{u}_1, \mathbf{u}_1) = 1.$$

A (2.8.34) egyenlőtlenség az alábbi alakban is írható:

$$(2.8.35) \quad (\mathbf{A}\mathbf{x}, \mathbf{x}) \geq \lambda_1(\mathbf{x}, \mathbf{x}).$$

(Ez az egyenlőtlenség tetszőleges hosszúságú \mathbf{x} vektorokra érvényes, mivel mindkét oldala szorozható az \mathbf{x} vektor abszolút értékének négyzetével.)

A (2.8.35) egyenlőtlenséget hozzuk a következő alakra:

$$(\mathbf{A}\mathbf{x} - \lambda_1\mathbf{x}, \mathbf{x}) = ((\mathbf{A} - \lambda_1\mathbf{E})\mathbf{x}, \mathbf{x}) \geq 0.$$

A mondottak értelmében $\mathbf{x} = \mathbf{u}_1$ esetén az egyenlőségjel érvényes:

$$((\mathbf{A} - \lambda_1\mathbf{E})\mathbf{u}_1, \mathbf{u}_1) = 0.$$

Az $A - \lambda_1 E$ transzformáció tehát kielégíti a 2.8.19. tétel feltételeit, ezért $(A - \lambda_1 E)\mathbf{u}_1 = \mathbf{0}$, azaz $A\mathbf{u}_1 = \lambda_1 \mathbf{u}_1$. Tehát \mathbf{u}_1 az A transzformáció λ_1 sajátértékéhez tartozó sajátvektor. A λ_1 sajátérték egyúttal az A transzformáció legkisebb sajátértéke. Ha ugyanis az A transzformációnak lenne egy $\lambda_0 < \lambda_1$ sajátértéke, akkor a hozzá tartozó \mathbf{u}_0 sajátvektorra fennállna a következő összefüggés:

$$(A\mathbf{u}_0, \mathbf{u}_0) = (\lambda_0 \mathbf{u}_0, \mathbf{u}_0) = \lambda_0,$$

tehát az $(A\mathbf{x}, \mathbf{x})$ kvadratikus alaknak az \mathbf{u}_0 helyen felvett értéke – feltevésünkkel ellentétben – kisebb lenne, mint λ_1 . Ezzel a tételt bebizonyítottuk. ■

A következő sajátérték ugyanazzal a gondolatmenettel határozható meg, ahogy a 2.8.10 tételt bizonyítottuk. Az E térnek az \mathbf{u}_1 sajátvektorra ortogonális vektorai az A transzformációra nézve invariáns $(n - 1)$ -dimenziós alteret alkotnak. Ha tehát ebben az altérben keressük az $(A\mathbf{x}, \mathbf{x})$ kvadratikus alaknak az $(\mathbf{x}, \mathbf{x}) = 1$ feltétel melletti minimumát, akkor egy λ_2 sajátértéket és hozzá tartozó \mathbf{u}_2 sajátvektort kapunk. Mivel $(A\mathbf{x}, \mathbf{x})$ minimuma az egész térben nem lehet nagyobb, mint bármely altérben, ezért nyilvánvalóan $\lambda_2 \geq \lambda_1$. További sajátértéket kapunk, ha ugyanezt a feltételes minimumfeladatot az \mathbf{u}_1 és \mathbf{u}_2 sajátvektorokra ortogonális, $(n - 2)$ -dimenziós invariáns altérben oldjuk meg. Az eljárást folytatva, meghatározható a transzformáció valamennyi sajátértéke és sajátvektora.

Az önadjungált A transzformáció sajátértékei, amelyek valós számok, ezzel a módszerrel nagyság szerinti sorrendben adódnak:

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n.$$

A (2.8.35) egyenlőtlenség alapján a legkisebb sajátérték az alábbi minimumfeladat megoldásaként nyerhető:

$$(2.8.36) \quad \lambda_1 = \min_x \frac{(A\mathbf{x}, \mathbf{x})}{(\mathbf{x}, \mathbf{x})},$$

ahol \mathbf{x} befutja a tér összes vektorainak halmazát.

Egészen hasonló megfontolással belátható, hogy a legnagyobb sajátértéket a megfelelő maximumfeladat megoldása adja:

$$(2.8.37) \quad \lambda_n = \max_x \frac{(A\mathbf{x}, \mathbf{x})}{(\mathbf{x}, \mathbf{x})}.$$

Látható, hogy az itt szereplő

$$(2.8.38) \quad R = \frac{(A\mathbf{x}, \mathbf{x})}{(\mathbf{x}, \mathbf{x})}$$

hányados szerepe jelentős a sajátértékek extrémális tulajdonságainak a vizsgálatában. Célszerű ezért néhány új fogalom bevezetése.

2.8.8 definíció. *Euklideszi térben értelmezett tetszőleges, önadjungált A transzformációhoz és a tér tetszőleges \mathbf{x} vektorához hozzárendelt (2.8.38) értéket az A transzformáció \mathbf{x} vektorhoz tartozó Rayleigh-féle* hányadosának nevezzük.*

2.8.9 definíció. *Az A önadjungált transzformáció Rayleigh-féle hányadosa által felvett értékek összességét, mialatt az \mathbf{x} vektor az egész teret befutja, az A transzformáció értékkészletének nevezzük és $[A]$ -val jelöljük.*

Az önadjungált transzformációhoz rendelt hermitikus alakok csak valós értékeket vehetnek fel; ebből következik, hogy minden önadjungált transzformáció értékkészlete a valós számok egy részhalmaza, s minthogy a Rayleigh-féle hányados folytonos, így legnagyobb és legkisebb értéke között minden értéket felvesz, vagyis ez a halmaz a számegyenes egy zárt intervalluma. Az imént bebizonyított 2.8.20 tételből pedig az következik, hogy az önadjungált transzformáció sajátértékei ebben az intervallumban helyezkednek el.

A 2.8.17 tétel meghatározza, hogy két hermitikus alak milyen feltételek mellett redukálható egyidejűleg négyzetösszeggé (ha legalább egyikük pozitív definit). Ennek kapcsán bevezettük az A, B transzformációpár általánosított sajátértékeinek és sajátvektorainak fogalmát. Az ott követett gondolatmenet alkalmazásával és a sajátértékek most megismert extrémális tulajdonságainak ismeretében eredményeink könnyen kiterjeszthetők az általánosított sajátértékek esetére. Ha ismét feltesszük, hogy A és B önadjungált lineáris transzformációk, és B ezen felül pozitív definit, akkor a skaláris szorzatot (2.8.20) szerint a

$$B(\mathbf{x}; \mathbf{y}) = (B\mathbf{x}, \mathbf{y})$$

előírással értelmezve, a legkisebb általánosított sajátértéket ismét az $(A\mathbf{x}, \mathbf{x})$ hermitikus alak minimuma fogja szolgáltatni, ezúttal azonban a $(B\mathbf{x}, \mathbf{x}) = 1$ mellékfeltétel mellett. Ha tehát az A, B transzformációpárra a Rayleigh-féle hányadost az

$$(2.8.39) \quad R = \frac{(A\mathbf{x}, \mathbf{x})}{(B\mathbf{x}, \mathbf{x})}$$

összefüggéssel definiáljuk, akkor – mivel a 2.8.18 tételben már bebizonyítottuk, hogy az általánosított sajátértékek valósak – a legnagyobb, ill. legkisebb általánosított sajátértéket a (2.8.39) Rayleigh-féle hányados maximuma,

*J. W. Rayleigh (1842–1919) angol matematikus.

ill. minimuma szolgáltatja:

$$\lambda_1 = \min_x \frac{(Ax, x)}{(Bx, x)}, \quad \lambda_n = \max_x \frac{(Ax, x)}{(Bx, x)}.$$

A Rayleigh-féle hányados által felvett értékek összességét – amely ebben az esetben is a valós számegyenes egy zárt intervallumát fedi le – az A, B *transzformációpár értékkészletének* fogjuk nevezni.

Megjegyzés. Az önadjungált transzformációk, ill. transzformációpárok értékkészletének fogalma kiterjeszthető általánosabb transzformációtípusokra is, ez azonban általában nem a valós számegyenes egy intervalluma, hanem a komplex számsík egy tartománya lesz.

2.9 LINEÁRIS TRANSZFORMÁCIÓK A VALÓS LINEÁRIS TÉRBEN

Ebben a szakaszban azt vizsgáljuk, hogy az eddig megismert főbb fogalmak és tételek hogyan módosulnak abban a speciális esetben, amikor a valós tér lineáris transzformációit tekintjük. Mint majd látjuk, a valós térben nem minden lineáris transzformációnak létezik valós sajátértéke és így hozzá tartozó sajátvektora, tehát nem mindig létezik egydimenziós invariáns altére. Megmutatjuk azonban, hogy a valós tér minden lineáris transzformációjának van egy- vagy kétdimenziós invariáns altére. Tárgyaljuk a komplex térben diagonalizálható transzformációk normálalakját a valós térben, majd részletesebben ismertetjük a szimmetrikus és az ortogonális transzformációk főbb tulajdonságait, és megadjuk ezek mátrixának kanonikus alakját.

Végül néhány geometriai feladaton szemléltetjük a kvadrátikus alakokra vonatkozó fontosabb megállapításainkat.

2.9.1 Lineáris transzformáció normálalakja

2.9.1 tétel. Az n -dimenziós \mathbf{R} valós lineáris tér minden A lineáris transzformációjának van egydimenziós vagy kétdimenziós invariáns altére.

Bizonyítás. Legyen \mathbf{A} az \mathbf{R} tér A transzformációjának mátrixa az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisban. Mivel \mathbf{R} valós, tehát a skalárral való szorzás csak valós számokra van értelmezve, ezért a koordináták minden bázisban csak valósak lehetnek. A 2.6.1 tétel bizonyításának gondolatmenetét követve, a sajátértékekre a

$$(2.9.1) \quad |\lambda \mathbf{E} - \mathbf{A}| = 0$$

karakterisztikus egyenletet kapjuk. Ennek gyökeitől függően, két esetet kell megkülönböztetnünk.

(1) Ha λ_0 a karakterisztikus egyenlet valós gyöke, akkor ezt behelyettesítve a sajátvektorokat meghatározó

$$(2.9.2) \quad (\lambda_0 \mathbf{E} - \mathbf{A})\mathbf{u} = \mathbf{0}$$

egyenletbe, ennek van nemtriviális valós megoldása, amely \mathbf{R} egy vektorának az $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisra vonatkozó koordinátáiból áll; tehát ebben az esetben van egydimenziós invariáns altér.

(2) Ha λ_0 a (2.9.1) karakterisztikus egyenlet komplex gyöke:

$$\lambda_0 = \alpha + i\beta,$$

akkor ezt a (2.9.2) egyenletbe behelyettesítve, annak megoldásaként komplex számokat kapunk; ezek nem határozzák meg a valós euklideszi tér egyetlen vektorát sem, vagyis ehhez a λ_0 értékhez nem található egydimenziós invariáns altér. A kapott megoldás viszont felírható $\mathbf{u} + i\mathbf{v}$ alakban, ahol \mathbf{u} és \mathbf{v} elemei valós számok:

$$\{(\alpha + i\beta)\mathbf{E} - \mathbf{A}\}(\mathbf{u} + i\mathbf{v}) = \mathbf{0},$$

illetve beszorzás után

$$\alpha\mathbf{u} - \beta\mathbf{v} - \mathbf{A}\mathbf{u} + i\{\alpha\mathbf{v} + \beta\mathbf{u} - \mathbf{A}\mathbf{v}\} = \mathbf{0}.$$

Ebből következik, hogy a bal oldalon mind a valós résznek, mind a képzetes résznek zérusnak kell lennie, vagyis

$$(2.9.3) \quad \mathbf{A}\mathbf{u} = \alpha\mathbf{u} - \beta\mathbf{v}, \quad \mathbf{A}\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{u}.$$

Ez azt jelenti, hogy az \mathbf{u} és \mathbf{v} valós vektorok által generált kétdimenziós altér invariáns az \mathbf{A} transzformációra nézve, hiszen (2.9.3)-ból következik, hogy \mathbf{u} és \mathbf{v} tetszőleges lineáris kombinációjára alkalmazva az \mathbf{A} transzformációt, ismét az \mathbf{u} és \mathbf{v} valamely lineáris kombinációját kapjuk. ■

A 2.6 szakaszban láttuk, hogy ha a komplex lineáris térben a transzformáció minden sajátértékéhez annyi lineárisan független sajátvektor tartozik, amennyi a sajátérték multiplicitása, akkor a transzformáció mátrixa egyszerű struktúrájú, vagyis valamilyen bázisban diagonálmátrix. Természetesen a sajátértékek és a diagonálmátrix elemei általában nem valós számok.

A valós lineáris tér egy lineáris transzformációja csak akkor diagonalizálható, ha karakterisztikus egyenletének csupa valós gyöke van és ezek mind egyikéhez annyi lineárisan független sajátvektor tartozik, amennyi a gyök multiplicitása. Ha viszont a karakterisztikus polinomnak van komplex gyöke (amelynek ekkor konjugáltja is gyök), akkor a transzformáció nem diagonalizálható. A 2.9.1 tétel szerint az ilyen konjugált komplex gyökpárhoz a transzformáció kétdimenziós invariáns altère tartozik.

2.9.2 tétel. *Legyen \mathbf{A} az n -dimenziós valós lineáris tér egy transzformációja, amelyhez tartozó karakterisztikus egyenlet gyökei*

$$\lambda_{2k-1} = \alpha_k + i\beta_k, \quad \lambda_{2k} = \alpha_k - i\beta_k, \quad \beta_k \neq 0 \quad (k = 1, 2, \dots, q),$$

$$\lambda_l = \alpha_l \quad (l = 2q + 1, \dots, n),$$

ahol az α és β értékek valósak. Ekkor létezik olyan bázis, amelyben a transzformáció mátrixa az alábbi normálalakot veszi fel:

$$(2.9.4) \quad \begin{bmatrix} \alpha_1 & -\beta_1 & & & & \\ \beta_1 & \alpha_1 & & & & \\ & & \alpha_2 & -\beta_2 & & \\ & & \beta_2 & \alpha_2 & & \\ & & & & \ddots & \\ & & & & & \alpha_q & -\beta_q \\ & & & & & \beta_q & \alpha_q \\ & & & & & & & \alpha_{2q+1} & & \\ & & & & & & & & \alpha_{2q+2} & & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & \alpha_n \end{bmatrix}.$$

Bizonyítás. Helyettesítsük az $\mathbf{Aw} = \lambda\mathbf{w}$ egyenletrendszerben λ helyébe a λ_{2k-1} és λ_{2k} értékeket; az ekkor kapott \mathbf{w}_k (triviálistól különböző) megoldások $k = 1, \dots, q$ esetén konjugált komplex elemű vektorok. Jelölje ezeket

$$(2.9.5) \quad \mathbf{w}_{2k-1} = \mathbf{u}_k + i\mathbf{v}_k, \quad \mathbf{w}_{2k} = \mathbf{u}_k - i\mathbf{v}_k, \quad (k = 1, 2, \dots, q).$$

A λ helyébe a λ_l értékeket írva, ezek sajátértékek, és a hozzájuk tartozó $\mathbf{w}_l = \mathbf{u}_l$ ($l = 2q+1, 2q+2, \dots, n$) megoldások sajátvektorok. Ha most a valós euklideszi tér bázisául az így nyert

$$(2.9.6) \quad \mathbf{v}_1, \mathbf{u}_1, \mathbf{v}_2, \mathbf{u}_2, \dots, \mathbf{v}_q, \mathbf{u}_q, \mathbf{u}_{2q+1}, \dots, \mathbf{u}_n$$

vektorokat választjuk, akkor ezek segítségével (2.9.3) alapján

$$\mathbf{Av}_k = \alpha_k \mathbf{v}_k + \beta_k \mathbf{u}_k \quad (k = 1, 2, \dots, q),$$

$$\mathbf{Au}_k = -\beta_k \mathbf{v}_k + \alpha_k \mathbf{u}_k,$$

$$\mathbf{Au}_l = \alpha_l \mathbf{u}_l \quad (l = 2q+1, \dots, n)$$

írható, és így a (2.9.6) bázisban valóban megkapjuk a (2.9.4) normálalakot. ■

Megjegyzés. A tételt itt csak arra az esetre mondtuk ki, ha a transzformáció karakterisztikus egyenletének valamennyi gyöke egyszeres. Ugyanezzel a gondolatmenettel bebizonyítható azonban, hogy többszörös gyökök esetén is érvényes a tétel, feltéve, hogy minden gyökhöz az $\mathbf{Aw} = \lambda\mathbf{w}$ egyenletrendszernek annyi lineárisan független megoldása van, amennyi a gyök multiplicitása.

A (2.9.4) normálalakot *kvázidiagonál* alakú mátrixnak nevezzük, és röviden a következőképpen jelöljük:

$$\left\langle \begin{bmatrix} \alpha_1 & -\beta_1 \\ \beta_1 & \alpha_1 \end{bmatrix} \begin{bmatrix} \alpha_2 & -\beta_2 \\ \beta_2 & \alpha_2 \end{bmatrix} \cdots \begin{bmatrix} \alpha_q & -\beta_q \\ \beta_q & \alpha_q \end{bmatrix} \alpha_{2q+1} \quad \alpha_{2q+2} \cdots \alpha_n \right\rangle.$$

Abban a speciális esetben, amikor a karakterisztikus egyenlet gyökei mind a képzetes tengelyen helyezkednek el, vagyis mindegyik α_i zérus, a transzformáció mátrixa a következő normálalakra hozható:

$$(2.9.7) \quad \left\langle \begin{bmatrix} 0 & -\beta_1 \\ \beta_1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -\beta_2 \\ \beta_2 & 0 \end{bmatrix} \cdots \begin{bmatrix} 0 & -\beta_q \\ \beta_q & 0 \end{bmatrix} 0 \dots 0 \right\rangle.$$

A 3.1.8 tételben majd bebizonyítjuk, hogy ilyenek pl. a ferdén szimmetrikus mátrixok, ezek tehát a (2.9.7) normálalakra hozhatók.

2.9.2 Szimmetrikus és ortogonális transzformációk

A következőkben a valós euklideszi térben értelmezett önadjungált és unitér transzformációkkal foglalkozunk. Az önadjungált transzformáció 2.8.1 definíciójából és a 2.8.1 tételből következik, hogy a valós tér önadjungált transzformációinak mátrixa ortonormált bázisban mindig valós szimmetrikus mátrix. Ez indokolja a következő elnevezés bevezetését.

2.9.1 definíció. *A valós euklideszi térben értelmezett önadjungált transzformációt **szimmetrikus transzformációnak** nevezzük.*

A 2.8.3 tétel alapján a szimmetrikus transzformációk karakterisztikus egyenletének valamennyi gyöke valós, a 2.9.1 tételből következik tehát, hogy mindegyik gyök sajátérték és mindegyikhez tartozik (legalább egy) sajátvektor. A 2.8.10 tétel bizonyításában követett gondolatmenet segítségével az is belátható, hogy az n -dimenziós valós euklideszi tér szimmetrikus transzformációjának van n egymásra páronként ortogonális sajátvektora. (Itt is azt kell először belátni, hogy az \mathbf{e}_1 sajátvektorra ortogonális vektorokból álló $(n-1)$ -dimenziós $\mathbf{E}^{(1)}$ altér invariáns az \mathbf{A} transzformációra nézve. A megfontolást folytatva, megkapjuk a teljes ortogonális sajátvektor-rendszert.) Ebből következik, hogy a sajátvektorokat választva bázisnak, a szimmetrikus transzformáció mátrixa valós elemű diagonálmátrix. Tehát valós térben a főtengely-transzformáció a következőképpen fogalmazható meg.

2.9.3 tétel. *A valós euklideszi tér egy \mathbf{A} szimmetrikus transzformációjához egyértelműen hozzárendelhető egy $A(\mathbf{x}; \mathbf{x}) = (\mathbf{A}\mathbf{x}, \mathbf{x})$ kvadratikus alak. Ha ezt a kvadratikus alakot az \mathbf{A} transzformáció sajátvektoraiból alkotott bázisban írjuk fel, akkor olyan négyzetösszegbe megy át, amelynek együtthatói a transzformáció sajátértékei (lásd a 2.8.16 tételt).*

A 2.8.17 tétel mintájára megfogalmazható két kvadratikus alak egyidejű négyzetösszeggé transzformálásra vonatkozó tétel.

2.9.4 tétel. *Legyen $A(\mathbf{x}; \mathbf{x})$ és $B(\mathbf{x}; \mathbf{x})$ az n -dimenziós valós euklideszi tér egy-egy kvadratikus alakja és legyen $B(\mathbf{x}; \mathbf{x})$ ezenkívül még pozitív definit. Ekkor mindig van a térnek olyan bázisa, amelyben mindkét kvadratikus alak*

négyzetösszegként áll elő. E bázis a kvadratikus alakokhoz az $A(\mathbf{x}; \mathbf{x}) = (\mathbf{A}\mathbf{x}, \mathbf{x})$ és $B(\mathbf{x}; \mathbf{x}) = (\mathbf{B}\mathbf{x}, \mathbf{x})$ összefüggésekkel hozzárendelt (szimmetrikus) transzformációpár általánosított sajátvektoraiból áll, amelyek B -ortogonális rendszert alkotnak.

Jelölje az általánosított sajátvektoroknak a tér egy ortonormált bázisára vonatkozó koordinátáiból képzett koordinátavektorokat $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ és az ezekből alkotott mátrixot \mathbf{U} ; ekkor az általánosított sajátérték-feladat az alábbi alakban írható fel:

$$\mathbf{A}\mathbf{U} = \mathbf{B}\mathbf{U}\mathbf{\Lambda}.$$

Ha ezt beszorozzuk balról az \mathbf{U}^T mátrixszal és figyelembe vesszük az

$$(2.9.8) \quad \mathbf{U}^T \mathbf{B} \mathbf{U} = \mathbf{E}$$

összefüggést, akkor azt kapjuk, hogy

$$(2.9.9) \quad \mathbf{U}^T \mathbf{A} \mathbf{U} = \mathbf{\Lambda}.$$

A (2.9.8) és (2.9.9) összefüggésekből kiolvasható az a kongruens transzformáció, amellyel az \mathbf{A} és \mathbf{B} mátrix egyidejűleg négyzetösszegbe vihető át.

A valós tér szimmetrikus transzformációjának definíciójából következik, hogy egy $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ ortonormált bázishoz tartozó mátrixa valós, szimmetrikus mátrix. A valós tér unitér transzformációjának egy ortonormált bázishoz tartozó \mathbf{Q} mátrixa valós unitér, tehát *ortogonális mátrix*, azaz

$$(2.9.10) \quad \mathbf{Q}^{-1} = \mathbf{Q}^T.$$

Bevezetjük tehát a következő elnevezést.

2.9.2 definíció. *A valós euklideszi térben értelmezett unitér transzformációt ortogonális transzformációnak nevezünk.*

Változatlanul érvényes a 2.8.4 és 2.8.5 tétel, csupán az „unitér” transzformáció helyett mindenütt ortogonális transzformációt kell érteni:

Bármely két vektor skaláris szorzata – tehát speciálisan bármely vektor hossza – a tér ortogonális transzformációjával szemben invariáns.

A valós euklideszi tér egy transzformációja akkor és csak akkor ortogonális, ha bármely ortonormált bázist ortonormált bázisba visz át.

A 2.8.6 tétel is érvényes, tehát ortogonális transzformációk karakterisztikus egyenletének a gyökei a komplex számsík egységkörén helyezkednek el; ez azonban azt jelenti, hogy amennyiben komplex gyökpár lép fel, akkor ehhez a 2.9.1 tétel értelmében kétdimenziós invariáns altér tartozik.

A (2.8.3) és a (2.9.10) összefüggésekből következik, hogy bármely ortogonális mátrix determinánsának négyzete egyenlő az egységgel; ugyanis (2.9.10)

alapján $|\mathbf{Q}\mathbf{Q}^{-1}| = |\mathbf{Q}\mathbf{Q}^T| = |\mathbf{E}|$, és így $|\mathbf{Q}|^2 = 1$. Ez azt jelenti, hogy ortogonális transzformáció ortonormált bázishoz tartozó mátrixának determinánsa vagy $+1$ vagy -1 .

2.9.3 definíció. Azokat az ortogonális transzformációkat, amelyek mátrixának a determinánsa $+1$, **valódiaknak**, azokat pedig, amelyeké -1 , **nem-valódiaknak** nevezzük.

Vizsgáljuk meg az egy- és a kétdimenziós euklideszi tér ortogonális transzformációit.

Tekintsük először az egydimenziós teret. Legyen \mathbf{e} egy, a teret generáló vektor, és legyen \mathbf{Q} a tér egy ortogonális transzformációja. Ekkor teljesülnie kell valamilyen λ -ra a $\mathbf{Q}\mathbf{e} = \lambda\mathbf{e}$ összefüggésnek. Mivel a (2.8.4) tétel szerint $(\mathbf{Q}\mathbf{e}, \mathbf{Q}\mathbf{e}) = (\mathbf{e}, \mathbf{e})$, ezért $\lambda^2(\mathbf{e}, \mathbf{e}) = (\mathbf{e}, \mathbf{e})$, vagyis $\lambda = \pm 1$. Az egydimenziós térben tehát csak két ortogonális transzformáció létezik: az egyik valódi ortogonális transzformáció, mégpedig $\mathbf{Q}\mathbf{x} = \mathbf{x}$, a másik pedig nemvalódi ortogonális transzformáció, mégpedig $\mathbf{Q}\mathbf{x} = -\mathbf{x}$.

Megjegyzés. Eszerint az egydimenziós valódi ortogonális transzformáció az azonos leképezés, azaz a tér mindegyik vektorát önmagába transzformálja; az egydimenziós nemvalódi ortogonális transzformációt pedig – a geometriai szemléletből kölcsönzött kifejezéssel – pontra való tükrözésnek nevezhetjük.

Vizsgáljuk a kétdimenziós, valós euklideszi tér ortogonális transzformációit. Legyen $\mathbf{e}_1, \mathbf{e}_2$ a tér egy ortonormált bázisa és a \mathbf{Q} transzformáció mátrixa ebben a bázisban

$$(2.9.11) \quad \mathbf{Q} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

(a) Tegyük fel először, hogy \mathbf{Q} determinánsa $+1$:

$$(2.9.12) \quad ad - bc = 1,$$

tehát hogy \mathbf{Q} valódi ortogonális transzformáció. Abból, hogy a \mathbf{Q} transzformáció ortogonális, következik (2.9.10) alapján, hogy $\mathbf{Q}^{-1} = \mathbf{Q}^T$, vagyis

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \begin{bmatrix} a & c \\ b & d \end{bmatrix},$$

másrészt (2.9.12) figyelembevételével

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Így a $\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$ egyenlőség miatt a transzformáció mátrixa:

$$\mathbf{Q} = \begin{bmatrix} a & -c \\ c & a \end{bmatrix},$$

ahol $a^2 + c^2 = 1$. Ezért a és c a következőképpen helyettesíthető: $a = \cos \varphi$, $c = \sin \varphi$. Érvényes tehát a következő tétel.

2.9.5 tétel. *A kétdimenziós tér minden valódi ortogonális transzformációjának mátrixa ortonormált bázisban*

$$\mathbf{Q} = \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix}.$$

A 2.5.4 pontban bemutatott 20. Példa alapján tudjuk, hogy ez a transzformáció a kétdimenziós geometriai térben, azaz síkban, φ szögű, origó körüli elforgatást jelent. A tételből következik, hogy a transzformáció karakterisztikus egyenlete

$$\lambda^2 - 2\lambda \cos \varphi + 1 = 0;$$

ennek diszkriminánsa

$$4(\cos^2 \varphi - 1) \leq 0.$$

A triviális $\varphi = 0$ esettől eltekintve – ami az azonos transzformációnak felel meg – a diszkrimináns negatív, tehát a transzformációnak nincs (valós) sajátértéke.

(b) Legyen most \mathbf{Q} nemvalódi ortogonális transzformáció, vagyis determinánsa -1 :

$$(2.9.13) \quad ad - bc = -1.$$

Ekkor a \mathbf{Q} mátrix ortogonalitásából az (a) pont alatti gondolatmenethez hasonlóan következik

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \begin{bmatrix} a & c \\ b & d \end{bmatrix},$$

és innen (2.9.13) figyelembevételével

$$\begin{bmatrix} -d & b \\ c & -a \end{bmatrix} = \begin{bmatrix} a & c \\ b & d \end{bmatrix}.$$

A transzformáció mátrixa tehát

$$\mathbf{Q} = \begin{bmatrix} a & c \\ c & -a \end{bmatrix},$$

ahol ismét $a^2 + c^2 = 1$. Tehát a és c értékét ismét trigonometrikus függvényekkel helyettesítve, pl. $a = \cos \varphi$ és $c = \sin \varphi$ választással

$$\mathbf{Q} = \begin{bmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{bmatrix}$$

adódik, vagyis a transzformáció karakterisztikus egyenlete

$$(\lambda + \cos \varphi)(\lambda - \cos \varphi) - \sin^2 \varphi = 0,$$

amiből $\lambda_1 = 1$ és $\lambda_2 = -1$. Tehát a nemvalódi ortogonális transzformációnak van két (valós) sajátértéke*. A hozzájuk tartozó \mathbf{u}_1 , \mathbf{u}_2 sajátvektorokat a

$$\mathbf{Q}\mathbf{u}_1 = \mathbf{u}_1 \quad \text{és} \quad \mathbf{Q}\mathbf{u}_2 = -\mathbf{u}_2$$

egyenletekből határozzuk meg. Azt kapjuk, hogy a sajátvektorok koordinátái az \mathbf{e}_1 , \mathbf{e}_2 bázisban

$$\mathbf{u}_1 = \begin{bmatrix} \cos \frac{\varphi}{2} \\ \sin \frac{\varphi}{2} \end{bmatrix} \quad \text{és} \quad \mathbf{u}_2 = \begin{bmatrix} -\sin \frac{\varphi}{2} \\ \cos \frac{\varphi}{2} \end{bmatrix}.$$

Az \mathbf{u}_1 , \mathbf{u}_2 sajátvektorok ortogonálisak és ha ezeket választjuk bázisként, akkor ebben a bázisban a nemvalódi ortogonális transzformáció mátrixa

$$(2.9.14) \quad \mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Hasonlóképpen kaphatjuk meg a kétdimenziós valós euklideszi tér másik nemvalódi ortogonális transzformációját, amelynek mátrixa

$$\mathbf{Q} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Mint a 2.5.4 pontból tudjuk, a síkban ez a két transzformáció egy-egy koordinátatengelyre való tükrözésnek felel meg.

Tetszőleges dimenziójú térben az ortogonális transzformációk mátrixának legegyszerűbb alakjára vonatkozik a következő tétel.

2.9.6 tétel. *Ha \mathbf{Q} az n -dimenziós \mathbf{E}_n valós euklideszi tér egy ortogonális transzformációja, akkor van az \mathbf{E}_n térnek olyan $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ ortonormált*

*Mivel a kétdimenziós tér minden nemvalódi ortogonális transzformációjának mátrixa szimmetrikus, ebből következik, hogy mindkét sajátértéke valós és egységnyi abszolút értékű, tehát csak $+1$ és -1 lehet.

bázisa, amelyben az ortogonális Q transzformáció mátrixa a következő alakú:

$$(2.9.15) \quad \begin{bmatrix} 1 & & & & & & & & \\ & 1 & & & & & & & \\ & & \ddots & & & & & & \\ & & & 1 & & & & & \\ & & & & -1 & & & & \\ & & & & & -1 & & & \\ & & & & & & \ddots & & \\ & & & & & & & -1 & \\ & & & & & \cos \varphi_1 & -\sin \varphi_1 & & \\ & & & & & \sin \varphi_1 & \cos \varphi_1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & \cos \varphi_k & -\sin \varphi_k \\ & & & & & & & & \sin \varphi_k & \cos \varphi_k \end{bmatrix}.$$

Bizonyítás. A 2.9.1 tétel értelmében az E térben található egy- vagy kétdimenziós $E^{(1)}$ invariáns altér. Ha $E^{(1)}$ egydimenziós, akkor jelölje e_1 ennek az altérnek az egységvektorát. Ha $E^{(1)}$ kétdimenziós, akkor jelölje ennek egy ortonormált bázisát e_1, e_2 .

Ha $E^{(1)}$ egydimenziós, akkor benne a Q transzformáció alakja

$$Qx = x, \quad \text{ill.} \quad Qx = -x.$$

Ha az $E^{(1)}$ altér kétdimenziós, akkor benne a transzformáció valódi ortogonális transzformáció (különben $E^{(1)}$ tartalmazna egydimenziós invariáns alteret), ezért a Q transzformációnak a mátrixa az $E^{(1)}$ altérben

$$(2.9.16) \quad \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix}.$$

Az $E^{(1)}$ altér összes vektorára ortogonális vektorok \tilde{E} halmaza ismét invariáns altér. Legyen ugyanis $x \in \tilde{E}$, vagyis legyen $(x, y) = 0$ minden $y \in E^{(1)}$ vektorra. Mivel a 2.8.4 tétel értelmében $(Qx, Qy) = (x, y)$, ezért egyúttal $(Qx, Qy) = 0$. Mivel továbbá $z = Qy$ befutja az egész $\tilde{E}^{(1)}$ alteret, ha y befutja azt, ezért $(Qx, z) = 0$ minden $z \in \tilde{E}^{(1)}$ vektorra, vagyis $Qx \in \tilde{E}$.

Az E altér $(n-1)$ -dimenziós, ha $E^{(1)}$ egydimenziós, és $(n-2)$ -dimenziós, ha $E^{(1)}$ kétdimenziós. Valóban, az első esetben \tilde{E} az e_1 vektorra, a második esetben pedig a e_1, e_2 vektorpárra ortogonális vektorok halmaza.

Az \tilde{E} invariáns altérben ismét található, egy egydimenziós vagy egy kétdimenziós invariáns altér; ebben az altérben ismét kiválasztva egy bázist, és az eljárást folytatva, n páronként ortogonális egységvektort kapunk, amelyet E bázisának választhatunk. A transzformáció mátrixa ebben a bázisban a (2.9.15) alakot veszi fel.

A főátló mentén elhelyezkedő ± 1 értékek az egydimenziós invariáns altereknek, a (2.9.16) alakú blokkok pedig a kétdimenziós invariáns altereknek felelnek meg. ■

Megjegyzés. Ha a geometriai térből vett szemlélet alapján *egyszerű forgatásnak* nevezzük az olyan valódi ortogonális transzformációt, amely kétdimenziós síkbeli forgatást jelent, de változatlanul hagyja a síkra merőleges $(n - 2)$ -dimenziós alteret, akkor ennek mátrixa az alábbi alakra hozható:

$$\begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \begin{matrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{matrix} & \\ & & & & \ddots \\ & & & & & 1 \end{bmatrix}.$$

Ha a geometriai térből vett szemlélet alapján *egyszerű tükrözésnek* nevezzük az olyan nemvalódi lineáris transzformációt, amely ellenkezőjére változtatja egy egydimenziós altern vektorainak előjelét, de változatlanul hagyja az erre ortogonális $(n - 1)$ -dimenziós altern vektorait, akkor ennek mátrixa valamilyen bázisban az alábbi alakot veszi fel:

$$\begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & -1 \\ & & & & & 1 \\ & & & & & & \ddots \\ & & & & & & & 1 \end{bmatrix}.$$

A 2.9.6 tétel segítségével megmutatható, hogy minden ortogonális transzformáció előállítható bizonyos számú egyszerű forgatás és egyszerű tükrözés szorzataként.

Megemlítjük még, hogy a valós euklideszi térben a 2.8.14 tételhez hasonlóan bizonyítható a következő, vele analóg tétel.

2.9.7 tétel. Az n -dimenziós valós euklideszi tér bármely nonszinguláris A transzformációja előállítható

$$(2.9.17) \quad A = SQ$$

alakban, ahol S pozitív definit (valós) szimmetrikus transzformáció, Q pedig (valós) ortogonális transzformáció.

Adott bázis felvételével megkapjuk a megfelelő tételt mátrixokra megfogalmazva:

2.9.8 tétel. Bármely nonszinguláris valós elemű mátrix előállítható egy pozitív definit szimmetrikus mátrix és egy ortogonális mátrix szorzataként.

26. Példa. Állítsuk elő az $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$ mátrixot egy pozitív definit szimmetrikus mátrix és egy ortogonális mátrix szorzataként!

Megoldás. Első lépésként írjuk fel az $\mathbf{A}\mathbf{A}^\top$ szorzatot:

$$\mathbf{A}\mathbf{A}^\top = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}.$$

Mivel \mathbf{A} nemszinguláris, $\mathbf{A}\mathbf{A}^\top$ pozitív definit mátrix. Abból a célból, hogy a négyzetgyökét megkapjuk, határozzuk meg a sajátértékeit és a sajátvektorait. A karakterisztikus polinomja

$$D(\lambda) \equiv |\lambda \mathbf{E} - \mathbf{A}\mathbf{A}^\top| = \begin{vmatrix} \lambda - 1 & 1 \\ 1 & \lambda - 2 \end{vmatrix} = \lambda^2 - 3\lambda + 1.$$

A sajátértékei tehát $\lambda_1 = \frac{3 + \sqrt{5}}{2}$ és $\lambda_2 = \frac{3 - \sqrt{5}}{2}$. Egyszerű számolással igazolható, hogy

$$\frac{3 + \sqrt{5}}{2} = \left(\frac{1 + \sqrt{5}}{2} \right)^2 \quad \text{és} \quad \frac{3 - \sqrt{5}}{2} = \frac{1}{\left(\frac{1 + \sqrt{5}}{2} \right)^2},$$

és vezessük be a további számolás egyszerűsítése kedvéért a $\phi = \frac{1 + \sqrt{5}}{2}$ jelölést.* Ezzel a sajátértékek $\lambda_1 = \phi^2$, $\lambda_2 = \frac{1}{\phi^2}$ alakban írhatók. A sajátvektorokat az alábbi egyenletrendszerek megoldása szolgáltatja:

$$\lambda_1 = \phi^2 = \phi + 1$$

$$\begin{bmatrix} \lambda_1 - 1 & 1 \\ 1 & \lambda_1 - 2 \end{bmatrix} \begin{bmatrix} u_{11} \\ u_{21} \end{bmatrix} = \begin{bmatrix} \phi & 1 \\ 1 & \phi - 1 \end{bmatrix} \begin{bmatrix} u_{11} \\ u_{21} \end{bmatrix} = 0.$$

A második sor elemeihez tartozó „aldeterminánsok” arányával $\begin{bmatrix} u_{11} \\ u_{21} \end{bmatrix} = \begin{bmatrix} -1 \\ \phi \end{bmatrix}$.

Mivel $\frac{3 - \sqrt{5}}{2} = 2 - \frac{1 + \sqrt{5}}{2}$, azaz $\lambda_2 = 2 - \phi$, a második sajátértékekhez tartozó sajátvektort a

$$\begin{bmatrix} \lambda_2 - 1 & 1 \\ 1 & \lambda_2 - 2 \end{bmatrix} \begin{bmatrix} u_{12} \\ u_{22} \end{bmatrix} = \begin{bmatrix} 1 - \phi & 1 \\ 1 & -\phi \end{bmatrix} \begin{bmatrix} u_{12} \\ u_{22} \end{bmatrix} = 0$$

*Csupán az érdekesség kedvéért jegyezzük meg, hogy ϕ az aranymetszés néven ismert nevezetes arány, amely kielégíti a $\phi^2 = \phi + 1$ összefüggést.

egyenletrendszer megoldása szolgáltatja, amelyet most az *első* sor elemeihez tartozó „aldeterminánsok“ arányával fejezünk ki: $\begin{bmatrix} u_{12} \\ u_{22} \end{bmatrix} = \begin{bmatrix} \phi \\ 1 \end{bmatrix}$. A sajátvektorok definíciója alapján felírható

$$\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} -1 & \phi \\ \phi & 1 \end{bmatrix} = \begin{bmatrix} -1 & \phi \\ \phi & 1 \end{bmatrix} \begin{bmatrix} \phi^2 & \\ & \frac{1}{\phi^2} \end{bmatrix}.$$

Most szorozzunk jobbról a

$$\begin{bmatrix} -1 & \phi \\ \phi & 1 \end{bmatrix}^{-1} = \frac{1}{\phi^2 + 1} \begin{bmatrix} -1 & \phi \\ \phi & 1 \end{bmatrix}$$

inverz mátrixszal:

$$\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} -1 & \phi \\ \phi & 1 \end{bmatrix} \begin{bmatrix} \phi^2 & \\ & \frac{1}{\phi^2} \end{bmatrix} \begin{bmatrix} -1 & \phi \\ \phi & 1 \end{bmatrix} \frac{1}{\phi^2 + 1}.$$

Azt látjuk, hogy az adott mátrix és a sajátértékekből alkotott diagonálmátrix hasonlóak. A mátrix négyzetgyökét tehát úgy kapjuk meg, hogy a sajátértékek pozitív négyzetgyökét vesszük és ugyanazt a hasonlósági transzformációt végezzük:

$$\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}^{\frac{1}{2}} = \begin{bmatrix} -1 & \phi \\ \phi & 1 \end{bmatrix} \begin{bmatrix} \phi & 0 \\ 0 & \frac{1}{\phi} \end{bmatrix} \begin{bmatrix} -1 & \phi \\ \phi & +1 \end{bmatrix} \frac{1}{\phi^2 + 1}.$$

Figyelembe véve, hogy $\frac{1}{\phi} = \phi - 1$, a szorzást elvégezve

$$\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}^{\frac{1}{2}} = \frac{\phi}{\phi^2 + 1} \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}$$

adódik. Mivel $\frac{\phi}{\phi^2 + 1} = \frac{1}{\phi + \frac{1}{\phi}} = \frac{1}{\sqrt{5}}$, a pozitív definit szimmetrikus mátrix

$$(\mathbf{A}\mathbf{A}^\top)^{\frac{1}{2}} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}.$$

A második tényezőt, az ortogonális mátrixot megkapjuk, ha ennek inverzét megszorozzuk az adott mátrixszal: $\mathbf{U} = (\mathbf{A}\mathbf{A}^\top)^{-\frac{1}{2}}\mathbf{A}$. A fentiek alapján

$$(\mathbf{A}\mathbf{A}^\top)^{-\frac{1}{2}} = \frac{\sqrt{5}}{5} \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix},$$

így az ortogonális mátrixot

$$\mathbf{U} = \frac{\sqrt{5}}{5} \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$$

alakban kapjuk. A keresett előállítás tehát

$$\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix} \cdot \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}.$$

2.9.3 Kvadratikus alakok

Legyen $A(\mathbf{x}; \mathbf{x})$ az n -dimenziós valós euklideszi térben értelmezett kvadratikus alak. Tekintsük az \mathbf{x} vektor egy $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bázisra vonatkozó x_1, x_2, \dots, x_n koordinátáit változóknak és legyen \mathbf{A} a kvadratikus alak szimmetrikus mátrixa ebben a bázisban. Megvizsgáljuk az $A(\mathbf{x}; \mathbf{x}) = 1$ egyenletet és annak geometriai jelentését az adott bázisban felírt

$$(2.9.18) \quad \mathbf{x}^T \mathbf{A} \mathbf{x} = 1$$

alakja segítségével $n = 2, 3$ esetén, ahol \mathbf{x} az x_i koordinátákból alkotott vektor.

(1) Ha \mathbf{A} másodrendű mátrix, akkor (2.9.18) részletesen kiírva a következő alakot veszi fel:

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = 1.$$

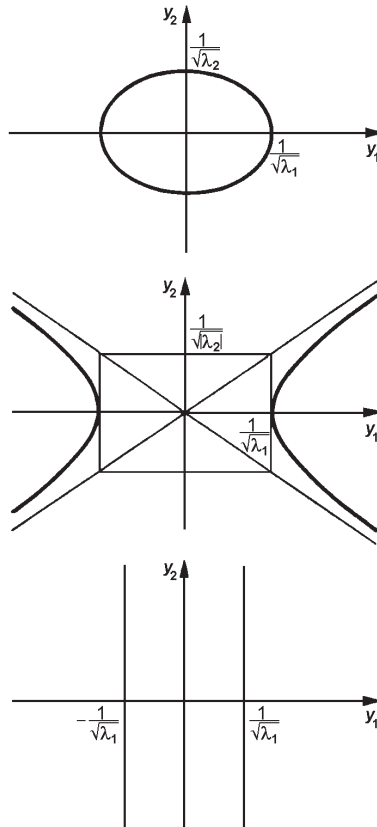
Ez – mint az analitikus geometriából ismeretes – egy másodrendű görbe egyenlete, amelyet a főtengetytétel értelmében négyzetösszegre transzformálhatunk. Legyen az \mathbf{A} mátrix λ_1 és λ_2 sajátértékéhez tartozó sajátvektor \mathbf{u}_1 és \mathbf{u}_2 . A sajátvektorokat válasszuk új bázisként, és az \mathbf{x} vektor új bázisra vonatkozó koordinátái alkossák az \mathbf{y} vektort. Vagyis legyen $\mathbf{U}\mathbf{y} = \mathbf{x}$, azaz

$$(2.9.19) \quad \mathbf{y} = \mathbf{U}^T \mathbf{x}.$$

Behelyettesítve a (2.9.18) összefüggésbe, az új bázisra vonatkozó y_k változókra nézve

$$(2.9.20) \quad \lambda_1 y_1^2 + \lambda_2 y_2^2 = 1.$$

Innen kiolvasható, hogy a másodrendű görbe jellegét a sajátértékek előjele határozza meg.



A következő eseteket különböztetjük meg (csak a valós alakzatokat tekintjük, lásd az ábrát):

- (a) mindkét sajátérték pozitív (ellipszis);
- (b) a sajátértékek különböző előjelűek, (hiperbola);
- (c) az egyik sajátérték 0, a másik pozitív (két párhuzamos egyenes).

Parabolát nem kaphatunk az egyenletből, hiszen az nem tartalmaz a változóknak elsőfokú tagot.

A (2.9.20) képlet alapján könnyen megállapítható a sajátértékek és az ellipszis, ill. a hiperbola tengelyei közötti kapcsolat. A (2.9.20) összefüggésből

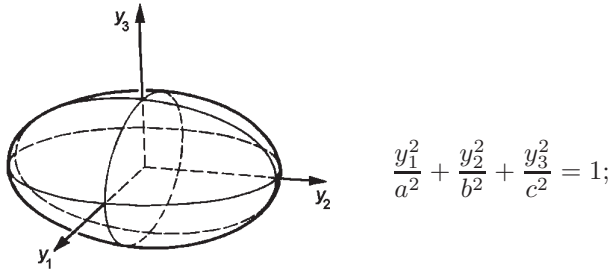
$$\frac{y_1^2}{\frac{1}{\lambda_1}} + \frac{y_2^2}{\frac{1}{\lambda_2}} = 1, \text{ ahonnan az } a = \frac{1}{\sqrt{|\lambda_1|}} \text{ és } b = \frac{1}{\sqrt{|\lambda_2|}}$$

jelölés bevezetésével $\lambda_1 > 0$ és $\lambda_2 > 0$, ill. $\lambda_1 \lambda_2 < 0$ esetén megkapjuk az ellipszis, ill. hiperbola jól ismert középponti egyenletét: $\frac{y_1^2}{a^2} \pm \frac{y_2^2}{b^2} = 1$. Ha

$\lambda_2 = 0$, $\lambda_1 > 0$, akkor az egyenlet a következő alakban írható fel: $\lambda_1 y_1^2 = 1$, azaz $y_1 = \pm \frac{1}{\sqrt{\lambda_1}}$, ami valóban két párhuzamos egyenes egyenlete.

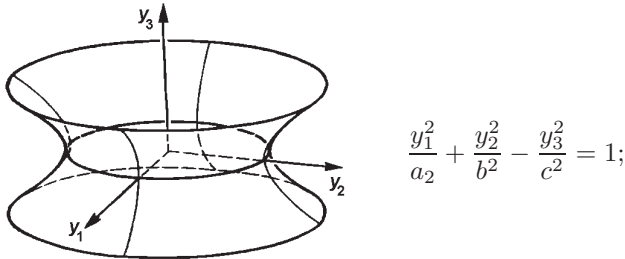
(2) Abban az esetben, ha \mathbf{A} harmadrendű mátrix, a (2.9.18) összefüggés egy másodrendű felület egyenletét adja. Az eljárás ekkor sem különbözik lényegesen a kétdimenziós esettől. A sajátértékek előjelétől függően most a következő eseteket különböztethetjük meg:

(a) mindhárom sajátérték pozitív (ellipszoid):



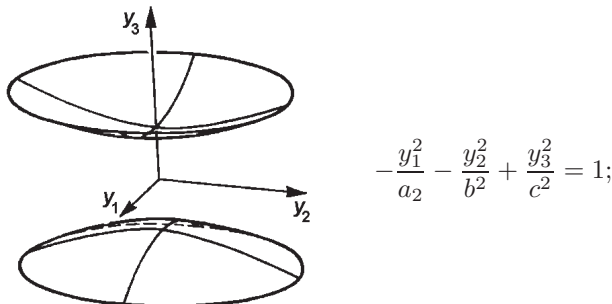
a) ábra

(b) két sajátérték pozitív, egy sajátérték negatív (egyköpenyű hiperboloid):



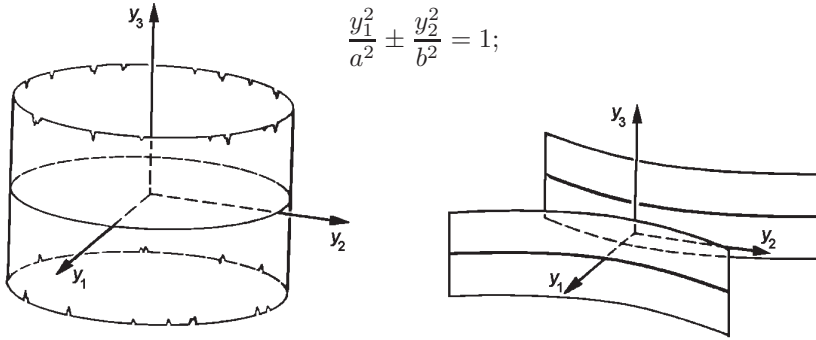
b) ábra

(c) egy sajátérték pozitív, két sajátérték negatív (kétköpenyű hiperboloid):



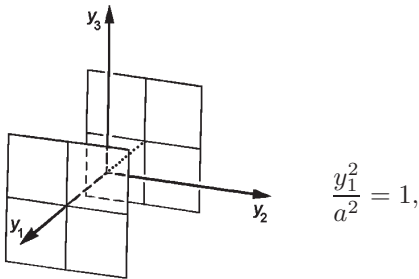
c) ábra

(d) egy sajátérték zérus (ha a másik két sajátérték pozitív, akkor elliptikus henger, ha különböző előjelű, hiperbolikus henger)



d) ábra

(e) két sajátérték zérus, a harmadik pozitív (két párhuzamos síkká degenerált másodrendű felület):



e) ábra

azaz $y_1 = \pm a$.

Hangsúlyozzuk, hogy mind a két-, mind a háromdimenziós feladatban a sajátvektorok a főtengelek irányát jelölik ki, a sajátértékek abszolút értékének reciprokai pedig a féltengelek hosszának a négyzetével egyenlők. Ebből a geometriai interpretációból ered a „főtengelek-transzformáció” elnevezés.

3. fejezet

MÁTRIXFÜGGVÉNYEK

*A világegyetem annyira egyszerű, hogy egyváltozós differenciálegyenlettel lehet elemezni – minden csak az idő függvénye.
(James Peebles princetoni kozmológus)*

Ebben a fejezetben mátrixok függvényét definiáljuk és módszereket adunk meghatározásukra. Felhasználjuk a mátrixalgebrának az 1. fejezetben közölt alapfogalmait, továbbá a lineáris transzformációk 2. fejezetben megismert elméletét, és megmutatjuk, hogy a lineáris transzformációk és a mátrixok közötti szoros kapcsolat miként teszi lehetővé a mátrixok szerkezetének mélyebb elemzését. Itt kapott helyet (téglalap alakú) mátrixoknak a szinguláris értékek szerinti felbontása, ami az újabb kutatások nélkülözhetetlen eszközevé vált és hozzájárul a mátrixok szerkezetéről alkotott képünkhöz. Látni fogjuk, milyen összefüggés található a sajátérték-probléma és a mátrixok diagonalizálhatósága között, majd megmutatjuk, hogy a mátrixok a diagonalizálható (egyszerű struktúrájú) és a nem diagonalizálható (nem egyszerű struktúrájú) mátrixok osztályára oszlanak. Bebizonyítjuk, hogy ha egy mátrix minimálpolinomjának csupa egyszeres gyöke van, akkor a mátrix függvénye a Lagrange-féle interpoláció alkalmazásával a minimálpolinom fokszámánál alacsonyabb fokú polinomra redukálható és ez hasonlósági transzformációval diagonalizálható. Itt tárgyaljuk a kommutatív blokkokból álló hipermátrixok spektrálfelbontását, és ennek speciális eseteként Kronecker-polinomok spektrálfelbontását.

Ha a minimálpolinom gyökei nem mind egyszeres gyökök, akkor az Hermite-féle interpoláció alkalmazásával megmutatjuk, hogy a mátrix függvénye (és így a mátrix is) nem diagonalizálható, hanem idempotens és nilpotens mátrixok lineáris kombinációjaként állítható elő. Bevezetjük a fővektorok fogalmát és konstruktív módszert adunk nilpotens mátrixok Jordan-féle normálalakra való transzformációjára, majd az elemi osztók elméletét ismertetjük.

Végül a mátrixfüggvények legfontosabb alkalmazási területével, lineáris differenciálegyenlet-rendszerek megoldási módszereivel foglalkozunk és számos kidolgozott példa bemutatásával igyekszünk az Olvasó érdeklődését felkelteni a mátrixelmélet alkalmazásai iránt.

3.1 EGYSZERŰ STRUKTÚRÁJÚ MÁTRIXOK SPEKTRÁLIS TULAJDONSÁGAI

3.1.1 Mátrix spektrálfelbontása

A 2.6 szakaszban megismerkedtünk a lineáris transzformációk sajátérték-feladatának fogalmával. Láttuk, hogy a transzformáció sajátértékeinek és sajátvektorainak meghatározása visszavezethető a transzformáció adott bázishoz tartozó mátrixára vonatkozó sajátérték-feladat megoldására, vagyis mindig megfogalmazható úgy is, mint mátrix sajátértékeinek és sajátvektorainak meghatározása. Az alábbiakban megmutatjuk, hogy a mátrix egy speciális diadikus előállítása is a sajátértékek és sajátvektorok fogalmához vezet.

Legyen \mathbf{A} egy n -edrendű kvadratikus mátrix. Azt vizsgáljuk, hogy \mathbf{A} milyen feltételek mellett állítható elő olyan diádok lineáris kombinációjaként, amelyeknek az oszlop-, ill. sorvektorai biortogonális vektorrendszert (lásd az 1.6.2 definíciót) alkotnak. Tegyük fel, hogy \mathbf{A} olyan mátrix, amely előállítható a mondott alakban, és vizsgáljuk meg közelebbről, mi ekkor a diádok vektortényezőinek, ill. a lineáris kombináció együtthatóinak a jelentése.

Legyen tehát

$$(3.1.1) \quad \mathbf{A} = \sum_{k=1}^n \lambda_k \mathbf{u}_k \mathbf{v}_k^T,$$

ahol

$$(3.1.2) \quad \mathbf{v}_k^T \mathbf{u}_l = \delta_{kl}.$$

Vezessük be az alábbi jelöléseket:

$$(3.1.3) \quad \mathbf{U} = [\mathbf{u}_1 \quad \mathbf{u}_2 \dots \mathbf{u}_n], \quad \mathbf{V}^T = \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix}, \quad \mathbf{\Lambda} = \langle \lambda_1 \quad \lambda_2 \dots \lambda_n \rangle;$$

ezzel a jelöléssel (3.1.1) és (3.1.2) az alábbi alakot ölti:

$$(3.1.4) \quad \mathbf{A} = \mathbf{U} \mathbf{\Lambda} \mathbf{V}^T,$$

ahol

$$(3.1.5) \quad \mathbf{V}^T \mathbf{U} = \mathbf{E}.$$

3.1 EGYSZERŰ STRUKTÚRÁJÚ MÁTRIXOK SPEKTRÁLIS TULAJDONSÁGAI 257

A (3.1.5) összefüggésből

$$(3.1.6) \quad \mathbf{V}^T = \mathbf{U}^{-1};$$

ezt behelyettesítve a (3.1.4) egyenletbe,

$$(3.1.7) \quad \mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1}.$$

Innen kiolvasható: *egy mátrix előállítása biortogonális vektorrendszert alkotó diádok lineáris kombinációjaként és diagonalizálása hasonlósági transzformáció segítségével ekvivalens egymással.*

A (3.1.4) és (3.1.5) összefüggések alapján felírható:

$$(3.1.8) \quad \mathbf{A}\mathbf{U} = \mathbf{U}\mathbf{\Lambda},$$

ill.

$$(3.1.9) \quad \mathbf{V}^T \mathbf{A} = \mathbf{\Lambda} \mathbf{V}^T;$$

azaz bármelyik \mathbf{u}_k , ill. \mathbf{v}_k^T vektorra fennáll, hogy

$$(3.1.10) \quad \mathbf{A}\mathbf{u}_k = \lambda_k \mathbf{u}_k,$$

ill.

$$(3.1.11) \quad \mathbf{v}_k^T \mathbf{A} = \lambda_k \mathbf{v}_k^T.$$

A 2.6 szakaszban használt elnevezések szerint a (3.1.10) egyenlettel meghatározott λ_k értékek a mátrix sajátértékei, az \mathbf{u}_k vektorok pedig a λ_k sajátértékhez tartozó sajátvektorok.

Mint a következőkben látni fogjuk, a (3.1.10) és (3.1.11) egyenleteknek ugyanazokra a λ_k értékekre van megoldásuk. Ezért a λ_k értékeket továbbra is – mindkét egyenletre vonatkozóan – sajátértékeknek nevezzük. A (3.1.11) egyenlet megoldásait *bal oldali sajátvektoroknak* nevezzük, míg a sajátvektor elnevezésen továbbra is a (3.1.10) egyenlet megoldásait értjük. Ha hangsúlyozni akarjuk, hogy nem a bal oldali sajátvektorokról van szó, akkor a (3.1.10) megoldásaira a *jobb oldali sajátvektor* elnevezést használjuk. Könnyű belátni, hogy a mátrix (jobb oldali) sajátvektorai megegyeznek transzponáltjának bal oldali sajátvektoraival – és megfordítva. Ezért minden olyan tétel, amelyet a sajátvektorokra kimondunk, értelemszerűen fennáll a bal oldali sajátvektorokra is. A mátrix (3.1.7) alakú előállítását a mátrix *spektrálfelbontásának*, a hasonlósági transzformáció jobb oldali sajátvektorokból alkotott \mathbf{U} mátrixát pedig *modálmátrixnak* (lásd a 2.8.4 definíciót) nevezzük.

A λ_k sajátértékeket a (3.1.10), ill. (3.1.11) összefüggés alapján úgy határozzuk meg, hogy a

$$(3.1.12) \quad (\lambda \mathbf{E} - \mathbf{A})\mathbf{u} = \mathbf{0},$$

ill. a

$$(3.1.13) \quad \mathbf{v}^T(\lambda \mathbf{E} - \mathbf{A}) = \mathbf{0}$$

homogén lineáris egyenletrendszernek a triviálistól különböző megoldását keressük. Ilyen megoldás létezésének szükséges és elégséges feltétele mindkét egyenletre nézve az, hogy a $\lambda \mathbf{E} - \mathbf{A}$ mátrix szinguláris legyen, azaz teljesüljön

$$(3.1.14) \quad |\lambda \mathbf{E} - \mathbf{A}| = 0.$$

A 2.6.3 definíciónak megfelelően a $\lambda \mathbf{E} - \mathbf{A}$ mátrixot most is az \mathbf{A} mátrix *karakterisztikus mátrixának*, a

$$(3.1.15) \quad D(\lambda) \equiv |\lambda \mathbf{E} - \mathbf{A}|$$

λ -ban n -edfokú polinomot *karakterisztikus polinomjának*, a (3.1.14) egyenletet pedig *karakterisztikus egyenletének* nevezzük. Ha a karakterisztikus polinom fokszámát is fel akarjuk tüntetni, $D_n(\lambda)$ jelölést használunk.

A karakterisztikus polinom λ hatványai szerint felírva:

$$(3.1.16) \quad D(\lambda) = \lambda^n - \lambda^{n-1} \sum_{k=1}^n a_{kk} + \lambda^{n-2} \sum_{\binom{n}{2}} \begin{vmatrix} a_{ii} & a_{ik} \\ a_{ki} & a_{kk} \end{vmatrix} - \dots + (-1)^n |\mathbf{A}|,$$

ahol az összegzést az $1, 2, \dots, n$ indexek minden megfelelő kombinációjára kell elvégezni, azaz a k -adik hatvány együtthatója a mátrix $(n - k)$ -adrendű főminorainak az összege. Innen az is kiolvasható, hogy szinguláris mátrixoknak mindig van zérus sajátértékük. A (3.1.14) karakterisztikus egyenletnek a komplex számok körében mindig n számú megoldása van, amelyek között azonban lehetnek többszörös gyökök is.

3.1.1 tétel. Az n -edrendű \mathbf{A} mátrix karakterisztikus polinomját az alábbi alakban írva:

$$(3.1.17) \quad D_n(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{\alpha_k}, \quad \text{ahol} \quad \sum_{k=1}^s \alpha_k = n,$$

a karakterisztikus mátrix rangja a $\lambda = \lambda_k$ helyen legalább $n - \alpha_k$, azaz

$$(3.1.18) \quad \varrho(\lambda_k \mathbf{E} - \mathbf{A}) \geq n - \alpha_k.$$

Bizonyítás. Jelölje $D_{ij}(\lambda)$ a karakterisztikus mátrix adjungáltjának elemeit:

$$(3.1.19) \quad \text{adj}(\lambda \mathbf{E} - \mathbf{A}) = [D_{ij}(\lambda)].$$

3.1 EGYSZERŰ STRUKTÚRÁJÚ MÁTRIXOK SPEKTRÁLIS TULAJDONSÁGAI 259

A determináns deriválási szabályát alkalmazva azt kapjuk, hogy a karakterisztikus polinom deriváltjai a karakterisztikus mátrix főminorainak összegeként fejezhetők ki:

$$(3.1.20) \quad D'(\lambda) = \sum_{i=1}^n D_{ii}(\lambda).$$

$$(3.1.21) \quad D''(\lambda) = 2 \sum_{\binom{n}{2}} D \begin{pmatrix} i_1 & i_2 \\ i_1 & i_2 \end{pmatrix} (\lambda),$$

ahol $D \begin{pmatrix} i_1 & i_2 \\ i_1 & i_2 \end{pmatrix} (\lambda)$ a karakterisztikus mátrix i_1 és i_2 indexű sorának és oszlopának elhagyásával adódó $(n-2)$ -edrendű főminort jelenti, és az összegezést az $1, 2, \dots, n$ indexek valamennyi másodosztályú kombinációjára kell elvégezni. Hasonlóképpen,

$$(3.1.22) \quad D^{(\alpha_k)}(\lambda) = \alpha_k! \sum_{\binom{n}{\alpha_k}} D \begin{pmatrix} i_1 & i_2 \dots i_{\alpha_k} \\ i_1 & i_2 \dots i_{\alpha_k} \end{pmatrix} (\lambda),$$

ahol $D \begin{pmatrix} i_1 & i_2 \dots i_{\alpha_k} \\ i_1 & i_2 \dots i_{\alpha_k} \end{pmatrix} (\lambda)$ a karakterisztikus mátrix $i_1, i_2, \dots, i_{\alpha_k}$ indexű sorának és oszlopának elhagyásával adódó $(n - \alpha_k)$ -adrendű főminort jelenti, ahol az összegezést az $1, 2, \dots, n$ indexek valamennyi α_k elemű kombinációjára kell elvégezni. A $D^{(\alpha_k)}(\lambda_k) \neq 0$ feltétel miatt a (3.1.22) jobb oldalán szereplő $(n - \alpha_k)$ -adrendű főminorok a $\lambda = \lambda_k$ helyen nem lehetnek valamennyien egyenlők zérussal, tehát $\lambda_k \mathbf{E} - \mathbf{A}$ rangja legalább $n - \alpha_k$. ■

Ennek a tételnek a következményeként adódik a λ_k sajátértékhez tartozó lineárisan független sajátvektorok számára vonatkozó következő tétel.

3.1.2 tétel. Egy mátrix λ_k sajátértékéhez tartozó sajátvektorok száma legfeljebb annyi, amennyi λ_k multiplicitása a karakterisztikus egyenletben.

A sajátérték definíciójából következik, hogy minden sajátértékhez legalább egy sajátvektor tartozik.

Bizonyítás. Helyettesítsük be a karakterisztikus egyenlet λ_k gyökét a (3.1.12) egyenletbe, és jelölje \mathbf{U} az egyenlet lineárisan független partikuláris megoldásaiból alkotott mátrixot. Mivel a 3.1.1 tétel alapján a $\lambda_k \mathbf{E} - \mathbf{A}$ mátrix rangja legalább $n - \alpha_k$, ezért az 1.3.13 tétel értelmében a

$$(3.1.23) \quad (\lambda_k \mathbf{E} - \mathbf{A})\mathbf{U} = \mathbf{0}$$

egyenlet csak akkor állhat fenn, ha \mathbf{U} rangja legfeljebb α_k , és így valóban fennáll a λ_k sajátértékhez tartozó lineárisan független sajátvektorok számára

vonatkozó

$$(3.1.24) \quad 1 \leq \varrho(\mathbf{U}) \leq \alpha_k$$

becslés. ■

E tételből következik, hogy a mátrix (3.1.7) alakú spektrálfelbontása, vagyis n lineárisan független jobb és bal oldali sajátvektor csak akkor létezik, ha a mátrix minden sajátértékéhez pontosan annyi lineárisan független sajátvektor tartozik, amennyi a sajátérték multiplicitása. Speciálisan, ha n különböző sajátérték van, akkor ez a feltétel biztosan teljesül.

A 2.6.6. definícióban a lineáris transzformációk tulajdonságai alapján definiáltuk az egyszerű struktúrájú mátrix fogalmát. Most erre a fogalomra egy ezzel ekvivalens definíciót adunk.

3.1.1 definíció. *Azokat az n -edrendű mátrixokat, amelyeknek n független jobb és bal oldali sajátvektoruk van, egyszerű struktúrájú mátrixoknak nevezzük.*

Az eddig mondottakat összefoglalva megállapíthatjuk, hogy azok a mátrixok egyszerű struktúrájúak, amelyeknek a jobb és bal oldali sajátvektorai teljes biortogonális vektorrendszert alkotnak. Ezek a mátrixok a módálmátrixszal végzett hasonlósági transzformációval diagonalizálhatók. A lineáris transzformációk terminológiájával élve azt mondhatjuk, hogy ha a lineáris transzformáció mátrixa egyszerű struktúrájú, akkor található olyan bázis, amelyben diagonálmátrix.

Azt is láttuk, hogy ha egy mátrix sajátértékei mind különbözőek, akkor biztosan egyszerű struktúrájú; ha van legalább egy többszörös sajátértéke, akkor további vizsgálat szükséges annak eldöntésére, hogy egyszerű struktúrájú-e.

Ezért van jelentősége annak, hogy viszonylag egyszerű feltételeket találjunk, amelyek alapján eldönthető, hogy egy tetszőleges mátrix egyszerű struktúrájú-e vagy sem.

3.1.2 Projektormátrix spektrálfelbontása

Vizsgáljuk most meg az 1.6 szakaszban megismert projektormátrixok sajátértékeinek és sajátvektorainak tulajdonságait (lásd az 1.6.1 definíciót). Bizonyítjuk ezek spektrálfelbontására vonatkozó következő tételt.

3.1.3 tétel. *Ha az r -edrangu, n -edrendű \mathbf{P} projektormátrix egy minimális diadikus felbontása*

$$(3.1.25) \quad \mathbf{P} = \sum_{k=1}^r \mathbf{u}_k \mathbf{v}_k^T = \mathbf{U} \mathbf{V}^T,$$

3.1 EGYSZERŰ STRUKTÚRÁJÚ MÁTRIXOK SPEKTRÁLIS TULAJDONSÁGAI 261

az $\mathbf{E} - \mathbf{P}$ ún. komplementer projektor egy minimális diadikus felbontása pedig

$$(3.1.26) \quad \mathbf{E} - \mathbf{P} = \sum_{l=1}^{n-r} \mathbf{w}_l \mathbf{z}_l^T = \mathbf{W} \mathbf{Z}^T,$$

akkor a \mathbf{P} projektormátrix r -szeres sajátértéke 1 és $(n-r)$ -szeres sajátértéke 0, teljes spektrálfelbontása pedig a

$$(3.1.27) \quad \mathbf{P} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \dots \quad \mathbf{u}_r \quad \mathbf{w}_1 \quad \dots \quad \mathbf{w}_{n-r}] \begin{bmatrix} \overbrace{1}^r & & & & & \\ & \overbrace{1}^{n-r} & & & & \\ & & \ddots & & & \\ & & & 1 & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_r^T \\ \mathbf{z}_1^T \\ \vdots \\ \mathbf{z}_{n-r}^T \end{bmatrix}$$

alakban írható.

Bizonyítás. Az 1.6.3 tétel értelmében az \mathbf{u}_k és \mathbf{v}_k^T vektorok minden minimális diadikus előállításban biortogonális vektorrendszert alkotnak. A (3.1.25) előállításból tehát kiolvasható, hogy $\lambda = 1$ a projektormátrix r -szeres sajátértéke, és ehhez r lineárisan független sajátvektor tartozik: \mathbf{u}_k a jobb oldali, \mathbf{v}_k^T a bal oldali sajátvektorokat adja. Ha pedig a projektormátrixok $\mathbf{P}(\mathbf{E} - \mathbf{P}) = \mathbf{0}$ definiáló egyenletébe behelyettesítjük az $\mathbf{E} - \mathbf{P}$ komplementer projektor (3.1.26) alatti felbontását, akkor $\mathbf{P} \mathbf{W} \mathbf{Z}^T = \mathbf{0}$ adódik, ahonnan az 1.3.9 tétel felhasználásával következik, hogy $\mathbf{P} \mathbf{W} = \mathbf{0}$. Ez pedig azt jelenti, hogy \mathbf{W} bármely oszlopvektora a \mathbf{P} mátrix 0 sajátértékéhez tartozó jobb oldali sajátvektor. Hasonlóképpen az

$$(\mathbf{E} - \mathbf{P})\mathbf{P} = \mathbf{W} \mathbf{Z}^T \mathbf{P} = \mathbf{0}$$

összefüggésből az 1.3.8 tétel felhasználásával következik $\mathbf{Z}^T \mathbf{P} = \mathbf{0}$, azaz \mathbf{Z}^T bármely sorvektora a \mathbf{P} mátrix 0 sajátértékéhez tartozó bal oldali sajátvektor. Ha a (3.1.25) és a (3.1.26) egyenleteket összeadjuk, akkor

$$(3.1.28) \quad \mathbf{E}_n = \mathbf{P} + (\mathbf{E} - \mathbf{P}) = \sum_{k=1}^r \mathbf{u}_k \mathbf{v}_k^T + \sum_{l=1}^{n-r} \mathbf{w}_l \mathbf{z}_l^T = [\mathbf{U} \quad \mathbf{W}] \begin{bmatrix} \mathbf{V}^T \\ \mathbf{Z}^T \end{bmatrix},$$

és így megkaptuk a \mathbf{P} mátrix sajátvektorainak egy teljes biortogonális rendszerét. Tehát \mathbf{P} spektrálfelbontása az alábbi alakban írható:

$$(3.1.29) \quad \mathbf{P} = 1 \cdot \mathbf{U} \mathbf{V}^T + 0 \cdot \mathbf{W} \mathbf{Z}^T = [\mathbf{U} \quad \mathbf{W}] \begin{bmatrix} \mathbf{E}_r & \\ & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}^T \\ \mathbf{Z}^T \end{bmatrix},$$

és ezzel a tételt bebizonyítottuk. ■

A 3.1.3 tétel lehetővé teszi, hogy bármilyen nemteljes biortogonális vektorrendszert teljessé tegyünk. Legyen $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ és $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ nemteljes biortogonális vektorrendszer, azaz

$$(3.1.30) \quad \mathbf{V}^T \mathbf{U} = \mathbf{E}_r, \quad \text{ahol } r < n.$$

Ha a \mathbf{V}^T és \mathbf{U} mátrixot fordított sorrendben szorozzuk össze, akkor r -edrangú projektormátrixot kapunk; ugyanis (3.1.30) miatt $\mathbf{U} \mathbf{V}^T \mathbf{U} \mathbf{V}^T = \mathbf{U} \mathbf{V}^T$. Ha tehát elkészítjük az $\mathbf{E} - \mathbf{U} \mathbf{V}^T$ mátrix (amely $(n-r)$ -edrangú projektormátrix) egy minimális diadikus felbontását:

$$\mathbf{E} - \mathbf{U} \mathbf{V}^T = \mathbf{W} \mathbf{Z}^T,$$

akkor (3.1.28) alapján egy teljes biortogonális vektorrendszert kapunk.

Az 1.6.4 tétel alkalmazásával speciális esetként adódik a következő feladat megoldása. Adott egy $\mathbf{u}_1, \dots, \mathbf{u}_r$ nemteljes unitér vektorrendszer, amelyre tehát

$$(3.1.31) \quad \mathbf{U}^H \mathbf{U} = \mathbf{E}_r \quad r < n.$$

Egészítsük ki ezt a vektorrendszert teljes unitér vektorrendszerré. A fentiek alapján az \mathbf{U}^H és \mathbf{U} mátrixot fordított sorrendben összeszorozva, r -edrangú hermitikus projektort kapunk, és a komplementer projektor szintén hermitikus projektor; ezt hermitikus diádok összegére bontva, megkapjuk az adott vektorrendszert teljes unitér vektorrendszerré kiegészítő vektorokat:

$$(3.1.32) \quad \mathbf{E} - \mathbf{U} \mathbf{U}^H = \mathbf{W} \mathbf{W}^H,$$

ahol

$$(3.1.33) \quad \begin{bmatrix} \mathbf{U} & \mathbf{W} \end{bmatrix} \begin{bmatrix} \mathbf{U}^H \\ \mathbf{W}^H \end{bmatrix} = \mathbf{E}_n.$$

3.1.3 Unitér transzformációval diagonalizálható mátrixok

Most egy tételt közlünk, amely tetszőleges kvadratikus mátrixok háromszög-mátrixszá transzformálására vonatkozik.

3.1.4 tétel (Schur tétele). Minden kvadratikus mátrix alkalmasan választott unitér transzformációval felső (vagy alsó) háromszögmátrixra transzformálható.

Bizonyítás. Legyen \mathbf{A} tetszőleges kvadratikus mátrix, amelynek \mathbf{x}_1 a λ_1 sajátértékéhez tartozó, az

$$(3.1.34) \quad \mathbf{x}_1^H \mathbf{x}_1 = 1$$

3.1 EGYSZERŰ STRUKTÚRÁJÚ MÁTRIXOK SPEKTRÁLIS TULAJDONSÁGAI 263

összefüggéssel normált* sajátvektora. A (3.1.32) összefüggés segítségével ez az egyetlen – a (3.1.34) összefüggés miatt unitérnek tekinthető – vektor teljes unitér vektorrendszerre egészíthető ki:

$$\mathbf{E} - \mathbf{x}_1 \mathbf{x}_1^H = \sum_{k=2}^n \mathbf{y}_k \mathbf{y}_k^H = \mathbf{Y} \mathbf{Y}^H.$$

A (3.1.33) összefüggés alapján tehát az így nyert $\mathbf{X}_1 = [\mathbf{x}_1 \ \mathbf{y}_2 \dots \mathbf{y}_n]$ mátrix unitér mátrix, azaz $\mathbf{X}_1^H \mathbf{X}_1 = \mathbf{E}$. Ha most az \mathbf{A} mátrixot az \mathbf{X}_1 mátrix segítségével végrehajtott unitér transzformációnak vetjük alá, akkor $\mathbf{A} \mathbf{x}_1 = \lambda_1 \mathbf{x}_1$ miatt

$$\mathbf{X}_1^H \mathbf{A} \mathbf{X}_1 = \begin{bmatrix} \mathbf{x}_1^H \\ \mathbf{y}_2^H \\ \vdots \\ \mathbf{y}_n^H \end{bmatrix} [\mathbf{A} \mathbf{x}_1 \ \mathbf{A} \mathbf{y}_2 \dots \mathbf{A} \mathbf{y}_n] = \left[\begin{array}{c|ccc} \lambda_1 & * & * & \dots * \\ \hline 0 & & & \\ \vdots & & \mathbf{A}_2 & \\ 0 & & & \end{array} \right].$$

ahol a csillaggal jelölt helyeken zérustól különböző elemek állnak. Ha az $(n-1)$ -edrendű \mathbf{A}_2 mátrixnak egy λ_2 sajátértékéhez tartozó normált sajátvektora \mathbf{x}_2 , amelyet a $\mathbf{z}_3, \mathbf{z}_4, \dots, \mathbf{z}_n$ vektorokkal ismét unitér mátrixszá egészítünk ki, akkor ennek segítségével az \mathbf{A}_2 mátrix úgy transzformálódik, hogy az első oszlop $\lambda_2 \mathbf{e}_2$ lesz. Egészítsük ki most ezt az $(n-1)$ -edrendű unitér mátrixot a következőképpen:

$$\mathbf{X}_2 = \left[\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & \mathbf{x}_2 \mathbf{z}_3 \dots \mathbf{z}_n & \\ 0 & & & \end{array} \right],$$

ami ismét unitér mátrix, és írjuk fel a második transzformációt:

$$\mathbf{X}_2^H (\mathbf{X}_1^H \mathbf{A} \mathbf{X}_1) \mathbf{X}_2 = \left[\begin{array}{cc|ccc} \lambda_1 & * & * & \dots & * \\ 0 & \lambda_2 & * & \dots & * \\ \hline 0 & 0 & & & \\ \dots & \dots & & \mathbf{A}_3 & \\ 0 & 0 & & & \end{array} \right].$$

Az eljárást folytatva, $n-1$ lépésben az adott mátrixot felső háromszögmátrixszá transzformáltuk, mégpedig unitér transzformációval, mivel az 1.4.2 tétel

*Általában normálnak nevezzük az egységnyi hosszúságú, azaz 1 abszolút értékű vektort.

szerint az $\mathbf{X}_1 \mathbf{X}_2 \dots \mathbf{X}_{n-1} = \mathbf{U}$ szorzat is unitér mátrix:

$$(3.1.35) \quad \mathbf{U}^H \mathbf{A} \mathbf{U} = \mathbf{B} = \begin{bmatrix} \lambda_1 & b_{12} & b_{13} \dots b_{1n} \\ 0 & \lambda_2 & b_{23} \dots b_{2n} \\ 0 & 0 & \lambda_3 \dots b_{3n} \\ \dots & \dots & \dots \\ 0 & 0 & 0 \dots \lambda_n \end{bmatrix}.$$

■

A lineáris transzformációk elméletében láttuk, hogy a normális transzformációknak kitüntetett szerepük van. A 2.8.10 tételben bebizonyítottuk, hogy a normális és csak a normális transzformációk mátrixa diagonalizálható unitér transzformációval (lásd a (2.8.11) képletet). Az alábbiakban ugyanennek a tételnek egy másik bizonyítását adjuk a 3.1.4 tétel felhasználásával.

3.1.5 tétel. *A normális és csak a normális mátrixok transzformálhatók unitér transzformációval diagonálmátrixszá; ennek elemei a mátrix sajátértékei.*

Bizonyítás. Először feltesszük, hogy \mathbf{A} normális mátrix, és belátjuk, hogy ekkor a 3.1.4 tétel alapján alkalmas unitér transzformációval \mathbf{B} felső háromszög-mátrix alakra hozható, amely szintén normális mátrix, sőt diagonálmátrix.

Ugyanis

$$\mathbf{B}^H \mathbf{B} = \mathbf{U}^H \mathbf{A}^H \mathbf{U} \mathbf{U}^H \mathbf{A} \mathbf{U},$$

$$\mathbf{B} \mathbf{B}^H = \mathbf{U}^H \mathbf{A} \mathbf{U} \mathbf{U}^H \mathbf{A}^H \mathbf{U},$$

tehát $\mathbf{A}^H \mathbf{A} = \mathbf{A} \mathbf{A}^H$ miatt (mivel $\mathbf{U} \mathbf{U}^H = \mathbf{E}$),

$$\mathbf{B}^H \mathbf{B} = \mathbf{B} \mathbf{B}^H,$$

vagyis \mathbf{B} normális mátrix. A $\mathbf{B}^H \mathbf{B}$ mátrix bal felső sarokeleme $\bar{\lambda}_1 \lambda_1$, a $\mathbf{B} \mathbf{B}^H$ mátrixé viszont $\bar{\lambda}_1 \lambda_1 + \bar{b}_{12} b_{12} + \bar{b}_{13} b_{13} + \dots + \bar{b}_{1n} b_{1n}$, tehát e két elem egyenlőségéből következik $b_{12} = b_{13} = \dots = b_{1n} = 0$. Hasonló megfontolással kapjuk, hogy a \mathbf{B} mátrix főátlója feletti valamennyi elem zérus, tehát \mathbf{B} diagonálmátrix.

Ezek után feltesszük, hogy \mathbf{B} diagonálmátrix és belátjuk, hogy \mathbf{A} normális mátrix. Mivel bármely diagonálmátrix a konjugáltjával felcserélhető, az ebből unitér transzformációval nyert mátrix is felcserélhető a transzponált konjugáltjával. Mivel ugyanis a \mathbf{B} diagonálmátrixra fennáll $\mathbf{B} \mathbf{B}^H = \mathbf{B}^H \mathbf{B}$, ezért $\mathbf{U} \mathbf{B} \mathbf{U}^H \mathbf{U} \mathbf{B}^H \mathbf{U}^H = \mathbf{U} \mathbf{B}^H \mathbf{U}^H \mathbf{U} \mathbf{B} \mathbf{U}^H$, tehát $\mathbf{A} \mathbf{A}^H = \mathbf{A}^H \mathbf{A}$, vagyis \mathbf{A} normális mátrix, és ezzel a tételt bebizonyítottuk. ■

Mivel az unitér, az hermitikus és a ferdén hermitikus – továbbá az ezeknek megfelelő valós elemű ortogonális, szimmetrikus és ferdén szimmetrikus – mátrixok mind normális mátrixok, ezért a 3.1.5 tétel következménye, hogy valamennyinek a sajátvektorai unitér vektorrendszert alkotnak, azaz a bal

3.1 EGYSZERŰ STRUKTÚRÁJÚ MÁTRIXOK SPEKTRÁLIS TULAJDONSÁGAI 265

oldali sajátvektorok elemei a hozzájuk tartozó jobb oldali sajátvektorok megfelelő elemeinek komplex konjugáltjai. Nézzük most meg, mit mondhatunk e mátrixok sajátértékeinek elhelyezkedéséről. Az unitér és az önadjungált transzformációk vizsgálata során bebizonyítottuk, hogy az előbbi sajátértékei a komplex számsík 0 középpontú egységsugarú körén, az utóbbi sajátértékei pedig a valós számegyenesen helyezkednek el. Ebből következik, hogy ugyanez vonatkozik az unitér, ill. az hermitikus mátrixok sajátértékeinek elhelyezkedésére is (lásd a 2.8.6 és 2.8.3 tételt). Gyakorlás céljából mátrixokra külön is bebizonyítjuk ezeket a fontos tételeket.

3.1.6 tétel. *Unitér mátrix sajátértékeinek abszolút értéke 1.*

Bizonyítás. Legyen \mathbf{U} unitér mátrix, azaz $\mathbf{U}^{-1} = \mathbf{U}^H$.

A sajátvektorokat definiál

$$(3.1.36) \quad \mathbf{U}\mathbf{x}_k = \lambda_k \mathbf{x}_k$$

egyenletben vegyük mindkét oldal transzponált konjugáltját: $\mathbf{x}_k^H \mathbf{U}^H = \overline{\lambda}_k \mathbf{x}_k^H$, és szorozzuk az így kapott kifejezéssel balról a (3.1.36) egyenletet: $\mathbf{x}_k^H \mathbf{U}^H \mathbf{U} \mathbf{x}_k = \lambda_k \overline{\lambda}_k \mathbf{x}_k^H \mathbf{x}_k$. Innen az $\mathbf{U}^H \mathbf{U} = \mathbf{E}$ összefüggés figyelembevételével:

$$(3.1.37) \quad \mathbf{x}_k^H \mathbf{x}_k (1 - |\lambda_k|^2) = 0,$$

majd $\mathbf{x}_k^H \mathbf{x}_k \neq 0$ miatt $|\lambda_k| = 1$ adódik. ■

3.1.7 tétel. *Hermitikus mátrix sajátértékei valós számok.*

Bizonyítás. Legyen \mathbf{A} hermitikus mátrix, azaz

$$(3.1.38) \quad \mathbf{A}^H = \mathbf{A}.$$

A sajátvektorokat definiál

$$(3.1.39) \quad \mathbf{A}\mathbf{x}_k = \lambda_k \mathbf{x}_k$$

egyenletnek vegyük ismét a transzponált konjugáltját:

$$(3.1.40) \quad \mathbf{x}_k^H \mathbf{A}^H = \overline{\lambda}_k \mathbf{x}_k^H,$$

majd szorozzuk meg a (3.1.39) egyenletet balról az \mathbf{x}_k^H , a (3.1.40) egyenletet jobbról az \mathbf{x}_k vektorral, és képezzük e két kifejezés különbségét:

$$(3.1.41) \quad \mathbf{x}_k^H \mathbf{A} \mathbf{x}_k - \mathbf{x}_k^H \mathbf{A}^H \mathbf{x}_k = (\lambda_k - \overline{\lambda}_k) \mathbf{x}_k^H \mathbf{x}_k.$$

A (3.1.38) összefüggés miatt a (3.1.41) egyenlet bal oldala zérus, így $\mathbf{x}_k^H \mathbf{x}_k \neq 0$ következtében fennáll

$$(3.1.42) \quad \lambda_k - \overline{\lambda}_k = 0,$$

tehát a sajátértékek valósak. ■

Ennek alapján egyszerűen belátható a ferdén hermitikus mátrixok sajátértékeinek elhelyezkedésére vonatkozó következő tétel.

3.1.8 tétel. *Ferdén hermitikus mátrix sajátértékei tiszta képzetes számok.*

Bizonyítás. Legyen \mathbf{F} ferdén hermitikus mátrix, azaz

$$(3.1.43) \quad \mathbf{F}^H = -\mathbf{F}.$$

A sajátvektorokat definiáló egyenletnek ismét véve a transzponált konjugáltját, majd az előbbi balról az \mathbf{x}_k^H , az utóbbit jobbról az \mathbf{x}_k vektorral szorozva és a két egyenletet összeadva,

$$(3.1.44) \quad \mathbf{x}_k^H \mathbf{F} \mathbf{x}_k + \mathbf{x}_k^H \mathbf{F}^H \mathbf{x}_k = (\lambda_k + \bar{\lambda}_k) \mathbf{x}_k^H \mathbf{x}_k.$$

A (3.1.43) összefüggés alapján a (3.1.44) egyenlet bal oldala ismét zérus, így $\mathbf{x}_k^H \mathbf{x}_k \neq 0$ miatt most

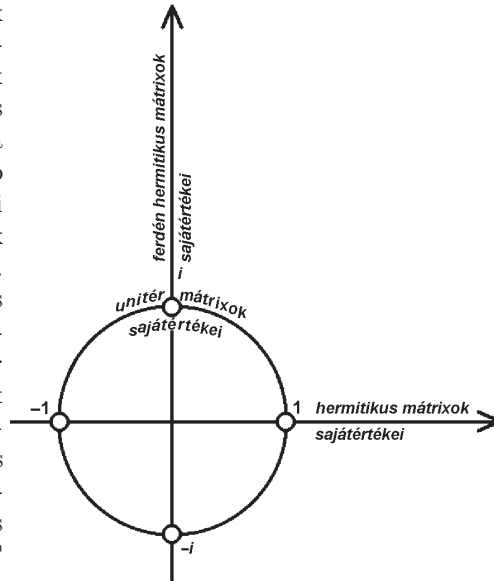
$$(3.1.45) \quad \lambda_k + \bar{\lambda}_k = 0,$$

tehát a sajátértékek tiszta képzetes számok. ■

Megjegyezzük, hogy a valós elemű szimmetrikus mátrixok az hermitikus mátrixok speciális esetét alkotják, tehát ezek sajátértékei is mindig valós számok. Ezen felül – mivel a sajátvektorokat meghatározó egyenletrendszer együtthatói valós számok – a sajátvektorok elemei szintén valós számok. Ez azt jelenti, hogy a valós szimmetrikus mátrixok modálmátrixa valós elemű unitér mátrix – azaz *ortogonális* mátrix –, és ezért a valós szimmetrikus mátrixok ortogonális transzformációval diagonalizálhatók. Tehát ha \mathbf{A} valós elemű szimmetrikus mátrix és \mathbf{T} a modálmátrixa, akkor

$$(3.1.46) \quad \mathbf{A} = \mathbf{T} \mathbf{\Lambda} \mathbf{T}^T,$$

ahol $\mathbf{\Lambda} = \langle \lambda_k \rangle$ a valós sajátértékekből alkotott diagonálmátrix. A (3.1.46) összefüggésből az is kiolvasható, hogy valós szimmetrikus mátrixok jobb és bal oldali sajátvektorai megegyeznek.



3.1 EGYSZERŰ STRUKTÚRÁJÚ MÁTRIXOK SPEKTRÁLIS TULAJDONSÁGAI 267

Összefoglalva a kapott eredményeket, az ábrán szemléltetjük az egyes speciális normális mátrixok sajátértékeinek elhelyezkedését a komplex számsíkon.

3.1.4 Mátrixok szinguláris értékek szerinti felbontása

A következőkben megmutatjuk, hogy $\mathbf{A}^H \mathbf{A}$, illetve $\mathbf{A} \mathbf{A}^H$ teljes spektrálfelbontásával hogyan jellemezhető az adott \mathbf{A} mátrix. Bevezetjük a szinguláris értékek fogalmát, és bebizonyítjuk a mátrixok szinguláris értékek szerinti felbontására vonatkozó tételt.

Legyen \mathbf{A} komplex elemű, n -edrendű és r -edrangú négyzetes mátrix, és tekintsük az $\mathbf{A}^H \mathbf{A}$ hermitikus mátrixot. Az $\mathbf{A}^H \mathbf{A}$ mátrix pozitív szemidefinit, ugyanis minden \mathbf{x} vektorra

$$\mathbf{x}^H \mathbf{A}^H \mathbf{A} \mathbf{x} = (\mathbf{A} \mathbf{x})^H \mathbf{A} \mathbf{x} = \mathbf{y}^H \mathbf{y} = |\mathbf{y}|^2 \geq 0,$$

ahol $\mathbf{y} = \mathbf{A} \mathbf{x}$. Ezért az $\mathbf{A}^H \mathbf{A}$ mátrix minden sajátértéke nemnegatív; jelölje ezeket $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$, és legyen

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_n = 0.$$

3.1.2 definíció. Ha \mathbf{A} komplex elemű, n -edrendű és r -edrangú négyzetes mátrix és σ_i^2 ($i = 1, 2, \dots, n$) jelöli a pozitív szemidefinit $\mathbf{A}^H \mathbf{A}$ mátrix sajátértékeit, akkor a

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_n = 0$$

számokat **\mathbf{A} szinguláris értékeinek** nevezzük (lásd [30]).

Ezek után bebizonyítjuk a szinguláris értékek szerinti felbontásra vonatkozó tételt téglalap alakú, komplex elemű mátrixokra (lásd [5], [9]).

3.1.9 tétel. Legyen \mathbf{A} tetszőleges $m \times n$ típusú komplex elemű mátrix, és tegyük fel, hogy $m \geq n$. Ekkor létezik olyan m -edrendű \mathbf{U} és n -edrendű \mathbf{V} unitér mátrix, hogy

$$(3.1.47) \quad \mathbf{A} = \mathbf{U} \mathbf{D} \mathbf{V}^H$$

ahol

$$\mathbf{D} = \left[\begin{array}{ccc} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ \hline & & \mathbf{0} \end{array} \right] \}_{m-n}$$

$m \times n$ típusú valós elemű mátrix, és $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$. A σ_i ($i = 1, 2, \dots, n$) számok az \mathbf{A} mátrix szinguláris értékei, \mathbf{V} az $\mathbf{A}^H \mathbf{A}$ mátrix modálmátrixa, \mathbf{U} pedig az $\mathbf{A} \mathbf{A}^H$ mátrixé.

Bizonyítás. Tekintsük az $\mathbf{A}^H \mathbf{A}$ n -edrendű pozitív szemidefinit mátrixot, és a sajátvektoraiból alkotott unitér modálmátrixot jelölje \mathbf{V} . Ekkor az $\mathbf{A}^H \mathbf{A}$ mátrix unitér transzformációval diagonalizálható: $\mathbf{V}^H \mathbf{A}^H \mathbf{A} \mathbf{V} = \tilde{\mathbf{D}}^2$, ahol $\tilde{\mathbf{D}}$ a nemnegatív szinguláris értékekből alkotott diagonálmátrix:

$$\tilde{\mathbf{D}} = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \end{bmatrix}.$$

Definiáljuk az $m \times n$ típusú \mathbf{F} mátrixot az

$$(3.1.48) \quad \mathbf{F} = \mathbf{A} \mathbf{V}$$

összefüggéssel. Ekkor $\mathbf{F}^H \mathbf{F} = \tilde{\mathbf{D}}^2$, ami diagonálmátrix. Ebből következik, hogy ha $\varrho(\mathbf{A}) = r = n$, akkor az \mathbf{F} mátrix oszlopvektoraiként páronként kielégítik az $\mathbf{f}_i^H \mathbf{f}_i = \sigma_i^2$ ($i = 1, \dots, n$) és $\mathbf{f}_i^H \mathbf{f}_k = 0$ ($i \neq k$; $i, k = 1, \dots, n$) összefüggéseket. Ha $\varrho(\mathbf{A}) = r < n$, akkor az \mathbf{F} mátrix $(r+1)$ -edik, ..., n -edik oszlopvektora zérusvektor. Legyen tehát $\varrho(\mathbf{A}) = r \leq n$. Az \mathbf{F} mátrix oszlopvektorainak segítségével képezzük az

$$(3.1.49) \quad \mathbf{u}_i = \frac{1}{\sigma_i} \mathbf{f}_i \quad (i = 1, \dots, r)$$

nemt teljes unitér vektorrendszert, és egészítsük ki azt $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ teljes unitér vektorrendszerré (lásd a 3.1.3 tételt és az utána következő alkalmazását). Vezessük be az $\mathbf{U} = [\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_m]$ unitér mátrixot: $\mathbf{U}^H \mathbf{U} = \mathbf{E}$. A (3.1.48) és (3.1.49) összefüggésekből következik, hogy $\mathbf{F} = \mathbf{A} \mathbf{V} = \mathbf{U} \mathbf{D}$, ahol

$$\mathbf{D} = \left[\begin{array}{c} \tilde{\mathbf{D}} \\ \hline \mathbf{0} \end{array} \right]_{m-n},$$

és innen közvetlenül adódik a (3.1.47) összefüggés.

Végül megmutatjuk, hogy \mathbf{U} oszlopvektoraiként az $\mathbf{A} \mathbf{A}^H$ mátrix sajátvektoraiként, és megadunk két összefüggést az \mathbf{U} és \mathbf{V} mátrixok oszlopvektoraiként.

A (3.1.47) összefüggésből $\mathbf{A}^H = \mathbf{V} \mathbf{D}^T \mathbf{U}^H$ adódik. Ezért

$$\mathbf{A} \mathbf{A}^H \mathbf{U} = \mathbf{U} \left[\begin{array}{c|c} \tilde{\mathbf{D}}^2 & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right]$$

vagy más alakban, $\mathbf{A} \mathbf{A}^H \mathbf{u}_i = \sigma_i^2 \mathbf{u}_i$, tehát az \mathbf{U} mátrix oszlopvektoraiként az $\mathbf{A} \mathbf{A}^H$ mátrix sajátvektoraiként.

Az $\mathbf{A} \mathbf{V} = \mathbf{U} \mathbf{D}$ és $\mathbf{A}^H \mathbf{U} = \mathbf{V} \mathbf{D}^T$ összefüggésekből kapjuk az \mathbf{u}_i és \mathbf{v}_i vektorok közti kapcsolatokat: $\mathbf{A} \mathbf{v}_i = \sigma_i \mathbf{u}_i$ és $\mathbf{A}^H \mathbf{u}_i = \sigma_i \mathbf{v}_i$. ■

3.1 EGYSZERŰ STRUKTÚRÁJÚ MÁTRIXOK SPEKTRÁLIS TULAJDONSÁGAI 269

Megjegyzés. A kapott összefüggések értelmében elegendő az $\mathbf{A}^H \mathbf{A}$ mátrix spektrálfelbontását meghatározni, a \mathbf{v}_i sajátvektorok ismeretében \mathbf{u}_i számítható:

$$\mathbf{u}_i = \frac{1}{\sigma_i} \mathbf{A} \mathbf{v}_i \quad i = 1, 2, \dots, r.$$

A többi szinguláris vektort ($i = r + 1, r + 2, \dots, m$) a nemteljes unitér rendszer teljessé tételével kaphatjuk meg.

A 3.1.9 tétel speciális eseteként megfogalmazzuk a következő tételt, amelynek a bizonyítása a 3.1.9 tételéhez hasonlóan végezhető el.

3.1.10 tétel. *Ha \mathbf{A} valós elemű, n -edrendű négyzetes mátrix, akkor létezik olyan valós ortogonális \mathbf{U} és \mathbf{V} mátrix, hogy $\mathbf{A} = \mathbf{U} \mathbf{D} \mathbf{V}^T$, ahol a \mathbf{D} diagonál-mátrix elemei \mathbf{A} szinguláris értékei, az \mathbf{U} és \mathbf{V} ortogonális mátrixok pedig az $\mathbf{A} \mathbf{A}^T$, illetve az $\mathbf{A}^T \mathbf{A}$ valós szimmetrikus pozitív szemidefinit mátrixok modálmátrixai.*

Megemlítjük még a 3.1.10 tétel néhány következményét. Ha $\varrho(\mathbf{A}) = r = n$, akkor \mathbf{D} és \mathbf{A} nonszinguláris, és inverzük előállítható a

$$\mathbf{D}^{-1} = \langle \sigma_1^{-1} \quad \sigma_2^{-1} \dots \sigma_n^{-1} \rangle, \quad \text{ill.} \quad \mathbf{A}^{-1} = \mathbf{V} \mathbf{D}^{-1} \mathbf{U}^T$$

alakban. Az \mathbf{A}^{-1} mátrix szinguláris értékei tehát $\sigma_1^{-1}, \sigma_2^{-1}, \dots, \sigma_n^{-1}$.

Az \mathbf{A} mátrix determinánsának abszolút értékét is kiszámíthatjuk a szinguláris értékek segítségével. Mivel $|\mathbf{A}| = |\mathbf{U}| \cdot |\mathbf{D}| \cdot |\mathbf{V}^T|$, az ortogonális mátrixok determinánsa pedig 1 vagy -1 , ezért $|\mathbf{A}| = \pm |\mathbf{D}| = \pm \sigma_1 \cdot \sigma_2 \cdot \dots \cdot \sigma_n$.

A 3.1.9 tétel segítségével tetszőleges mátrix pszeudoinverze (lásd az 1.6.8 definíciót) előállítható szinguláris értékek szerinti felbontásban. Az erre vonatkozó tétel bizonyításához felhasználjuk, hogy a

$$(3.1.50) \quad \mathbf{D} = \left\{ \begin{array}{c} \left[\begin{array}{ccc|ccc} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ \hline & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{array} \right] \\ \underbrace{\hspace{10em}}_n \end{array} \right\}_m^r$$

3.1 EGYSZERŰ STRUKTÚRÁJÚ MÁTRIXOK SPEKTRÁLIS TULAJDONSÁGAI 271

Helyettesítsük be a (3.1.52) mátrixot és vezessük be az \mathbf{U} mátrix első r oszlopából alkotott minormátrixra az \mathbf{U}_r jelölést; ezzel $\mathbf{A}\mathbf{A}^\dagger = \mathbf{U}_r\mathbf{U}_r^H$ adódik, vagyis $\mathbf{A}\mathbf{A}^\dagger$ hermitikus projektor. A fentiekhez hasonlóan $\mathbf{A}^\dagger\mathbf{A} = \mathbf{V}\mathbf{D}^\dagger\mathbf{U}^H\mathbf{U}\mathbf{D}\mathbf{V}^H = \mathbf{V}\mathbf{D}^\dagger\mathbf{D}\mathbf{V}^\dagger$, majd bevezetve \mathbf{V} első r oszlopából alkotott minormátrixra a \mathbf{V}_r jelölést és behelyettesítve a (3.1.53) mátrixot, $\mathbf{A}^\dagger\mathbf{A} = \mathbf{V}_r\mathbf{V}_r^H$ írható, vagyis $\mathbf{A}^\dagger\mathbf{A}$ is hermitikus projektor. ■

1. Példa. Határozzuk meg az

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 3 \\ 2 & 2 & -2 \\ -1 & 2 & 2 \end{bmatrix}$$

mátrix szinguláris értékek szerinti felbontását!

Megoldás. Először határozzuk meg az $\mathbf{A}^\top\mathbf{A}$ szimmetrikus mátrix modál-mátrixát. Az

$$\mathbf{A}^\top\mathbf{A} = \begin{bmatrix} 9 & 2 & 0 \\ 2 & 8 & 0 \\ 0 & 0 & 17 \end{bmatrix}$$

mátrix sajátértékei 3 tizedesjegy pontossággal $\lambda_1 = 17$, $\lambda_2 = 10,561$, $\lambda_3 = 6,438$, és a hozzájuk tartozó normált (egységnyi abszolút értékű) sajátvektorok:

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0,788 \\ 0,615 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -0,615 \\ 0,788 \\ 0 \end{bmatrix}.$$

Tehát az $\mathbf{A}^\top\mathbf{A}$ mátrix modál-mátrixa $\mathbf{V} = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3]$. Az \mathbf{A} mátrix szinguláris értékei $\sigma_1 = 4,123$, $\sigma_2 = 3,250$, $\sigma_3 = 2,537$. Az \mathbf{U} mátrix előállításához használjuk fel az $\mathbf{A}\mathbf{v}_i = \sigma_i\mathbf{u}_i$ összefüggést. Így

$$\mathbf{u}_1 = \begin{bmatrix} 0,728 \\ -0,485 \\ 0,485 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 0,485 \\ 0,863 \\ 0,136 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -0,485 \\ 0,136 \\ 0,863 \end{bmatrix}.$$

Az \mathbf{U} modál-mátrix tehát $\mathbf{U} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3]$. Az \mathbf{A} mátrix szinguláris értékek szerinti felbontása

$$\begin{aligned} \mathbf{A} &= \mathbf{U}\mathbf{D}\mathbf{V}^\top = \\ &= \begin{bmatrix} 0,728 & 0,485 & -0,485 \\ -0,485 & 0,863 & 0,136 \\ 0,485 & 0,136 & 0,863 \end{bmatrix} \begin{bmatrix} 4,133 & 0 & 0 \\ 0 & 3,250 & 0 \\ 0 & 0 & 2,537 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0,788 & 0,615 & 0 \\ -0,615 & 0,788 & 0 \end{bmatrix}. \end{aligned}$$

* * *

3.2 A MÁTRIXFÜGGVÉNY FOGALMA ÉS ELŐÁLLÍTÁSA A MINIMÁLPOLINOM EGYSZERES GYÖKEI ESETÉN

Már az 1.2.2 pontban értelmeztük a kvadratikus mátrixok hatványát és ennek alapján a mátrixpolinomokat. Ebben a szakaszban bebizonyítjuk a továbbiakban alapvetően fontos Cayley–Hamilton-tételt. Mátrixpolinomok segítségével – hatványsorba fejthető skalárfüggvényekből kiindulva – általánosabb típusú mátrixfüggvényeket értelmezzünk. Ezt az teszi lehetővé, hogy kvadratikus mátrix tetszőleges fokszerű polinomja redukálható a minimálpolinomnál alacsonyabb fokú mátrixpolinomra. Ily módon értelmezhetjük a mátrix-hatványsor konvergenciáját és ezzel a megfelelő mátrixfüggvényt is. (Lásd pl. [26].)

3.2.1 A Cayley–Hamilton-tétel és élesítése

Minden n -edrendű kvadratikus mátrix kielégít egy legfeljebb n^2 -fokú egyenletet. Ha ugyanis adott \mathbf{A} mátrixra felírjuk az alábbi egyenletet:

$$c_0 \mathbf{E} + c_1 \mathbf{A} + c_2 \mathbf{A}^2 + \cdots + c_{n^2-1} \mathbf{A}^{n^2-1} + \mathbf{A}^{n^2} = \mathbf{0},$$

ahol a $c_0, c_1, \dots, c_{n^2-1}$ együtthatók egyelőre ismeretlenek, akkor ezek meghatározására – az egyenletet a mátrix minden elemére külön felírva – éppen n^2 egyenletből álló lineáris egyenletrendszert kapunk. Tehát az adott \mathbf{A} mátrix ezt a legfeljebb n^2 -fokú egyenletet biztosan kielégíti. Kérdés, hogy létezik-e n^2 -nél alacsonyabb fokú egyenlet, amelyiket a mátrix kielégít. Erre a választ a Cayley–Hamilton-tétel adja meg.

3.2.1 tétel (Cayley–Hamilton*-tétel). *Minden kvadratikus mátrix kielégíti karakterisztikus egyenletét, azaz ha*

$$(3.2.1) \quad D(\lambda) \equiv \prod_{k=1}^s (\lambda - \lambda_k)^{\alpha_k} \quad \text{ahol} \quad \sum_{k=1}^s \alpha_k = n,$$

és $\lambda_1, \dots, \lambda_s$ az \mathbf{A} mátrix különböző sajátértékei, akkor

$$(3.2.2) \quad D(\mathbf{A}) = \mathbf{0}.$$

Bizonyítás. Írjuk fel a karakterisztikus mátrix adjungáltját:

$$(3.2.3) \quad \text{adj}(\lambda \mathbf{E} - \mathbf{A}) = [D_{ij}(\lambda)],$$

*A. Cayley (1821–1895) angol matematikus; W. R. Hamilton (1805–1865) ír matematikus.

ahol a $D_{ij}(\lambda)$ elemek λ -nak legfeljebb $(n-1)$ -edfokú polinomjai. Jelölje \mathbf{A}_k azt a mátrixot, amelynek ij indexű eleme a $D_{ij}(\lambda)$ polinomban λ^k együtthatója. Ekkor a karakterisztikus mátrix adjungáltja felírható λ mátrixegyütthatós polinomjaként:

$$(3.2.4) \quad \text{adj}(\lambda \mathbf{E} - \mathbf{A}) = \mathbf{A}_0 + \lambda \mathbf{A}_1 + \lambda^2 \mathbf{A}_2 + \cdots + \lambda^{n-1} \mathbf{A}_{n-1}.$$

Mivel bármely mátrix felcserélhető az adjungáltjával [vö. (1.2.27), (1.2.28)], ezért felírható az alábbi azonosság.

$$(3.2.5) \quad (\lambda \mathbf{E} - \mathbf{A}) \cdot \text{adj}(\lambda \mathbf{E} - \mathbf{A}) \equiv \text{adj}(\lambda \mathbf{E} - \mathbf{A}) \cdot (\lambda \mathbf{E} - \mathbf{A});$$

ebbe az azonosságba behelyettesítve a (3.2.4) összefüggést, az azonos kitevőjű hatványok együtthatóinak összehasonlításával azt kapjuk, hogy az \mathbf{A} mátrix valamennyi \mathbf{A}_k ($k = 1, 2, \dots, n-1$) mátrixszal felcserélhető. Ha most az (1.2.27) szerinti

$$(3.2.6) \quad (\lambda \mathbf{E} - \mathbf{A}) \text{adj}(\lambda \mathbf{E} - \mathbf{A}) \equiv D(\lambda) \mathbf{E}$$

azonosságban λ helyére az \mathbf{A} mátrixot helyettesítjük – amit éppen az \mathbf{A} és az \mathbf{A}_k mátrixok felcserélhetősége miatt megtehetünk –, akkor a (3.2.6) egyenlet bal oldalán zérust kapunk:

$$(3.2.7) \quad (\mathbf{A} \mathbf{E} - \mathbf{A})(\mathbf{A}_0 + \mathbf{A} \mathbf{A}_1 + \mathbf{A}^2 \mathbf{A}_2 + \cdots + \mathbf{A}^{n-1} \mathbf{A}_{n-1}) = D(\mathbf{A}) \cdot \mathbf{E},$$

tehát $D(\mathbf{A}) = \mathbf{0}$, amit bizonyítani kellett. ■

Ezek után azt a kérdést tehetjük fel, vajon létezik-e a mátrix rendszámánál alacsonyabb fokú egyenlet is, amelyet a mátrix kielégít. Hogy bizonyos mátrixokra létezik ilyen, az következik például a projektormátrixok definíciójából: ugyanis a rendszámától függetlenül minden projektormátrix kielégíti a $\mathbf{P}^2 - \mathbf{P} = \mathbf{0}$ másodfokú egyenletet.

Ahhoz, hogy a kérdésre választ kapjunk, vezessük be a *minimálpolinom* fogalmát.

3.2.1 definíció. Ha $D(\lambda)$ az \mathbf{A} mátrix karakterisztikus polinomja és $\theta(\lambda)$ a (3.1.19) adjungált mátrix $D_{ij}(\lambda)$ elemeinek legnagyobb közös osztója, akkor a

$$(3.2.8) \quad \Delta(\lambda) = \frac{D(\lambda)}{\theta(\lambda)}$$

polinomot az \mathbf{A} mátrix **minimálpolinomjának**, a $\Delta(\lambda) = 0$ egyenletet az \mathbf{A} mátrix **minimálegyenletének** nevezzük.

Megjegyzés. A minimálegyenlet a legalacsonyabb fokú egyenlet, amelyet egy mátrix kielégít.

3.2.2 definíció. Az

$$(3.2.9) \quad \mathbf{F}(\lambda) = \frac{\text{adj}(\lambda \mathbf{E} - \mathbf{A})}{\theta(\lambda)}$$

mátrixot a karakterisztikus mátrix **redukált adjungáltjának** nevezzük.

Mivel a $D(\lambda)$ minimálpolinom az $\text{adj}(\lambda \mathbf{E} - \mathbf{A})$ mátrix $D_{ij}(\lambda)$ elemeinek homogén lineáris kombinációja, ezért biztosan osztható ezek közös osztói-val, így a $\theta(\lambda)$ legnagyobb közös osztóval is, tehát $\Delta(\lambda)$ valóban polinom. A redukált adjungált elemei polinomok, ezért (3.2.4) mintájára felírható λ mátrixegyütthatós polinomjaként:

$$(3.2.10) \quad \mathbf{F}(\lambda) = \mathbf{F}_0 + \lambda \mathbf{F}_1 + \lambda^2 \mathbf{F}_2 + \cdots + \lambda^q \mathbf{F}_q,$$

ahol q a maximális fokszámú polinomelem fokszáma.

Ezek után bebizonyítható a 3.2.1 tétel élesítése:

3.2.2 tétel. Minden mátrix kielégíti saját redukált karakterisztikus egyenletét.

Bizonyítás. A tétel bizonyítása egészen hasonlóan végezhető el mint a 3.2.1 tételé. A (3.2.6) egyenlet mindkét oldalát elosztjuk az adjungált mátrix elemeinek legnagyobb közös osztójával, a $\theta(\lambda)$ polinommal:

$$(3.2.11) \quad (\lambda \mathbf{E} - \mathbf{A}) \frac{\text{adj}(\lambda \mathbf{E} - \mathbf{A})}{\theta(\lambda)} \equiv \frac{D(\lambda)}{\theta(\lambda)} \mathbf{E}.$$

A (3.2.8)–(3.2.10) összefüggések behelyettesítése után látható, hogy mivel (3.2.11) a λ skálárban racionális egész azonosság, ahol az \mathbf{A} mátrix valamennyi \mathbf{F}_k mátrixszal felcserélhető, ezért λ helyére beírható az \mathbf{A} mátrix, és így

$$(3.2.12) \quad \Delta(\mathbf{A}) = \mathbf{0}$$

adódik. ■

A minimálpolinom fontos tulajdonsága azonban, hogy nem „veszítheti el” a karakterisztikus polinom egyetlen gyökét sem, vagyis a karakterisztikus polinom valamennyi gyöke a minimálpolinomnak is gyöke. Ezt mondja ki a következő tétel.

3.2.3 tétel. Ha az n -edrendű \mathbf{A} mátrix karakterisztikus polinomja

$$D(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{\alpha_k} \quad \text{ahol} \quad \sum_{k=1}^s \alpha_k = n,$$

minimálpolinomja pedig

$$\Delta(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{\gamma_k}, \quad \text{ahol} \quad \sum_{k=1}^s \gamma_k = m \leq n,$$

akkor $\gamma_k \geq 1$ ($k = 1, \dots, s$).

Bizonyítás. Jelölje $\theta(\lambda)$ az $\text{adj}(\lambda \mathbf{E} - \mathbf{A}) = [D_{ij}(\lambda)]$ elemeinek legnagyobb közös osztóját. Mivel a karakterisztikus polinom osztható $\theta(\lambda)$ -val, ezért $\theta(\lambda)$ gyökei a λ_k sajátértékek $\beta_k \leq \alpha_k$ multiplicitással:

$$\theta(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{\beta_k}.$$

A (3.1.20) összefüggés szerint viszont a karakterisztikus polinom deriváltja

$$D'(\lambda) = \sum_{i=1}^s D_{ii}(\lambda),$$

és ebből következik, hogy a $D_{ii}(\lambda)$ minorok közös osztója a $D'(\lambda)$ deriváltnak is osztója. Mivel pedig λ_k a $D(\lambda)$ polinom α_k multiplicitású gyöke, ugyanez a $D'(\lambda)$ polinomnak $\alpha_k - 1$ multiplicitású gyöke. Ez azt jelenti, hogy $\beta_k \leq \alpha_k - 1$, azaz $\gamma_k \geq 1$. ■

3.2.2 A mátrixfüggvény értelmezése és redukciója mátrixpolinomra

Tekintsünk egy $f(z)$ komplex változós függvényt, amely az origó körüli R sugarú kör belsejében $\sum_{k=0}^{\infty} c_k z^k$ alakú hatványsorba fejthető. Vagyis, ha $S_N(z)$ e hatványsor N -edik részletösszegét jelenti, akkor létezik a

$$(3.2.13) \quad \lim_{N \rightarrow \infty} S_N(z) = \lim_{N \rightarrow \infty} \sum_{k=0}^N c_k z^k$$

határérték a $|z| < R$ nyílt körlemez minden belső pontjában és ez a határérték egyenlő az e pontbeli $f(z)$ függvényértékkel.

3.2.3 definíció. Ha \mathbf{A} kvadratikus mátrixot jelöl, akkor az $f(\mathbf{A})$ mátrixfüggvényt az

$$(3.2.14) \quad f(\mathbf{A}) = \lim_{N \rightarrow \infty} S_N(\mathbf{A}) = \lim_{N \rightarrow \infty} \sum_{k=0}^N c_k \mathbf{A}^k$$

határértékkel értelmezzük, feltéve, hogy ez létezik.

Az $S_N(\mathbf{A})$ mátrixpolinom-sorozat konvergenciáját úgy vizsgáljuk, hogy az N -edfokú részletösszeget a minimálpolinomnál alacsonyabb fokú polinomra redukáljuk, és így a határátmenetet csupán számsorozatokon kell elvégeznünk.

Az \mathbf{A} mátrixot illetően különbséget kell tennünk aszerint, hogy minimálpolinomjának csupa egyszeres gyökei vannak-e, vagy pedig van legalább egy többszörös gyöke is. A gyakorlati problémák túlnyomó többségében az első eset fordul elő. A következőkben csak ezt az esetet vizsgáljuk. Ehhez szükségünk van a Lagrange-féle interpolációs alappolinomokra.

Mint ismeretes, az

$$(3.2.15) \quad L_k(\lambda_l) = \delta_{kl} \quad (k, l = 1, 2, \dots, s)$$

összefüggésekkel definiált Lagrange-féle interpolációs alappolinomok az alábbi alakban írhatók:

$$(3.2.16) \quad L_k(z) = \frac{\Delta(z)}{\Delta'(\lambda_k)(z - \lambda_k)}, \quad \text{ahol} \quad \Delta(z) = \prod_{k=1}^s (z - \lambda_k)$$

és $\Delta'(\lambda_k)$ a $\Delta(z)$ polinom deriváltjának értékét jelenti a $z = \lambda_k$ helyen.

3.2.4 tétel. *Ha az \mathbf{A} mátrix karakterisztikus polinomja*

$$(3.2.17) \quad D(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{\alpha_k}, \quad \text{ahol} \quad \sum_{k=1}^s \alpha_k = n,$$

minimálpolinomja pedig

$$(3.2.18) \quad \Delta(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k),$$

és \mathbf{A} sajátértékei az $f(z)$ függvényt előállító hatványsor (R sugarú) konvergenciakörének belsejébe esnek, azaz

$$(3.2.19) \quad |\lambda_k| < R,$$

akkor a (3.2.14) határértékkel definiált $f(\mathbf{A})$ mátrixfüggvény létezik, és a minimálpolinom fokszámanál alacsonyabb fokú mátrixpolinomra redukálható. Mégpedig, ha $L_k(z)$ a (3.2.18) minimálpolinom gyökhelyeihez mint alappolinomokhoz tartozó Lagrange-féle interpolációs alappolinomokat jelöli, akkor*

$$(3.2.20) \quad f(\mathbf{A}) = \sum_{k=1}^s f(\lambda_k) L_k(\mathbf{A}).$$

* J. L. Lagrange (1736–1813) francia matematikus.

Bizonyítás. Legyen $N > s$ és képezzük az $S_N(z)$ részletösszegnek a minimálpolinomra vett osztási maradékát:

$$(3.2.21) \quad S_N(z) \equiv \Delta(z)q(z) + R_N(z),$$

ahol $R_N(z)$ egy legfeljebb $(s-1)$ -edfokú polinom. A $z = \lambda_k$ helyettesítéssel a (3.2.21) összefüggésből

$$(3.2.22) \quad S_N(\lambda_k) = R_N(\lambda_k) \quad (k = 1, 2, \dots, s)$$

adódik, és ezáltal az $R_N(z)$ polinom egyértelműen meghatározható a minimálpolinom gyökhelyeihez tartozó Lagrange-féle alappolinomok segítségével. Mivel a (3.2.22) feltételek egyértelműen meghatározzák az $R_N(z)$ polinomot, ezért az $L_k(z)$ alappolinomoknak és a (3.2.22) feltételeknek a segítségével a legfeljebb $(s-1)$ -ed fokú $R_N(z)$ polinom azonosan előállítható a következőképpen:

$$(3.2.23) \quad R_N(z) \equiv \sum_{k=1}^s S_N(\lambda_k) L_k(z).$$

Ha most a (3.2.21) összefüggésben, amely a z skalárváltozóban racionális egész azonosság, behelyettesítjük z helyére az \mathbf{A} mátrixot, akkor – felhasználva, hogy a 3.2.2 tétel értelmében $\Delta(\mathbf{A}) = \mathbf{0}$ – azt kapjuk, hogy

$$(3.2.24) \quad S_N(\mathbf{A}) = R_N(\mathbf{A});$$

majd a (3.2.23) összefüggés felhasználásával

$$(3.2.25) \quad S_N(\mathbf{A}) = \sum_{k=1}^s S_N(\lambda_k) L_k(\mathbf{A}).$$

Ezzel elértük azt, hogy az N -edfokú részletösszegben előforduló mátrixpolinomok már nem függenek N -től, ezért a határátmenetet csak az $S_N(\lambda_k)$ számsorozatokon kell elvégezni. Tehát a (3.2.25) mátrixpolinom-sorozat konvergens, ha teljesül a (3.2.19) feltétel. Ekkor a (3.2.25) összefüggés jobb oldalán a határátmenet tagonként elvégezhető, és így a keresett mátrixfüggvényre

$$f(\mathbf{A}) = \lim_{N \rightarrow \infty} S_N(\mathbf{A}) = \sum_{k=1}^s \lim_{N \rightarrow \infty} S_N(\lambda_k) L_k(\mathbf{A}),$$

és innen (3.2.20)

$$f(\mathbf{A}) = \sum_{k=1}^s f(\lambda_k) L_k(\mathbf{A})$$

adódik. Így az $f(\mathbf{A})$ mátrixfüggvényt sikerült egy legfeljebb $(s-1)$ -edfokú mátrixpolinomra redukálni. ■

3.2.3 A Lagrange-féle mátrixpolinomok tulajdonságai

A kapott $L_k(\mathbf{A})$ mátrixpolinomokat *Lagrange-féle mátrixpolinomoknak* nevezzük és a következőkben ezek tulajdonságait vizsgáljuk meg.

3.2.5 tétel. Az $L_k(\mathbf{A})$ mátrixpolinom projektor.

Bizonyítás. A Lagrange-féle alappolinom (3.2.15) definíciója alapján az $L_k(z)(L_k(z) - 1)$ szorzat osztható a minimálpolinommal, azaz

$$(3.2.26) \quad L_k(z)(L_k(z) - 1) = \Delta(z)q_1(z).$$

A z változó helyére az \mathbf{A} mátrixot helyettesítve és felhasználva a 3.2.2 tételt, azt kapjuk, hogy

$$(3.2.27) \quad L_k(\mathbf{A})(L_k(\mathbf{A}) - \mathbf{E}) = \mathbf{0}, \quad \text{vagyis} \quad L_k(\mathbf{A})^2 = L_k(\mathbf{A}),$$

amit bizonyítani akartunk. ■

3.2.6 tétel. Az $L_k(\mathbf{A})$ projektor rangja α_k .

Bizonyítás. A Lagrange-féle alappolinomok (3.2.16) előállítására alapján

$$(3.2.38) \quad L_k(z)(z - \lambda_k) = \frac{\Delta(z)}{\Delta'(\lambda_k)}.$$

A z változó helyére az \mathbf{A} mátrixot helyettesítve és ismét felhasználva a 3.2.2 tételt, a következő adódik:

$$(3.2.29) \quad L_k(\mathbf{A})(\mathbf{A} - \lambda_k \mathbf{E}) = \mathbf{0}.$$

Mivel a 3.1.1 tétel értelmében

$$(3.2.30) \quad \varrho(\mathbf{A} - \lambda_k \mathbf{E}) \geq n - \alpha_k,$$

és a (3.2.29) szorzat csak akkor lehet 0, ha a tényezők rangjának összege az 1.3.13 tétel alapján nem nagyobb, mint n , tehát

$$(3.2.31) \quad \varrho(L_k(\mathbf{A})) \leq \alpha_k.$$

Másrészt az $f(z) \equiv 1$ függvényt – amely $(s - 1)$ -nél nem magasabb fokszámú polinom – a Lagrange-féle interpolációs alappolinomok segítségével előállítva, felírható, hogy

$$(3.2.32) \quad \sum_{k=1}^s L_k(z) \equiv 1,$$

amiből

$$(3.2.33) \quad \sum_{k=1}^s L_k(\mathbf{A}) = \mathbf{E}.$$

Alkalmazva az 1.3.5 tételt azt kapjuk, hogy

$$(3.2.34) \quad n = \varrho(\mathbf{E}) \leq \sum_{k=1}^s \varrho(L_k(\mathbf{A})).$$

Innen következik, hogy ha a (3.2.31) összefüggésben csak egyetlen k esetén is az egyenlőség nem állna fenn, akkor ellentmondásba jutnánk a (3.2.34) összefüggéssel. Tehát minden k -ra fennáll a

$$(3.2.35) \quad \varrho(L_k(\mathbf{A})) = \alpha_k$$

összefüggés. ■

3.2.7 tétel. *Különböző indexű Lagrange-féle mátrixpolinomok szorzata zérus:*

$$(3.2.36) \quad L_k(\mathbf{A}) \cdot L_l(\mathbf{A}) = \mathbf{0}, \quad \text{ha } k \neq l.$$

Bizonyítás. A (3.2.15) összefüggés alapján különböző indexű Lagrange-féle alappolinomok szorzata osztható a minimálpolinommal, azaz

$$(3.2.37) \quad L_k(z)L_l(z) = \Delta(z)q_2(z);$$

a z változó helyére az \mathbf{A} mátrixot helyettesítve és felhasználva a 3.2.2 tételt, a bizonyítandó (3.2.36) összefüggés adódik. ■

Képezzük az α_k rangú $L_k(\mathbf{A})$ projektormátrix egy minimális diadikus felbontását:

$$(3.2.38) \quad L_k(\mathbf{A}) = \sum_{v=1}^{\alpha_k} \mathbf{u}_{kv} \mathbf{v}_{kv}^T = \mathbf{U}_k \mathbf{V}_k^T \quad (k = 1, 2, \dots, s).$$

Ezt a (3.2.36) összefüggésbe helyettesítve, és felhasználva az 1.3.10 tételt, az $\mathbf{U}_k(\mathbf{V}_k^T \mathbf{U}_l) \mathbf{V}_l^T = \mathbf{0}$ ($k \neq l$) egyenletből

$$(3.2.39) \quad \mathbf{V}_k^T \mathbf{U}_l = \mathbf{0} \quad (k \neq l)$$

következik. Ha most a (3.2.38) előállítást a (3.2.33) összefüggésbe helyettesítjük, akkor az alábbi adódik:

$$(3.2.40) \quad \sum_{k=1}^s \mathbf{U}_k \mathbf{V}_k^T = \mathbf{E}.$$

A (3.2.39) és (3.2.40) összefüggésekből kiolvasható, hogy az összes $L_k(\mathbf{A})$ mátrixpolinom minimális diadikus felbontása teljes biortogonális vektorrendszeret alkot. Bevezetve a (3.2.38) alatti diádok oszlopvektoraiból, ill. sorvektoraiból alkotott \mathbf{U} , ill. \mathbf{V}^T mátrixokat:

$$(3.2.41) \quad \mathbf{U} = [\mathbf{U}_1 \mathbf{U}_2 \dots \mathbf{U}_s]; \quad \mathbf{V}^T = \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \\ \vdots \\ \mathbf{V}_s^T \end{bmatrix},$$

(3.2.40) alapján felírható

$$(3.2.42) \quad \mathbf{U}\mathbf{V}^T = \mathbf{E},$$

és ebből

$$(3.2.43) \quad \mathbf{V}^T = \mathbf{U}^{-1}$$

következik.

3.2.4 Mátrixfüggvény spektrálfelbontása

Helyettesítsük be a (3.2.38) diádokat a mátrixpolinommal redukált $f(\mathbf{A})$ mátrixfüggvény (3.2.20) alatti kifejezésébe; az így adódó eredmény a következőképpen interpretálható.

A (3.2.20) összefüggés segítségével az $f(\mathbf{A})$ mátrixfüggvényt olyan projektorok lineáris kombinációjaként állítottuk elő, amelyek diadikus felbontása teljes biortogonális vektorrendszert eredményez. Ezek az oszlopvektorok, ill. sorvektorok tehát az $f(\mathbf{A})$ mátrixfüggvény jobb, ill. bal oldali sajátvektorai, és így nemcsak azt bizonyítottuk be, hogy az $f(\mathbf{A})$ mátrixfüggvény – a 3.2.4 tétel feltételei mellett – hasonlósági transzformációval diagonalizálható, hanem ténylegesen előállítottuk a modálmátrixot is, amellyel ez a transzformáció elvégezhető. Ezt a jelentős eredményt a következő tételben fogalmazzuk meg.

3.2.8 tétel. *Ha az \mathbf{A} mátrix karakterisztikus polinomja*

$$D_n(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{\alpha_k}, \quad \text{ahol} \quad \sum_{k=1}^s \alpha_k = n,$$

minimálpolinomja

$$\Delta(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k),$$

és a λ_k sajátértékek az $f(z)$ függvényt előállító hatványsor (R sugarú) konvergenciakörének belsejébe esnek, vagyis $|\lambda_k| < R$, akkor a λ_k gyökhelyekhez tartozó

$$L_k(z) = \frac{\Delta(z)}{\Delta'(\lambda_k)(z - \lambda_k)}$$

Lagrange-féle alappolinomok segítségével képezett $L_k(\mathbf{A})$ mátrixpolinomok α_k rangú projektorok, amelyeknek

$$L_k(\mathbf{A}) = \sum_{v=1}^{\alpha_k} \mathbf{u}_{kv} \mathbf{v}_{kv}^T$$

alakú diadikus felbontása az $f(\mathbf{A})$ mátrixfüggvény jobb és bal oldali sajátvektorait adja meg; az $f(\mathbf{A})$ mátrixfüggvény spektrálfelbontása tehát

$$(3.2.44) \quad f(\mathbf{A}) = \sum_{k=1}^s \sum_{v=1}^{\alpha_k} f(\lambda_k) \mathbf{u}_{kv} \mathbf{v}_{kv}^T$$

alakban írható fel.

Ebből következik, hogy bevezetve az $f(\lambda_k)$ elemekből alkotott $f(\mathbf{A})$ diagonálmátrixot, ahol minden $f(\lambda_k)$ elem α_k -szor szerepel, az $f(\mathbf{A})$ mátrixfüggvény az \mathbf{u}_{kv} sajátvektorokból alkotott \mathbf{U} modálmátrixszal képzett hasonlósági transzformációval diagonalizálható:

$$(3.2.45) \quad f(\mathbf{A}) = \mathbf{U}^{-1} f(\mathbf{A}) \mathbf{U}.$$

Abban a speciális esetben, amikor $f(z) \equiv z$, megkapjuk az \mathbf{A} mátrix spektrálfelbontását, ill. a sajátértékeiből alkotott diagonálmátrixra való transzformációját:

$$(3.2.46) \quad \mathbf{A} = \mathbf{U} \mathbf{A} \mathbf{U}^{-1}.$$

Tehát az a feltétel, hogy a minimálpolinom gyökei egyszeresek, a mátrix diagonalizálhatóságának *elégséges feltétele*. A 3.2.8 tétel tulajdonképpen ennél még többet mond, azt ugyanis, hogy ebben az esetben nemcsak a mátrix diagonalizálható, hanem a mátrixnak bármilyen – a konvergenciafeltételeket kielégítő – hatványsorral definiált függvénye is. A transzformáció mátrixa ugyanez a modálmátrix, tehát a mátrixnak és függvényének közös a sajátvektorrendszere, a mátrix függvényének sajátértékei pedig a mátrix sajátértékeinek a függvényei (a megfelelő multiplicitással).

Most néhány példát mutatunk arra, hogyan lehet konkrét esetekben meghatározni egy diagonalizálható mátrix függvényét, ill. spektrálfelbontását.

2. Példa. Meghatározandó $\sin \mathbf{A}$, valamint spektrálfelbontása, ha

$$\mathbf{A} = \begin{bmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{bmatrix}.$$

Megoldás. A karakterisztikus egyenlet

$$\begin{aligned}
 |\lambda \mathbf{E} - \mathbf{A}| &= \begin{vmatrix} \lambda - 3 & 3 & -2 \\ 1 & \lambda - 5 & 2 \\ 1 & -3 & \lambda \end{vmatrix} = \\
 &= (\lambda - 3)[\lambda(\lambda - 5) + 6] - (3\lambda - 6) + (6 + 2(\lambda - 5)) = \\
 &= (\lambda - 3)^2(\lambda - 2) - 3(\lambda - 2) + 2(\lambda - 2) = (\lambda - 2)(\lambda^2 - 6\lambda + 8) = \\
 &= (\lambda - 2)^2(\lambda - 4) = 0.
 \end{aligned}$$

Az utolsó alakból leolvasható, hogy $\lambda_1 = 2$ kétszeres sajátérték, $\lambda_2 = 4$ pedig egyszeres sajátérték.

A karakterisztikus mátrix adjungáltja:

$$\text{adj}(\lambda \mathbf{E} - \mathbf{A}) = \begin{bmatrix} (\lambda - 2)(\lambda - 3) & -3(\lambda - 2) & 2(\lambda - 2) \\ -(\lambda - 2) & (\lambda - 2)(\lambda - 1) & -2(\lambda - 2) \\ -(\lambda - 2) & 3(\lambda - 2) & (\lambda - 2)(\lambda - 6) \end{bmatrix}.$$

Az adjungált mátrix elemeinek legnagyobb közös osztója $(\lambda - 2)$, a minimál-egyenlet tehát

$$\Delta(\lambda) \equiv (\lambda - 2)(\lambda - 4) = 0.$$

Mivel ennek gyökei egyszeresek, a mátrixfüggvényt a gyökhelyekhez tartozó (elsőfokú) Lagrange-féle interpolációs alappolinomok segítségével meghatározhatjuk:

$$L_1(z) = -\frac{1}{2}(z - 4); \quad L_2(z) = \frac{1}{2}(z - 2).$$

Behelyettesítve z helyére az \mathbf{A} mátrixot,

$$\begin{aligned}
 L_1(\mathbf{A}) &= \frac{1}{2}(4\mathbf{E} - \mathbf{A}) = \frac{1}{2} \begin{bmatrix} 1 & 3 & -2 \\ 1 & -1 & 2 \\ 1 & -3 & 4 \end{bmatrix}, \\
 L_2(\mathbf{A}) &= \frac{1}{2}(\mathbf{A} - 2\mathbf{E}) = \frac{1}{2} \begin{bmatrix} 1 & -3 & 2 \\ -1 & 3 & -2 \\ -1 & 3 & -2 \end{bmatrix}.
 \end{aligned}$$

A keresett mátrixfüggvény most már felírható:

$$\sin \mathbf{A} = \frac{\sin 2}{2} \begin{bmatrix} 1 & 3 & -2 \\ 1 & -1 & 2 \\ 1 & -3 & 4 \end{bmatrix} + \frac{\sin 4}{2} \begin{bmatrix} 1 & -3 & 2 \\ -1 & 3 & -2 \\ -1 & 3 & -2 \end{bmatrix} =$$

$$= \frac{1}{2} \begin{bmatrix} \sin 2 + \sin 4 & 3(\sin 2 - \sin 4) & -2(\sin 2 - \sin 4) \\ \sin 2 - \sin 4 & -\sin 2 + 3\sin 4 & 2(\sin 2 - \sin 4) \\ \sin 2 - \sin 4 & -3(\sin 2 - \sin 4) & 2(2\sin 2 - \sin 4) \end{bmatrix}.$$

A mátrixelemek numerikus felírásától eltekintünk. A kapott mátrix spektrálfelbontása a Lagrange-féle mátrixpolinomok minimális diadikus felbontásával nyerhető. Mivel $\lambda_1 = 2$ a karakterisztikus egyenletnek kétszeres gyöke, azért $L_1(\mathbf{A})$ rangja 2, és mivel $\lambda_2 = 4$ egyszeres gyöke, $L_2(\mathbf{A})$ rangja 1. A Lagrange-féle mátrixpolinomok diadikus felbontása

$$L_1(\mathbf{A}) = \frac{1}{2} \begin{bmatrix} 1 & 3 & -2 \\ 1 & -1 & 2 \\ 1 & -3 & 4 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 3 & -2 \\ 0 & -2 & 2 \end{bmatrix},$$

$$L_2(\mathbf{A}) = \frac{1}{2} \begin{bmatrix} 1 & -3 & 2 \\ -1 & 3 & -2 \\ -1 & 3 & -2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & -3 & 2 \end{bmatrix},$$

tehát

$$\sin \mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & -1 \\ 1 & 3 & -1 \end{bmatrix} \begin{bmatrix} \sin 2 & 0 & 0 \\ 0 & \sin 2 & 0 \\ 0 & 0 & \sin 4 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{3}{2} & -1 \\ 0 & -1 & 1 \\ \frac{1}{2} & -\frac{3}{2} & 1 \end{bmatrix}.$$

* * *

3. Példa. Határozzuk meg az

$$\mathbf{A} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

harmadrendű, ferdén szimmetrikus mátrix $e^{\mathbf{A}}$ exponenciális függvényét \mathbf{A} polinomjaként!

Megoldás. A karakterisztikus egyenlet

$$D(\lambda) \equiv |\lambda \mathbf{E} - \mathbf{A}| \equiv \lambda^3 + (a_1^2 + a_2^2 + a_3^2)\lambda = 0;$$

innen az $a = \sqrt{a_1^2 + a_2^2 + a_3^2}$ jelöléssel a sajátértékek:

$$\lambda_1 = 0; \quad \lambda_2 = ia; \quad \lambda_3 = -ia, \quad \text{ahol } i \text{ a képzetes egység.}$$

A Lagrange-féle interpolációs alappolinomok tehát a következők:

$$L_1(z) = \frac{(z - ia)(z + ia)}{a^2},$$

$$L_2(z) = \frac{z(z+ia)}{-2a^2},$$

$$L_3(z) = \frac{z(z-ia)}{-2a^2}.$$

Behelyettesítve z helyére az \mathbf{A} mátrixot, az alábbiakat kapjuk:

$$L_1(\mathbf{A}) = \frac{1}{a^2}(\mathbf{A}^2 + a^2\mathbf{E}),$$

$$L_2(\mathbf{A}) = -\frac{1}{2a^2}(\mathbf{A}^2 + ia\mathbf{A}),$$

$$L_3(\mathbf{A}) = -\frac{1}{2a^2}(\mathbf{A}^2 - ia\mathbf{A}).$$

A keresett exponenciális függvény az \mathbf{A} mátrix másodfokú polinomjaként a következő alakban írható:

$$e^{\mathbf{A}} = e^{\lambda_1} L_1(\mathbf{A}) + e^{\lambda_2} L_2(\mathbf{A}) + e^{\lambda_3} L_3(\mathbf{A}),$$

azaz

$$\begin{aligned} e^{\mathbf{A}} &= \frac{1}{a^2}\mathbf{A}^2 + \mathbf{E} - e^{ia}\left(\frac{1}{2a^2}\mathbf{A}^2 - \frac{1}{2ia}\mathbf{A}\right) - e^{-ia}\left(\frac{1}{2a^2}\mathbf{A}^2 + \frac{1}{2ia}\mathbf{A}\right) = \\ &= \mathbf{E} + \mathbf{A}\frac{\sin a}{a} + \mathbf{A}^2\frac{1 - \cos a}{a^2}. \end{aligned}$$

Egyszerű számolással közvetlenül meggyőződhetünk arról, hogy az $\mathbf{a}^T = [a_1 \ a_2 \ a_3]$ jelöléssel $\mathbf{A}^2 = \mathbf{a}\mathbf{a}^T - a^2\mathbf{E}$; ezt behelyettesítve a megelőző összefüggésbe,

$$e^{\mathbf{A}} = \mathbf{E} \cos a + \mathbf{A}\frac{\sin a}{a} + \frac{\mathbf{a}\mathbf{a}^T}{a^2}(1 - \cos a),$$

azaz

$$e^{\mathbf{A}} = \begin{bmatrix} \cos a + \frac{a_1^2}{a^2}(1 - \cos a) & -\frac{a_3}{a} \sin a + \frac{a_1 a_2}{a^2}(1 - \cos a) & \frac{a_2}{a} \sin a + \frac{a_1 a_3}{a^2}(1 - \cos a) \\ \frac{a_3}{a} \sin a + \frac{a_1 a_2}{a^2}(1 - \cos a) & \cos a + \frac{a_2^2}{a^2}(1 - \cos a) & -\frac{a_1}{a} \sin a + \frac{a_2 a_3}{a^2}(1 - \cos a) \\ -\frac{a_2}{a} \sin a + \frac{a_1 a_3}{a^2}(1 - \cos a) & \frac{a_1}{a} \sin a + \frac{a_2 a_3}{a^2}(1 - \cos a) & \cos a + \frac{a_3^2}{a^2}(1 - \cos a) \end{bmatrix}.$$

* * *

4. Példa. Mutassuk meg, hogy ha \mathbf{P} projektormátrix, akkor

$$e^{\mathbf{P}} = \mathbf{E} + (e - 1)\mathbf{P}.$$

Megoldás. A feladatra kétféle megoldást mutatunk. Az első a mátrixfüggvény definícióján alapszik. Tehát esetünkben az $f(z) = e^z$ exponenciális függvény közismert hatványsorából indulunk ki:

$$e^{\mathbf{P}} = \mathbf{E} + \mathbf{P} + \frac{1}{2}\mathbf{P}^2 + \frac{1}{3!}\mathbf{P}^3 + \cdots + \frac{1}{n!}\mathbf{P}^n + \cdots;$$

mivel \mathbf{P} projektor, így $\mathbf{P} = \mathbf{P}^2 = \mathbf{P}^3 = \cdots = \mathbf{P}^n = \cdots$, tehát

$$e^{\mathbf{P}} = \mathbf{E} + \mathbf{P} \left(1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!} + \cdots \right),$$

azaz $e^{\mathbf{P}} = \mathbf{E} + (e - 1)\mathbf{P}$.

Ugyanerre az eredményre jutunk, ha felhasználjuk a \mathbf{P} projektormátrix spektrálfelbontását (lásd a 3.1.3 tételt).

Legyen \mathbf{P} minimális diadikus felbontása $\mathbf{P} = \mathbf{U}\mathbf{V}^T$, továbbá $\mathbf{E} - \mathbf{P}$ minimális diadikus felbontása $\mathbf{E} - \mathbf{P} = \mathbf{W}\mathbf{Z}^T$. Mivel a \mathbf{P} mátrix 1 sajátértékéhez tartozó jobb, ill. bal oldali sajátvektorait \mathbf{U} oszlopai, ill. \mathbf{V}^T sorai, 0 sajátértékéhez tartozó jobb, ill. bal oldali sajátvektorait pedig \mathbf{W} oszlopai, ill. \mathbf{Z}^T sorai adják, tehát \mathbf{P} spektrálfelbontása

$$\mathbf{P} = [\mathbf{U} \quad \mathbf{W}] \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}^T \\ \mathbf{Z}^T \end{bmatrix}.$$

Innen (3.2.45) szerint a középső tényező diagonálelemeit helyettesítve az exponenciális függvénybe, kapjuk:

$$e^{\mathbf{P}} = [\mathbf{U} \quad \mathbf{W}] \begin{bmatrix} e^1 \mathbf{E} & \mathbf{0} \\ \mathbf{0} & e^0 \mathbf{E} \end{bmatrix} \begin{bmatrix} \mathbf{V}^T \\ \mathbf{Z}^T \end{bmatrix},$$

azaz

$$e^{\mathbf{P}} = e\mathbf{U}\mathbf{V}^T + \mathbf{W}\mathbf{Z}^T = e\mathbf{P} + (\mathbf{E} - \mathbf{P}) = \mathbf{E} + (e - 1)\mathbf{P}.$$

* * *

5. Példa. Határozzuk meg a ciklikus mátrixok spektrálfelbontását!

Megoldás. Minden ciklikus mátrix az Ω elemi ciklikus mátrix polinomjaként állítható elő (lásd (1.4.14)); ezért elegendő, ha Ω spektrálfelbontásának a meghatározására szorítkozunk. Tekintettel arra, hogy az Ω mátrix permutáló mátrix is, tehát ortogonális, ezért unitér transzformációval diagonalizálható (2.8.10 tétel), és azt is tudjuk, hogy sajátértékei a komplex számsík egységkörén helyezkednek el (2.8.6 tétel).

Számítsuk ki először az $\mathbf{\Omega}$ elemi ciklikus mátrix sajátértékeit. A karakterisztikus polinom

$$D(\lambda) \equiv |\lambda \mathbf{E} - \mathbf{\Omega}| = \begin{vmatrix} \lambda & -1 & 0 \dots 0 & 0 \\ 0 & \lambda & -1 \dots 0 & 0 \\ 0 & 0 & \lambda \dots 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 \dots \lambda & -1 \\ -1 & 0 & 0 \dots 0 & \lambda \end{vmatrix}.$$

Ezt az első oszlop szerint kifejtve és zérussal egyenlővé téve, a karakterisztikus egyenlet

$$(3.2.47) \quad D(\lambda) \equiv \lambda^n - 1 = 0,$$

tehát a sajátértékek az n -edik egységgyökök.

Jelölje ezeket a továbbiakban ω_k , vagyis

$$(3.2.48) \quad \omega_k = e^{\frac{2k\pi i}{n}} \quad (k = 0, 1, 2, \dots, n-1),$$

ahol i a képzetes egységet jelöli.

A sajátvektorok előállításához először a Lagrange-féle interpolációs alapolinomokat írjuk fel. Mivel a gyökök valamennyien különböznek egymástól, az alapolinomok az alábbi alakban írhatók:

$$L_k(z) \equiv \frac{D(z)}{D'(\lambda_k)(z - \lambda_k)}.$$

Behelyettesítve a (3.2.47) és (3.2.48) összefüggéseket,

$$L_k(z) \equiv \frac{z^n - 1}{n\omega_k^{n-1}(z - \omega_k)}.$$

Elvégezve az osztást, az alábbi kifejezést kapjuk:

$$\begin{aligned} L_k(z) &= \frac{1}{n\omega_k^{n-1}} (\omega_k^{n-1} + \omega_k^{n-2}z + \dots + \omega_k z^{n-2} + z^{n-1}) = \\ &= \frac{1}{n} (1 + \bar{\omega}_k z + \bar{\omega}_k^2 z^2 + \dots + \bar{\omega}_k^{n-1} z^{n-1}). \end{aligned}$$

Ha most a z skalárváltozó helyére beírjuk az $\mathbf{\Omega}$ mátrixot, akkor

$$L_k(\mathbf{\Omega}) = \frac{1}{n} (\mathbf{E} + \bar{\omega}_k \mathbf{\Omega} + \bar{\omega}_k^2 \mathbf{\Omega}^2 + \dots + \bar{\omega}_k^{n-1} \mathbf{\Omega}^{n-1}),$$

azaz – a ciklikus mátrixokra szokásos jelöléssel (lásd (1.4.13)) –

$$(3.2.49) \quad L_k(\mathbf{\Omega}) = \frac{1}{n} \mathbf{C} (1, \bar{\omega}_k, \bar{\omega}_k^2, \dots, \bar{\omega}_k^{n-1}).$$

Mivel a sajátértékek egyszeresek, ezért a (3.2.49) alatti mátrixpolinom egyetlen diádként írható fel:

$$(3.2.50) \quad L_k(\Omega) = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 \\ \omega_k \\ \omega_k^2 \\ \vdots \\ \omega_k^{n-1} \end{bmatrix} \begin{bmatrix} 1 & \bar{\omega}_k & \bar{\omega}_k^2 & \dots & \bar{\omega}_k^{n-1} \end{bmatrix} \frac{1}{\sqrt{n}}.$$

Ezzel a $\mathbf{C}(c_0, c_1, \dots, c_{n-1})$ ciklikus mátrix spektrálfelbontása:

$$(3.2.51) \quad \mathbf{C} = \mathbf{U} \left\langle \sum_{\nu=0}^{n-1} c_\nu \sum_{\nu=0}^{n-1} c_\nu \omega_1^\nu \sum_{\nu=0}^{n-1} c_\nu \omega_2^\nu \dots \sum_{\nu=0}^{n-1} c_\nu \omega_{n-1}^\nu \right\rangle \mathbf{U}^H, \text{ ahol}$$

$$\mathbf{U} = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & 1 & \dots & 1 & \dots & 1 \\ 1 & \omega_1 & \omega_2 & \dots & \omega_k & \dots & \omega_{n-1} \\ 1 & \omega_1^2 & \omega_2^2 & \dots & \omega_k^2 & \dots & \omega_{n-1}^2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & \omega_1^p & \omega_2^p & \dots & \omega_k^p & \dots & \omega_{n-1}^p \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & \omega_1^{n-1} & \omega_2^{n-1} & \dots & \omega_k^{n-1} & \dots & \omega_{n-1}^{n-1} \end{bmatrix}$$

vagy – tömörebb jelöléssel –

$$(3.2.52) \quad \mathbf{C}(c_0, c_1, c_2, \dots, c_{n-1}) = \frac{1}{\sqrt{n}} \begin{bmatrix} \omega_{k-1}^{p-1} \end{bmatrix} \left\langle \sum_{\nu=0}^{n-1} c_\nu \omega_{k-1}^\nu \right\rangle \begin{bmatrix} \bar{\omega}_{k-1}^{q-1} \end{bmatrix} \frac{1}{\sqrt{n}}.$$

* * *

3.2.5 Lagrange-féle mátrixpolinomok előállítása a karakterisztikus mátrix adjungáltjával

Bizonyos feladatok kapcsán – amikor az \mathbf{A} mátrix szerkezete szabályos, és tetszőleges n rendszám esetén kívánjuk a Lagrange-féle mátrixpolinomokat meghatározni – célszerű ezek kiszámítására a (3.2.16) képlet helyett egy másik eljárást alkalmazni, amely közvetlenül a karakterisztikus mátrix adjungáltjának a segítségével ad explicit képletet az $L_k(\mathbf{A})$ mátrixpolinomra.

Erre vonatkozik a következő tétel.

3.2.9 tétel. *Ha $\Delta(\lambda)$ az \mathbf{A} mátrix csupa egyszeres gyököt tartalmazó minimálpolinomja és $\theta(\lambda)$ az $\text{adj}(\lambda \mathbf{E} - \mathbf{A})$ mátrix elemeinek legnagyobb közös osztója, akkor az $L_k(\mathbf{A})$ mátrixpolinom az*

$$(3.2.53) \quad L_k(\mathbf{A}) = \frac{1}{\Delta'(\lambda_k)} \left\{ \frac{\text{adj}(\lambda \mathbf{E} - \mathbf{A})}{\theta(\lambda)} \right\}_{\lambda=\lambda_k} \quad (k = 1, 2, \dots, s)$$

képlettel számítható.

Abban az esetben, ha a karakterisztikus egyenlet gyökei egyszeresek, akkor (3.2.53) a következőképpen egyszerűsödik:

$$(3.2.54) \quad L_k(\mathbf{A}) = \frac{1}{D'(\lambda_k)} \text{adj}(\lambda_k \mathbf{E} - \mathbf{A}) \quad (k = 1, 2, \dots, s).$$

Bizonyítás. Legyen

$$D(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{\alpha_k}$$

az \mathbf{A} mátrix karakterisztikus polinomja,

$$\Delta(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)$$

a minimálpolinomja, és állítsuk elő a következő függvényt:

$$(3.2.55) \quad \frac{\Delta(x) - \Delta(y)}{x - y},$$

– amely y -ban legfeljebb $(s - 1)$ -edfokú racionális egész függvény – a $\lambda_1, \lambda_2, \dots, \lambda_s$ gyökhelyekhez tartozó Lagrange-féle alappolinomok segítségével:

$$(3.2.56) \quad \frac{\Delta(x) - \Delta(y)}{x - y} \equiv \sum_{k=1}^s \frac{\Delta(x) - \Delta(\lambda_k)}{x - \lambda_k} L_k(y).$$

Mivel $\Delta(\lambda_k) = 0$ és $\frac{\Delta(x)}{x - \lambda_k} = \Delta'(\lambda_k) L_k(x)$, a fenti kifejezés $\frac{\Delta(x) - \Delta(y)}{x - y} \equiv$

$\equiv \sum_{k=1}^s \Delta'(\lambda_k) L_k(x) L_k(y)$ alakban írható. Ha ebben az azonosságban az x

és y változó helyébe mátrixokat kívánunk helyettesíteni, akkor ezeket nem választhatjuk meg egészen önkényesen. Ügyelni kell ugyanis arra, hogy az előforduló műveletek tulajdonságai a mátrixok behelyettesítése után is megmaradjanak – tehát hogy a szorzás e mátrixokkal is kommutatív legyen és hogy a bal oldal nevezője helyére nemszinguláris mátrix kerüljön. Ha y helyére az \mathbf{A} és x helyére a $\lambda \mathbf{E}$ mátrixot írjuk olyan λ értékkel, amely nem gyöke a minimálegyenletnek, akkor e feltételek mind biztosítva vannak. Ezért az itt előforduló mátrixműveletek egyértelműsége miatt nem okozhat félreértést, ha az inverzmátrixszal végzett szorzás helyett formailag megtartjuk a skalár azonosságban fellépő tört alakot:

$$\frac{\Delta(\lambda \mathbf{E}) - \Delta(\mathbf{A})}{\lambda \mathbf{E} - \mathbf{A}} = \sum_{k=1}^s \Delta'(\lambda_k) L_k(\mathbf{A}) L_k(\lambda \mathbf{E}).$$

Figyelembe véve a reciprokmátrix definícióját, valamint a 3.2.2 tételt, ebből adódóan

$$\frac{\text{adj}(\lambda \mathbf{E} - \mathbf{A})}{\theta(\lambda)} = \sum_{k=1}^s \Delta'(\lambda_k) L_k(\mathbf{A}) L_k(\lambda \mathbf{E})$$

írható. Ha most λ helyére valamelyik λ_p sajátértéket helyettesítjük, akkor a jobb oldalon álló összegnek csupán egyetlen tagja marad meg, mivel

$$L_k(\lambda_p) = \delta_{kp}.$$

Innen azonnal adódik a (3.2.53) (ill. a (3.2.54)) képlet. ■

A kapott képletek begyakorlására előállítjuk az alkalmazások szempontjából fontos szerepet betöltő egyenletes kontinuáns mátrix spektrálfelbontását, majd néhány további kidolgozott példát közlünk.

6. Példa. Határozzuk meg a

$$(3.2.57) \quad \mathbf{K} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & 1 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix} \begin{matrix} (1) \\ (2) \\ \\ \vdots \\ (n-1) \\ (n) \end{matrix}$$

egyenletes kontinuáns mátrix (lásd (1.4.3)) spektrálfelbontását!

Megoldás. A karakterisztikus polinom ebben az esetben (1.4.20) és (1.4.31) alapján

$$(3.2.58) \quad D(\lambda) = |\lambda \mathbf{E} - \mathbf{K}| = U_n(\lambda).$$

Mint láttuk, a

$$(3.2.59) \quad \lambda = 2 \cos \theta$$

transzformációval

$$(3.2.60) \quad U_n(2 \cos \theta) = \frac{\sin(n+1)\theta}{\sin \theta} = 0,$$

azaz

$$(3.2.61) \quad (n+1)\theta_k = k\pi \quad (k = 1, 2, \dots, n)$$

adódik, ahonnan

$$(3.2.62) \quad \theta_k = \frac{k\pi}{n+1} \quad (k = 1, 2, \dots, n).$$

A (3.2.59) transzformáció segítségével a sajátértékek:

$$(3.2.63) \quad \lambda_k = 2 \cos \frac{k\pi}{n+1} \quad (k = 1, 2, \dots, n),$$

vagyis n különböző sajátérték van.

Az $L_k(\mathbf{K})$ Lagrange-féle mátrixpolinomokat a (3.2.54) képlettel számítjuk ki. A karakterisztikus polinom deriváltja

$$D'(\lambda) = \frac{dD}{d\theta} : \frac{d\lambda}{d\theta} = \frac{(n+1) \cos(n+1)\theta \sin \theta - \cos \theta \sin(n+1)\theta}{-2 \sin^3 \theta},$$

tehát

$$(3.2.64) \quad D'(\lambda_k) = \frac{n+1}{2} \cdot \frac{(-1)^{k-1}}{\sin^2 \frac{k\pi}{n+1}},$$

a karakterisztikus mátrix adjungáltja pedig (1.4.24) alapján

$$(3.2.65) \quad \{\text{adj}(\lambda \mathbf{E} - \mathbf{K})\}_{ij} = \begin{cases} \frac{\sin i\theta}{\sin \theta} \cdot \frac{\sin(n+1-j)\theta}{\sin \theta}, & \text{ha } i \leq j, \\ \frac{\sin j\theta}{\sin \theta} \cdot \frac{\sin(n+1-i)\theta}{\sin \theta}, & \text{ha } i \geq j. \end{cases}$$

Ha ide λ_k (3.2.63) alatti értékeit, azaz θ_k (3.2.62) alatti értékeit behelyettesítjük, akkor

$$(3.2.66) \quad \sin(n+1-j) \frac{k\pi}{n+1} = (-1)^{k-1} \sin \frac{jk\pi}{n+1}$$

miatt

$$(3.2.67) \quad \{\text{adj}(\lambda_k \mathbf{E} - \mathbf{K})\}_{ij} = \frac{(-1)^{k-1}}{\sin^2 \frac{k\pi}{n+1}} \sin \frac{ik\pi}{n+1} \sin \frac{jk\pi}{n+1}.$$

A (3.2.64) és (3.2.67) kifejezéseket a (3.2.54) képletbe behelyettesítve azt kapjuk, hogy

$$(3.2.68) \quad L_k(\mathbf{K}) = \left[\frac{2}{n+1} \sin \frac{ik\pi}{n+1} \sin \frac{jk\pi}{n+1} \right],$$

azaz a sajátvektorokra

$$(3.2.69) \quad \mathbf{u}_k^\top = \sqrt{\frac{2}{n+1}} \left[\sin \frac{k\pi}{n+1} \sin \frac{2k\pi}{n+1} \dots \sin \frac{ik\pi}{n+1} \dots \sin \frac{nk\pi}{n+1} \right]$$

adódik. Ezzel a \mathbf{K} mátrix spektrálfelbontása az alábbi alakban írható:

$$(3.2.70) \quad \mathbf{K} = \sqrt{\frac{2}{n+1}} \left[\sin \frac{ik\pi}{n+1} \right] \left\langle 2 \cos \frac{k\pi}{n+1} \right\rangle \left[\sin \frac{jk\pi}{n+1} \right] \sqrt{\frac{2}{n+1}}.$$

A valós szimmetrikus \mathbf{K} mátrixhoz ortogonális modálmátrix tartozik, a

$$(3.2.70) \text{ összefüggésből pedig kiolvasható, hogy az } \mathbf{U} = \left[\sqrt{\frac{2}{n+1}} \sin \frac{ik\pi}{n+1} \right]$$

modálmátrix szimmetrikus is. Ebből következik, hogy \mathbf{U} involutórius mátrix, tehát a \mathbf{K} mátrixot *involutórius* transzformáció viszi át a sajátértékeiből alkotott diagonálmátrixba.

* * *

7. Példa. Határozzuk meg $a\mathbf{E} + b\mathbf{K}$ spektrálfelbontását!

Megoldás. Mivel a \mathbf{K} kontinuáns mátrix spektrálfelbontása

$$(3.2.71) \quad \mathbf{K} = \mathbf{U} \langle \lambda_k \rangle \mathbf{U}^T,$$

ahol

$$u_{ik} = \sqrt{\frac{2}{n+1}} \sin \frac{ik\pi}{n+1} \text{ és } \lambda_k = 2 \cos \frac{k\pi}{n+1} \quad (i, k = 1, 2, \dots, n),$$

ezért $a\mathbf{E} + b\mathbf{K}$ spektrálfelbontása

$$(3.2.72) \quad a\mathbf{E} + b\mathbf{K} = \mathbf{U} \langle a + b\lambda_k \rangle \mathbf{U}^T =$$

$$= \left[\sqrt{\frac{2}{n+1}} \sin \frac{ik\pi}{n+1} \right] \left\langle a + 2b \cos \frac{k\pi}{n+1} \right\rangle \left[\sqrt{\frac{2}{n+1}} \sin \frac{jk\pi}{n+1} \right].$$

* * *

8. Példa. Határozzuk meg

$$\tilde{\mathbf{A}} = \begin{bmatrix} 1 & -1 & 0 \dots & 0 & 0 \\ -1 & 2 & -1 \dots & 0 & 0 \\ 0 & -1 & 2 \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 \dots & 2 & -1 \\ 0 & 0 & 0 \dots & -1 & 2 \end{bmatrix} \begin{matrix} (1) \\ (2) \\ \\ \vdots \\ (n) \end{matrix}$$

spektrálfelbontását!

Útmutatás. Tekintsük először a

$$(3.2.73) \quad \tilde{\mathbf{K}} = \begin{bmatrix} 1 & 1 & 0 \dots 0 & 0 \\ 1 & 0 & 1 \dots 0 & 0 \\ 0 & 1 & 0 \dots 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 \dots 0 & 1 \\ 0 & 0 & 0 \dots 1 & 0 \end{bmatrix}.$$

mátrixot, és az 5. Példában megismert módszerrel mutassuk meg, hogy ennek sajátértékeire

$$2 \cos \frac{2k-1}{2n+1} \pi \quad (k = 1, 2, \dots, n),$$

sajátvektorainak elemeire pedig

$$\frac{2}{\sqrt{2n+1}} \cos \frac{(2i-1) \cdot (2k-1) \pi}{2n+1} \frac{\pi}{2} \quad (k = 1, 2, \dots, n)$$

adódik. Mivel az adott mátrix $\tilde{\mathbf{A}} = \mathbf{E} - \tilde{\mathbf{K}}$, ezért $\tilde{\lambda}_k$ sajátértékei

$$(3.2.74) \quad \tilde{\lambda}_k = 2 - 2 \cos \frac{2k-1}{2n+1} \pi = 4 \sin^2 \frac{2k-1}{2n+1} \frac{\pi}{2}$$

alakban írhatók, a keresett spektrálfelbontás tehát

$$(3.2.75) \quad \tilde{\mathbf{A}} = 2\mathbf{E} - \tilde{\mathbf{K}} = \left[\frac{2}{\sqrt{2n+1}} \cos \frac{(2i-1)(2k-1)\pi}{2n+1} \frac{\pi}{2} \right] \times \\ \times \left\langle 4 \sin^2 \frac{2k-1}{2n+1} \frac{\pi}{2} \right\rangle \left[\frac{2}{\sqrt{2n+1}} \cos \frac{(2k-1)(2j-1)\pi}{2n+1} \frac{\pi}{2} \right].$$

* * *

9. Példa. Határozzuk meg

$$\tilde{\tilde{\mathbf{A}}} = \begin{bmatrix} 1 & -1 & 0 \dots 0 & 0 & 0 \\ -1 & 2 & -1 \dots 0 & 0 & 0 \\ 0 & -1 & 2 \dots 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 \dots 2 & -1 & \\ 0 & 0 & 0 \dots -1 & 1 & \end{bmatrix} \begin{matrix} (1) \\ (2) \\ \\ \vdots \\ (n) \end{matrix}$$

spektrálfelbontását!

Megoldás. Tekintsük most először az alábbi mátrixot:

$$(3.2.76) \quad \tilde{\mathbf{K}} = \begin{bmatrix} 1 & 1 & 0 \dots 0 & 0 \\ 1 & 0 & 1 \dots 0 & 0 \\ 0 & 1 & 0 \dots 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 \dots 0 & 1 \\ 0 & 0 & 0 \dots 1 & 1 \end{bmatrix} \begin{matrix} (1 \\ (2 \\ \vdots \\ (n \end{matrix}.$$

Ennek karakterisztikus polinomja:

$$D(\lambda) \equiv \left| \lambda \mathbf{E} - \tilde{\mathbf{K}} \right| \equiv \begin{vmatrix} \lambda - 1 & -1 & 0 \dots 0 & 0 & 0 \\ -1 & \lambda & -1 \dots 0 & 0 & 0 \\ 0 & -1 & \lambda \dots 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 \dots \lambda & -1 & -1 \\ 0 & 0 & 0 \dots -1 & \lambda - 1 & -1 \end{vmatrix},$$

ahonnan ismét a $\lambda = 2 \cos \theta$ transzformációval

$$(3.2.77) \quad D(\lambda) \equiv \frac{\sin(n+1)\theta}{\sin \theta} - \frac{2 \sin n\theta}{\sin \theta} + \frac{\sin(n-1)\theta}{\sin \theta} = \\ = \frac{2 \sin n\theta (\cos \theta - 1)}{\sin \theta} = \frac{-2 \sin n\theta \sin \frac{1}{2}\theta}{\cos \frac{1}{2}\theta}.$$

A sajátértékek tehát a $\sin \frac{1}{2}\theta_0 = 0$, $\sin n\theta_k = 0$ egyenletekből számított

$$\theta_k = \frac{k\pi}{n} \quad (k = 0, 1, 2, \dots, n-1)$$

alapján

$$\lambda_k = 2 \cos \frac{k\pi}{n} \quad (k = 0, 1, 2, \dots, n-1).$$

A karakterisztikus mátrix adjungáltjának elemei az előző feladat alapján

$$\left\{ \text{adj} \left(\lambda_k \mathbf{E} - \tilde{\mathbf{K}} \right) \right\}_{ij} = \frac{\cos \left(i - \frac{1}{2} \right) \frac{k\pi}{n} \cos \left(n - j + \frac{1}{2} \right) \frac{k\pi}{n}}{\cos^2 \frac{k\pi}{2n}}, \text{ ha } i \leq j.$$

Felhasználva az alábbi összefüggéseket:

$$\cos \left(n - j + \frac{1}{2} \right) \frac{k\pi}{n} = (-1)^k \cos \left(j - \frac{1}{2} \right) \frac{k\pi}{n},$$

továbbá

$$\begin{aligned}
 D'(\lambda_k) &= \left[\frac{dD}{d\theta} : \frac{d\lambda}{d\theta} \right]_{\theta_k} = \\
 &= \frac{-2n \cos n\theta_k \sin \frac{1}{2}\theta_k - \sin n\theta_k \cos^2 \frac{1}{2}\theta_k}{-2 \cos^2 \frac{1}{2}\theta_k \sin \theta_k} = \\
 &= \begin{cases} \frac{n(-1)^k}{2 \cos^2 \frac{k\pi}{2n}}, & \text{ha } k = 1, 2, \dots, n-1, \\ n, & \text{ha } k = 0, \end{cases}
 \end{aligned}$$

az $L_k(\tilde{\mathbf{K}})$ mátrixpolinom elemeire a következő adódik:

$$\begin{aligned}
 L_0(\tilde{\mathbf{K}})_{ij} &= \frac{1}{n}, \\
 L_k(\tilde{\mathbf{K}})_{ij} &= \frac{2}{n} \cos \frac{(2i-1)k\pi}{2n} \cos \frac{(2j-1)k\pi}{2n} \quad (k = 1, 2, \dots, n-1).
 \end{aligned}$$

A sajátvektorok elemei tehát

$$\begin{aligned}
 (3.2.78) \quad u_{i0} &= \frac{1}{\sqrt{n}}, \\
 u_{ik} &= \sqrt{\frac{2}{n}} \cos \frac{(2i-1)k\pi}{2n} \quad (k = 1, 2, \dots, n-1, \quad i = 1, 2, \dots, n).
 \end{aligned}$$

Mivel az adott mátrix $\tilde{\mathbf{A}} = 2\mathbf{E} - \tilde{\mathbf{K}}$, ezért sajátértékei

$$\tilde{\lambda}_k = 2 - 2 \cos \frac{k\pi}{n} = 4 \sin^2 \frac{k\pi}{2n} \quad (k = 0, 1, 2, \dots, n-1),$$

a keresett spektrálfelbontás tehát

$$(3.2.79) \quad \tilde{\mathbf{A}} = 2\mathbf{E} - \tilde{\mathbf{K}} = [u_{ik}] \left\langle 4 \sin^2 \frac{k\pi}{2n} \right\rangle [u_{kj}],$$

ahol a sajátvektorok u_{ik} elemei a (3.2.78) képletből számíthatók.

* * *

10. Példa. Mutassuk meg, hogy a $(2n+1)$ -edrendű szimmetrikus ciklikus mátrixoknak n számú kétszeres sajátértékük van, a $2n$ -edrendű szimmetrikus ciklikus mátrixoknak pedig $n-1$ számú kétszeres sajátértékük van, és a kétszeres sajátértékekhez tartozó sajátvektorok (valós) elemei kifejezhetők trigonometrikus függvények segítségével.

Megoldás. Legyen \mathbf{C} egy $(2n+1)$ -edrendű szimmetrikus ciklikus mátrix, amelyet az $\mathbf{\Omega}$ elemi ciklikus mátrix polinomjaként az alábbi alakban írhatunk:

$$(3.2.80) \quad \mathbf{C} = c_0 \mathbf{E} + \sum_{\nu=1}^{2n} c_{\nu} \mathbf{\Omega}^{\nu}.$$

A szimmetria miatt $c_{\nu} = c_{2n+1-\nu}$; továbbá $\mathbf{\Omega}^{2n+1} = \mathbf{E}$ miatt $\mathbf{\Omega}^{2n+1-\nu} = \mathbf{\Omega}^{-\nu}$, ezért

$$(3.2.81) \quad \mathbf{C} = c_0 \mathbf{E} + \sum_{\nu=1}^n c_{\nu} (\mathbf{\Omega}^{\nu} + \mathbf{\Omega}^{-\nu}).$$

Felhasználva az elemi ciklikus mátrix spektrálfelbontását, (3.2.52) alapján

$$\mathbf{C} = \sum_{k=0}^{2n} \left\{ c_0 + \sum_{\nu=1}^n c_{\nu} \left(e^{\frac{2\nu k \pi i}{2n+1}} + e^{\frac{-2\nu k \pi i}{2n+1}} \right) \right\} \mathbf{u}_k \mathbf{u}_k^H$$

adódik, ahol i a képzetes egység, és az \mathbf{u}_k sajátvektorok elemei

$$u_{pk} = \frac{1}{\sqrt{2n+1}} e^{\frac{2pk\pi i}{2n+1}} \quad (p = 0, 1, 2, \dots, 2n, \quad k = 0, 1, 2, \dots, 2n).$$

Figyelembe véve az alábbi összefüggéseket:

$$e^{\frac{2\nu k \pi i}{2n+1}} + e^{\frac{-2\nu k \pi i}{2n+1}} = 2 \cos \frac{2\nu k \pi}{2n+1}, \quad \text{valamint} \quad \cos \frac{2\nu k \pi}{2n+1} = \cos \frac{2\nu(2n+1-k)\pi}{2n+1},$$

látható, hogy az adott mátrix egyszeres sajátértéke

$$(3.2.82) \quad \lambda_0 = c_0 + \sum_{\nu=1}^n c_{\nu},$$

és kétszeres sajátértékei:

$$(3.2.83) \quad \lambda_k = c_0 + 2 \sum_{\nu=1}^n c_{\nu} \cos \frac{2\nu k \pi}{2n+1} \quad (k = 1, 2, \dots, n).$$

A kétszeres sajátértékekhez tartozó sajátvektorok \mathbf{u}_k és \mathbf{u}_{2n+1-k} , a spektrálfelbontás tehát

$$\begin{aligned} \mathbf{C} &= \left(c_0 + \sum_{\nu=1}^n c_{\nu} \right) \mathbf{u}_0 \mathbf{u}_0^H + \\ &+ \sum_{k=1}^n \left(c_0 + 2 \sum_{\nu=1}^n c_{\nu} \cos \frac{2\nu k \pi}{2n+1} \right) (\mathbf{u}_k \mathbf{u}_k^H + \mathbf{u}_{2n+1-k} \mathbf{u}_{2n+1-k}^H). \end{aligned}$$

Tekintettel arra, hogy az adott mátrix szimmetrikus, valós sajátvektorokat keresünk. Ezeket megkapjuk, ha a kétszeres sajátértékekhez tartozó lineárisan független sajátvektorokat megfelelő lineáris kombinációikkal helyettesítjük. Bevezetve a másodrendű

$$(3.2.84) \quad \mathbf{T} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{bmatrix}$$

unitér mátrixot, és figyelembe véve, hogy

$$\begin{aligned} \mathbf{u}_k \mathbf{u}_k^H + \mathbf{u}_{2n+1-k} \cdot \mathbf{u}_{2n+1-k}^H &= [\mathbf{u}_k \quad \mathbf{u}_{2n+1-k}] \mathbf{T} \mathbf{T}^H \begin{bmatrix} \mathbf{u}_k^H \\ \mathbf{u}_{2n+1-k}^H \end{bmatrix} = \\ &= \begin{bmatrix} \frac{1}{\sqrt{2}}(\mathbf{u}_k + \mathbf{u}_{2n+1-k}) & \frac{1}{i\sqrt{2}}(\mathbf{u}_k - \mathbf{u}_{2n+1-k}) \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}}(\mathbf{u}_k^H + \mathbf{u}_{2n+1-k}^H) \\ \frac{1}{i\sqrt{2}}(-\mathbf{u}_k^H + \mathbf{u}_{2n+1-k}^H) \end{bmatrix} = \\ &= [\mathbf{v}_k \quad \mathbf{w}_k] \begin{bmatrix} \mathbf{v}_k^T \\ \mathbf{w}_k^T \end{bmatrix} = \mathbf{v}_k \mathbf{v}_k^T + \mathbf{w}_k \mathbf{w}_k^T, \end{aligned}$$

ahol a \mathbf{v}_k , ill. \mathbf{w}_k sajátvektorok általános elemei

$$(3.2.85) \quad v_{pk} = \sqrt{\frac{2}{2n+1}} \cos \frac{2pk\pi}{2n+1} \quad \text{és} \quad w_{pk} = \sqrt{\frac{2}{2n+1}} \sin \frac{2pk\pi}{2n+1} \\ (p = 0, 1, 2, \dots, 2n, \quad k = 1, 2, \dots, n),$$

a spektrálfelbontás pedig

$$(3.2.86) \quad \mathbf{C} = \left(c_0 + 2 \sum_{\nu=1}^n c_\nu \right) \mathbf{u}_0 \mathbf{u}_0^T + \\ + \sum_{k=1}^n \left(c_0 + 2 \sum_{\nu=1}^n c_\nu \cos \frac{2\nu k\pi}{2n+1} \right) (\mathbf{v}_k \mathbf{v}_k^T + \mathbf{w}_k \mathbf{w}_k^T).$$

Ennek alapján felírható a páratlan, $(2n+1)$ -edrendű \mathbf{C} szimmetrikus ciklikus mátrix $f(\mathbf{C})$ függvényének általános eleme is:

$$(3.2.87) \quad \{f(\mathbf{C})\}_{pq} = \frac{1}{2n+1} f \left(c_0 + 2 \sum_{\nu=1}^n c_\nu \right) + \\ + \frac{2}{2n+1} \sum_{k=1}^n f \left(c_0 + 2 \sum_{\nu=1}^n c_\nu \cos \frac{2\nu k\pi}{2n+1} \right) \cos \frac{(p-q)2k\pi}{2n+1}.$$

Innen is látható, hogy a kapott mátrix szimmetrikus és ciklikus, mivel általános eleme csak a $p - q$ indexkülönbség abszolút értékétől függ.

Ha a szimmetrikus ciklikus mátrix rendszáma páros ($2n$), akkor az eredmény csupán annyiban módosul, hogy a mátrixnak nem egy, hanem két egyszeres sajátértéke lesz. Ebben az esetben ugyanis

$$(3.2.88) \quad \mathbf{C} = c_0 \mathbf{E} + \sum_{\nu=1}^{n-1} c_\nu (\mathbf{\Omega}^\nu + \mathbf{\Omega}^{-\nu}) + c_n \mathbf{\Omega}^n,$$

és most a $2n$ -edrendű elemi ciklikus mátrix spektrálfelbontását felhasználva, az egyszeres sajátértékek:

$$(3.2.89) \quad \begin{aligned} \lambda_0 &= c_0 + 2 \sum_{\nu=1}^{n-1} c_\nu + c_n \\ \lambda_n &= c_0 + 2 \sum_{\nu=1}^{n-1} (-1)^\nu c_\nu + (-1)^n c_n, \end{aligned}$$

a kétszeres sajátértékek:

$$(3.2.90) \quad \lambda_k = c_0 + 2 \sum_{\nu=1}^{n-1} c_\nu \cos \frac{\nu k \pi}{n} + (-1)^k c_n \quad (k = 1, 2, \dots, n-1).$$

A megfelelő sajátvektorok elemei ebben az esetben

$$(3.2.91) \quad u_{p0} = \frac{1}{\sqrt{2n}}, \quad u_{pn} = \frac{(-1)^p}{\sqrt{2n}} \quad (p = 0, 1, 2, \dots, n-1),$$

továbbá

$$(3.2.92) \quad v_{pk} = \frac{1}{\sqrt{n}} \cos \frac{pk\pi}{n}, \quad w_{pk} = \frac{1}{\sqrt{n}} \sin \frac{pk\pi}{n}$$

$$(p = 0, 1, 2, \dots, 2n-1; k = 1, 2, \dots, n-1).$$

Páros, $2n$ -edrendű szimmetrikus ciklikus mátrix függvényének általános elemét tehát ezek alapján a következő alakban nyerjük:

$$(3.2.93) \quad \begin{aligned} \{f(\mathbf{C})\}_{pq} &= \frac{1}{2} \left\{ f \left(c_0 + c_n + 2 \sum_{\nu=1}^{n-1} c_\nu \right) + \right. \\ &\quad \left. + (-1)^{p+q} f \left(c_0 + (-1)^n c_n + 2 \sum_{\nu=1}^{n-1} (-1)^\nu c_\nu \right) \right\} + \\ &\quad + \frac{1}{n} \sum_{k=1}^{n-1} f \left(c_0 + (-1)^k c_n + 2 \sum_{\nu=1}^{n-1} c_\nu \cos \frac{\nu k \pi}{n} \right) \cos \frac{(p-q)k\pi}{n}. \end{aligned}$$

* * *

3.3 KOMMUTATÍV BLOKKOKBÓL ÁLLÓ HIPERMÁTRIXOK

Ebben a szakaszban azt vizsgáljuk meg, hogy az 1.5 szakaszban tárgyalt hipermátrixok determinánsának, spektrálfelbontásának és függvényének a meghatározásában milyen egyszerűsítések válnak lehetővé, ha a hipermátrix blokkjai a szorzásra nézve kommutatívak.

3.3.1 A hipermátrix determinánsa

Tekintsünk először olyan másodrendű hipermátrixokat, amelyeknek blokkjai nemkommutatív (azonos rendű) kvadratikus mátrixok és határozzuk meg a hipermátrix determinását. Tegyük fel, hogy a fődiagonálisban álló blokkok nonszinguláris mátrixok. Az 1.5.3 tételben megmutattuk, hogy az ilyen hipermátrixok determinánsának kiszámítása kétféleképpen is redukálható alacsonyabb rendű determinánsok kiszámítására (lásd az (1.5.16) és (1.5.17) összefüggéseket).

Vegyük figyelembe, hogy azonos rendű determinánsok szorzása úgy is elvégezhető, hogy a megfelelő mátrixok szorzatának vesszük a determinánsát; mivel a determinánsokat bármilyen sorrendben összeszorozhatjuk, ezért a végeredményt nem befolyásolja, ha a megfelelő mátrixokat különböző sorrendben szorozzuk össze. Így a keresett determináns kiszámítására a következő négy képletet kapjuk, amelyek természetesen ugyanazt az eredményt adják:

$$(3.3.1) \quad \begin{array}{ll} |\mathbf{AD} - \mathbf{ACA}^{-1}\mathbf{B}|, & |\mathbf{AD} - \mathbf{BD}^{-1}\mathbf{CD}|, \\ |\mathbf{DA} - \mathbf{CA}^{-1}\mathbf{BA}|, & |\mathbf{DA} - \mathbf{DBD}^{-1}\mathbf{C}| \end{array}$$

Könnyen meggyőződhetünk arról, hogy ha a másodrendű hipermátrix valamelyik sorában, vagy valamelyik oszlopában álló blokkok kommutatívak, akkor a (3.3.1) alatti formulák valamelyike lényegesen leegyszerűsödik. Így a következő összefüggéseket nyerjük:

$$\begin{array}{ll} \text{ha } \mathbf{AC} = \mathbf{CA}, & \text{akkor a determináns } |\mathbf{AD} - \mathbf{CB}|; \\ \text{ha } \mathbf{AB} = \mathbf{BA}, & \text{akkor a determináns } |\mathbf{DA} - \mathbf{CB}|; \\ \text{ha } \mathbf{CD} = \mathbf{DC}, & \text{akkor a determináns } |\mathbf{AD} - \mathbf{BC}|; \\ \text{ha } \mathbf{BD} = \mathbf{DB}, & \text{akkor a determináns } |\mathbf{DA} - \mathbf{BC}|. \end{array}$$

Látható, hogy a keresett determináns ezeknek az eseteknek mindegyikében a hipermátrix blokkjaiból a másodrendű determináns képzési szabályának megfelelően számított mátrix determinánsa lesz, de attól függően, hogy melyik két blokk felcserélhető, a képletek közül csak az egyik érvényes.

Ha viszont a négy blokkból álló hipermátrix blokkjai mind kommutatívak, akkor az egyes blokkok sorrendje teljesen közömbös, és a fenti négy képlet bármelyike alkalmazható. Ezt a gondolatot tetszőleges rendű, kommutatív blokkokból álló hipermátrixra alkalmazva, célszerűnek mutatkozik a hiperdetermináns fogalmának bevezetése.

3.3.1 definíció. Ha $[A_{ij}]$ olyan n -edrendű hipermátrix, amely m -edrendű kommutatív blokkokból áll, akkor azt az m -edrendű mátrixot, amely ugyanúgy számítható az $[A_{ij}]$ hipermátrix A_{ij} blokkjaiból, mint ahogy a közönséges $[a_{ij}]$ mátrix determinánsa számítható az a_{ij} skalárelemekből, **hiperdeterminánsnak** nevezzük és $\text{Det}[A_{ij}]$ -vel jelöljük:

$$\text{Det}[A_{ij}] = \sum_{(n!)} (-1)^I A_{1\nu_1} A_{2\nu_2} \dots A_{n\nu_n}.$$

Megjegyzés. Ne tévesszük szem elől, hogy míg a közönséges mátrix determinánsa skalár érték, addig a hipermátrix hiperdeterminánsa m -edrendű kvadratikus mátrix!

E definíció segítségével a kommutatív blokkokból álló, másodrendű hipermátrix determinánsának kiszámítási szabálya a következőképpen fogalmazható meg: a másodrendű hipermátrix determinánsa egyenlő a hiperdetermináns determinánsával. Az alábbi tételben általánosítjuk ezt az eredményt tetszőleges rendszámú, kommutatív blokkokból álló hipermátrixokra.

3.3.1 tétel. Legyen $[A_{ij}]$ egy m -edrendű kommutatív blokkokból álló, n -edrendű hipermátrix. Az $[A_{ij}]$ hipermátrix determinánsa egyenlő a hiperdetermináns determinánsával (lásd [27]):

$$(3.3.2) \quad |[A_{ij}]| = |\text{Det}[A_{ij}]|.$$

Bizonyítás. Idézzük fel a mátrixok minimális diadikus felbontására vonatkozó 1.3.1 tételt. Ott feltettük, hogy – esetleg sorok és oszlopok átrendezésével – minden egyes diád levonása után a kapott mátrix bal felső sarokeleme zérustól különböző. Megmutattuk, hogy ebből szükségképpen következik a mátrix trianguláris faktorizációja, és az a tény, hogy a mátrix bal felső sarokminorai zérustól különbözőek. Az 1.7.3 tétel szerint ennek megfordítása is igaz: ha egy mátrix bal felső sarokminorai zérustól különbözőek, akkor lehetséges a mátrix trianguláris faktorizációja. Tegyük fel tehát, hogy az \mathbf{A} mátrix előállítható az alábbi alakban:

$$(3.3.3) \quad \mathbf{A} = \begin{bmatrix} 1 & & & \\ * & 1 & & \\ * & * & 1 & \\ & & & \ddots \\ * & * & * & \dots & 1 \end{bmatrix} \begin{bmatrix} u_1 & & & \\ & u_2 & & \\ & & u_3 & \\ & & & \ddots \\ & & & & u_n \end{bmatrix} \begin{bmatrix} 1 & * & * & \\ & 1 & * & \\ & & 1 & * \\ & & & \ddots & \vdots \\ & & & & 1 \end{bmatrix}$$

ahol az első tényező alsó, a harmadik tényező felső háromszögmátrix (a csillaggal jelölt elemek zérustól különbözőek lehetnek), a második tényező pedig olyan diagonálmátrix, amelynek elemei \mathbf{A} bal felső sarokminorai segítségével a következőképpen fejezhetők ki:

$$(3.3.4) \quad u_k = \frac{D_k}{D_{k-1}},$$

ahol D_k az \mathbf{A} mátrix k -adrendű bal felső sarokminora. Ha ugyanis generálóként mindig a bal felső sarokelemet választjuk, és az így adódó alsó, illetve felső háromszögmátrix főátlójában álló elemeket jobbra, illetve balra kiemeljük, akkor megkapjuk a középső tényező u_k diagonálelemét. Képezve mindkét oldal k -adrendű bal felső sarokminorát, a $D_k = u_1 u_2 \dots u_k$ összefüggést nyerjük, ahonnan azonnal következik (3.3.4).

Alkalmazzuk ezt a segédtevélt az $[\mathbf{A}_{ij}]$ hiperhármatrixra, amelynek blokkjai kommutatívak. A \mathbf{D}_k sarokminorok most olyan hiperdeterminánsok, amelyek ugyanúgy számíthatók az \mathbf{A}_{ij} blokkokból, mint ahogy a D_k minorok az a_{ij} skalár elemekből. Az \mathbf{A}_{ij} blokkok kommutativitása miatt a \mathbf{D}_k mátrixok is kommutatívak, a minorokra tett kikötés miatt pedig nonsingulárisak. Ezért a (3.3.3) alakú felbontást az $[\mathbf{A}_{ij}]$ hiperhármatrixra alkalmazva,

$$(3.3.5) \quad [\mathbf{A}_{ij}] = \begin{bmatrix} \mathbf{E} & 0 & & 0 \\ * & \mathbf{E} & 0 & 0 \\ * & * & \mathbf{E} & \dots & 0 \\ & & & \ddots & \\ * & * & * & & \mathbf{E} \end{bmatrix} \begin{bmatrix} \mathbf{U}_1 & & & \\ & \mathbf{U}_2 & & \\ & & \ddots & \\ & & & \mathbf{U}_n \end{bmatrix} \begin{bmatrix} \mathbf{E} & * & * & * \\ 0 & \mathbf{E} & * & * \\ 0 & 0 & \mathbf{E} & \dots & * \\ & & & \ddots & \\ 0 & 0 & 0 & & \mathbf{E} \end{bmatrix},$$

ahol $\mathbf{U}_k = \mathbf{D}_k \mathbf{D}_{k-1}^{-1}$. Képezzük az $[\mathbf{A}_{ij}]$ hiperhármatrix determinánsát. Mivel az alsó és a felső háromszögmátrix determinánsa 1, a keresett determináns a középső tényező diagonálisában álló blokkok determinánsának a szorzata:

$$|[\mathbf{A}_{ij}]| = \prod_{k=1}^n |\mathbf{U}_k| = \left| \prod_{k=1}^n \mathbf{U}_k \right|.$$

A kommutativitás miatt azonban

$$\prod_{k=1}^n \mathbf{U}_k = \mathbf{D}_1 \mathbf{D}_2 \mathbf{D}_1^{-1} \mathbf{D}_3 \mathbf{D}_2^{-1} \dots \mathbf{D}_n \mathbf{D}_{n-1}^{-1} = \mathbf{D}_n,$$

tehát $|[\mathbf{A}_{ij}]| = |\mathbf{D}_n| = |\mathbf{Det} [\mathbf{A}_{ij}]|$, amit bizonyítani kellett. ■

Megjegyzés. Az a feltétel, amelyet a bizonyítás során alkalmaztunk, hogy ti. a bal felső sarokminorok nem lehetnek zérussal egyenlőek, nem jelenti az általánosság megszorítását. Ha ugyanis ez nem teljesül, akkor nonsinguláris mátrixok esetén a sorok és oszlopok mindig átrendezhetők úgy, hogy a sarokminorok mind zérustól különbözőek legyenek és ez a determinánsnak legfeljebb az előjelét változtatja meg.

3.3.2 Mátrixok direkt szorzata

3.3.2 definíció. Az m -edrendű \mathbf{A} és az n -edrendű \mathbf{B} mátrixok $\mathbf{A} \otimes \mathbf{B}$ -vel jelölt direkt szorzatán az

$$(3.3.6) \quad \mathbf{A} \otimes \mathbf{B} = [\mathbf{A}b_{ij}] = \begin{bmatrix} \mathbf{A}b_{11} & \mathbf{A}b_{12} \dots \mathbf{A}b_{1n} \\ \mathbf{A}b_{21} & \mathbf{A}b_{22} \dots \mathbf{A}b_{2n} \\ \dots & \dots & \dots & \dots \\ \mathbf{A}b_{n1} & \mathbf{A}b_{n2} \dots \mathbf{A}b_{nn} \end{bmatrix}$$

n -edrendű hipermátrixot értjük.

A direkt szorzat tehát olyan n -edrendű hipermátrix, amelynek a blokkjai valamennyien az \mathbf{A} mátrix skalár többszörösei, mégpedig az ij indexű blokkot úgy kapjuk, hogy az \mathbf{A} mátrixot megszorozzuk a \mathbf{B} mátrix ij indexű elemével. Ezért a blokkok természetesen kommutatívak (lásd pl. [12]).

11. Példa. A 3.3.1 tétel alkalmazásával számítsuk ki az $\mathbf{A} \otimes \mathbf{B}$ direkt szorzat determinánsát.

Megoldás. A hiperdetermináns úgy számítható, hogy az \mathbf{A} mátrixot először formálisan skalárként kezeljük és kiemeljük az A tényezőt $[Ab_{ij}]$ minden sorából: $|[Ab_{ij}]| = A^n |[b_{ij}]|$. Ezután figyelembe vesszük, hogy \mathbf{A} egy m -rendű mátrix:

$$(3.3.7) \quad \mathbf{Det} [\mathbf{A}b_{ij}] = \mathbf{A}^n |\mathbf{B}|.$$

Az így kapott m -edrendű mátrix a hiperdetermináns. Ennek determinánsát $|\mathbf{A}^n| |\mathbf{B}| = |\mathbf{B}|^m |\mathbf{A}^n|$ alakban kapjuk, vagyis $(|\mathbf{A}^n| = |\mathbf{A}|^n$ miatt)

$$(3.3.8) \quad |\mathbf{A} \otimes \mathbf{B}| = |\mathbf{B}|^m |\mathbf{A}|^n.$$

* * *

Hogy jobban megismerkedjünk a direkt szorzat fogalmával, néhány egyszerű példán megmutatjuk alkalmazását.

12. Példa. Írjuk fel direkt szorzatként az alábbi hiperdiagonál-mátrixot:

$$\begin{bmatrix} \mathbf{A} & 0 & 0 \dots 0 \\ 0 & \mathbf{A} & 0 \dots 0 \\ 0 & 0 & \mathbf{A} \dots 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \dots \mathbf{A} \end{bmatrix} \begin{matrix} (1) \\ (2) \\ \vdots \\ (n) \end{matrix}$$

Megoldás. Vegyük észre, hogy ez a hiperdiagonál-mátrix, melynek diagonál-blokkjai az m -edrendű \mathbf{A} mátrixszal egyenlőek, felírható az m -edrendű \mathbf{A} mátrix és az n -edrendű egységmátrix direkt szorzataként:

$$(3.3.9) \quad \mathbf{A}_m \otimes \mathbf{E}_n = \begin{bmatrix} \mathbf{A} & 0 \dots 0 \\ 0 & \mathbf{A} \dots 0 \\ \dots & \dots \dots \dots \\ 0 & 0 \dots \mathbf{A} \end{bmatrix} \begin{matrix} (1) \\ (2) \\ \vdots \\ (n) \end{matrix} \cdot$$

* * *

Ha a (3.3.9) direkt szorzatban a tényezők sorrendjét megfordítjuk, akkor

$$(3.3.10) \quad \mathbf{E}_n \otimes \mathbf{A}_m = \begin{bmatrix} a_{11}\mathbf{E} & a_{12}\mathbf{E} \dots a_{1m}\mathbf{E} \\ a_{21}\mathbf{E} & a_{22}\mathbf{E} \dots a_{2m}\mathbf{E} \\ \dots & \dots \dots \dots \\ a_{m1}\mathbf{E} & a_{m2}\mathbf{E} \dots a_{mm}\mathbf{E} \end{bmatrix}$$

adódik, azaz olyan hipermátrix, amelynek a blokkjai az n -edrendű egység-mátrixnak a többszörösei. Innen is látható, hogy a direkt szorzat nyilván nem kommutatív. Nem nehéz azonban belátni, hogy a sorok és oszlopok ugyanolyan átrendezésével, vagyis alkalmasan választott permutáló mátrixszal végzett ortogonális transzformációval a direkt szorzat tényezői felcserélődnek; tehát létezik olyan \mathbf{P} permutáló mátrix, hogy $\mathbf{P}(\mathbf{A}_m \otimes \mathbf{E}_n)\mathbf{P}^\top = \mathbf{E}_n \otimes \mathbf{A}_m$, vagy általánosabban

$$(3.3.11) \quad \mathbf{P}(\mathbf{A}_m \otimes \mathbf{B}_n)\mathbf{P}^\top = \mathbf{B}_n \otimes \mathbf{A}_m.$$

Gyakorlásképpen meggyőződhetünk arról, hogy ez a $\mathbf{P} = [\mathbf{P}_{pj}]$ permutáló mátrix olyan $n \times m$ típusú blokkokból álló hipermátrix, amelynek blokkjai n -edrendű, ill. m -edrendű megfelelő egységvektorokból képzett diádok:

$$(3.3.12) \quad \mathbf{P}_{pj} = \mathbf{e}_j^{(n)} \mathbf{e}_p^{(m)\top}.$$

13. Példa. Alakítsuk át a másodrendű \mathbf{A} és a harmadrendű \mathbf{B} mátrix direkt szorzatát az elemek megfelelő átrendezésével a harmadrendű \mathbf{B} és a másodrendű \mathbf{A} direkt szorzatára.

Megoldás.

$$\begin{aligned}
 & \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right] \left[\begin{array}{ccc|ccc} \mathbf{A}b_{11} & \mathbf{A}b_{12} & \mathbf{A}b_{13} & & & \\ \hline & & & & & \\ & & & & & \\ \hline \mathbf{A}b_{21} & \mathbf{A}b_{22} & \mathbf{A}b_{23} & & & \\ \hline & & & & & \\ \hline \mathbf{A}b_{31} & \mathbf{A}b_{32} & \mathbf{A}b_{33} & & & \end{array} \right] \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right] = \\
 & = \left[\begin{array}{ccc|ccc} \mathbf{B}a_{11} & & & \mathbf{B}a_{12} & & \\ \hline & & & & & \\ & & & & & \\ \hline \mathbf{B}a_{21} & & & \mathbf{B}a_{22} & & \end{array} \right].
 \end{aligned}$$

* * *

Az 3.3.2 definícióból következik a direkt szorzatok szorzására vonatkozó következő tétel.

3.3.2 tétel. Ha \mathbf{A} és \mathbf{C} két m -edrendű, \mathbf{B} és \mathbf{D} pedig két n -edrendű mátrix, akkor

$$(3.3.13) \quad (\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}.$$

Bizonyítás. A definíció alapján felírható:

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = [\mathbf{A}b_{ik}][\mathbf{C}d_{kj}].$$

A hipermátrixok szorzására vonatkozó tételből (lásd (1.5.3)) következik, hogy a szorzat olyan hipermátrix lesz, melynek minden blokkja \mathbf{AC} többszöröse, az ij indexű blokk együtthatója pedig $\sum_{k=1}^n b_{ik}d_{kj}$. Vagyis

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \left[\mathbf{AC} \sum_{k=1}^n b_{ik}d_{kj} \right] = \mathbf{AC} \otimes \mathbf{BD},$$

amit bizonyítani kellett. ■

14. Példa. Határozzuk meg az $\mathbf{A} \otimes \mathbf{B}$ direkt szorzat inverzét!

Megoldás. A 3.3.2 tétel alapján közvetlenül meggyőződhetünk arról, hogy két mátrix direkt szorzatának az inverze egyenlő a mátrixok inverzének a direkt szorzatával:

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1},$$

hiszen a beszorzással valóban az mn -edrendű \mathbf{E}_{mn} egységmátrix adódik:

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{A}^{-1} \otimes \mathbf{B}^{-1}) = \mathbf{E}_m \otimes \mathbf{E}_n = \mathbf{E}_{mn}.$$

3.3.3 Hipermátrix spektrálfelbontása

Amint az előzőekben láttuk, a hipermátrix blokkjainak kommutativitása lényeges egyszerűsítéseket tesz lehetővé hipermátrixok determinánsának a kiszámításában. Önként adódik tehát a kérdés: vajon a sajátértékek és sajátvektorok meghatározásában is előnyt jelent-e az a megszorítás, hogy a hipermátrix blokkjai kommutatívak. Mint látni fogjuk, az ilyen speciális hipermátrixok spektrálfelbontása – bizonyos korlátozások mellett – általában valóban egyszerűbben határozható meg, ha kihasználjuk a blokkok kommutativitását. Az alábbi eljárás fő előnye röviden úgy fogalmazható meg, hogy mn -edrendű mátrixok spektrálfelbontásának a feladata m -edrendű és n -edrendű mátrixok spektrálfelbontásának meghatározására redukálható. Ehhez azonban bizonyos megszorításokat kell alkalmaznunk, amelyek biztosítják az összes előforduló mátrixok teljes sajátvektor-rendszerének létezését, vagyis azt, hogy a mátrixok *egyszerű struktúrájúak* legyenek. A tétel megfogalmazásának egyszerűsítése, továbbá bonyolult jelölésmód elkerülése érdekében a szükségesnél erősebb megszorítást teszünk, mégpedig azt, hogy a hipermátrix blokkjai egy hermitikus \mathbf{A} mátrix függvényei. Ezek lehetnek racionális függvények, vagy olyan hatványsorral definiált függvények, melyekre a konvergencia feltételei teljesülnek. A későbbiek kedvéért még egy korlátozást teszünk ezekre a függvényekre. Jelölje a hipermátrix ij indexű blokkját

$$\mathbf{A}_{ij} = f_{ij}(\mathbf{A}).$$

Feltesszük, hogy az $f_{ij}(z)$ függvényekre teljesül

$$f_{ij}(z) = \overline{f_{ji}(\overline{z})}.$$

Ez a korlátozás is gyengíthető, azonban a tétel megfogalmazását nehezebbé tenné. Végül felhasználjuk két vektor direkt szorzatának a fogalmát. Bár ez a fogalom a 3.3.2 definíció speciális eseteként természetesen adódik, érdemes külön is megfogalmazni.

3.3.3 definíció. Az m elemű \mathbf{a}^\top és az n elemű \mathbf{b}^\top sorvektorból képezett

$$(3.3.14) \quad \mathbf{a}^\top \otimes \mathbf{b}^\top = [\mathbf{a}^\top b_1 \quad \mathbf{a}^\top b_2 \dots \mathbf{a}^\top b_n] = \\ = [[a_1 b_1 \quad a_2 b_1 \dots a_m b_1][a_1 b_2 \quad a_2 b_2 \dots a_m b_2] \dots [a_1 b_n \quad a_2 b_n \dots a_m b_n]]$$

hipervektort \mathbf{a}^\top és \mathbf{b}^\top direkt szorzatának nevezzük.

Oszlopvektorokra teljesen hasonlóan értelmezhető a direkt szorzat. Ezek után a következőképpen fogalmazható meg a hipermátrixok spektrálfelbontásának előállítására vonatkozó tétel.*

*A tétel lényegében megtalálható Egerváry dolgozatában (lásd [27]), bár annak megfogalmazása és bizonyítása is eltér az itt közölt változattól.

3.3.3 tétel. Legyen \mathbf{A} egy m -edrendű hermitikus mátrix, melynek spektrálfelbontása

$$(3.3.15) \quad \mathbf{A} = \mathbf{U} \langle a_k \rangle \mathbf{U}^H = \sum_{k=1}^m a_k \mathbf{u}_k \mathbf{u}_k^H;$$

az $f_{ij}(z)$ hatványsorba fejthető függvények ($i, j = 1, 2, \dots, n$) elégítsék ki az

$$(3.3.16) \quad f_{ij}(z) = \overline{f_{ji}(z)}$$

feltételeket, és az a_k sajátértékek legyenek valamennyi $f_{ij}(z)$ függvény konvergenciakörének a belsejében. Legyen továbbá az n -edrendű $[f_{ij}(a_k)]$ (hermitikus) mátrixok spektrálfelbontása

$$(3.3.17) \quad [f_{ij}(a_k)] = \mathbf{V}^{(k)} \langle \lambda_l^{(k)} \rangle \mathbf{V}^{(k)H} = \sum_{l=1}^n \lambda_l^{(k)} \mathbf{v}_l^{(k)} \mathbf{v}_l^{(k)H} \quad (k = 1, 2, \dots, m).$$

Ekkor az $[f_{ij}(\mathbf{A})]$ hipermátrix spektrálfelbontása

$$(3.3.18) \quad [f_{ij}(\mathbf{A})] = \sum_{l=1}^n \sum_{k=1}^m \lambda_l^{(k)} (\mathbf{u}_k \otimes \mathbf{v}_l^{(k)}) (\mathbf{u}_k^H \otimes \mathbf{v}_l^{(k)H})$$

alakban állítható elő.

A (3.3.18) összefüggésből kiolvasható, hogy az $f_{ij}(\mathbf{A})$ blokkokból álló hipermátrix sajátértékeit az $[f_{ij}(a_k)]$ mátrixok sajátértékei ($k = 1, 2, \dots, m$), sajátvektorait pedig az \mathbf{A} mátrix sajátvektorainak és az $[f_{ij}(a_k)]$ mátrixok sajátvektorainak direkt szorzatai szolgáltatják.

Bizonyítás. Tegyük fel, hogy ismert az m -edrendű hermitikus \mathbf{A} mátrix (3.3.15) alakú spektrálfelbontása. Ennek felhasználásával az $f_{ij}(\mathbf{A})$ mátrix-függvények spektrálfelbontására

$$(3.3.19) \quad f_{ij}(\mathbf{A}) = \sum_{k=1}^m f_{ij}(a_k) \mathbf{u}_k \mathbf{u}_k^H \quad (i, j = 1, 2, \dots, n)$$

adódik.

Ha most figyelembe vesszük, hogy az $f_{ij}(\mathbf{A})$ blokkokból álló $[f_{ij}(\mathbf{A})]$ hipermátrix az n -edrendű

$$\mathbf{E}_{ij} = \mathbf{e}_i^{(n)} \mathbf{e}_j^{(n)T}$$

mátrixegységek segítségével az

$$(3.3.20) \quad [f_{ij}(\mathbf{A})] = \sum_{i=1}^n \sum_{j=1}^n f_{ij}(\mathbf{A}) \otimes \mathbf{e}_i^{(n)} \mathbf{e}_j^{(n)\top}$$

alakban írható, és behelyettesítjük a (3.3.19) kifejezéseket, akkor az

$$[f_{ij}(\mathbf{A})] = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^m f_{ij}(a_k) \mathbf{u}_k \mathbf{u}_k^H \otimes \mathbf{e}_i^{(n)} \mathbf{e}_j^{(n)\top},$$

majd az összegezés sorrendjének felcserélésével az

$$[f_{ij}(\mathbf{A})] = \sum_{k=1}^m \mathbf{u}_k \mathbf{u}_k^H \otimes \sum_{i=1}^n \sum_{j=1}^n f_{ij}(a_k) \mathbf{e}_i^{(n)} \mathbf{e}_j^{(n)\top}$$

összefüggést nyerjük. Mivel a direkt szorzatok második tényezőjében szereplő kettős szummák az $f_{ij}(a_k)$ elemű n -edrendű hermitikus mátrixokat állítják elő:

$$\sum_{i=1}^n \sum_{j=1}^n f_{ij}(a_k) \mathbf{e}_i^{(n)} \mathbf{e}_j^{(n)\top} = [f_{ij}(a_k)] \quad (k = 1, 2, \dots, m),$$

a (3.3.17) összefüggésekkel adott spektrálfelbontásuk behelyettesítésével

$$[f_{ij}(\mathbf{A})] = \sum_{k=1}^m \mathbf{u}_k \mathbf{u}_k^H \otimes \sum_{l=1}^n \lambda_l^{(k)} \mathbf{v}_l^{(k)} \mathbf{v}_l^{(k)H}$$

adódik. Az itt fellépő direkt szorzatokra alkalmazva a 3.3.2 tételt:

$$(3.3.21) \quad \mathbf{u}_k \mathbf{u}_k^H \otimes \mathbf{v}_l^{(k)} \mathbf{v}_l^{(k)H} = (\mathbf{u}_k \otimes \mathbf{v}_l^{(k)}) (\mathbf{u}_k^H \otimes \mathbf{v}_l^{(k)H}),$$

végül a (3.3.18) alatti spektrálfelbontást kapjuk. ■

Megjegyzés. A bizonyításból következően az \mathbf{A} mátrixról elég lenne feltenni, hogy egyszerű struktúrájú, továbbá az $f_{ij}(z)$ függvényekről elég lenne feltenni, hogy az \mathbf{A} mátrix a_k sajátértékeinek behelyettesítésével adódó $[f_{ij}(a_k)]$ mátrixok mind egyszerű struktúrájúak. Ezek a feltételek biztosítanak még a teljes sajátvektor-rendszer létezését, vagyis a spektrálfelbontás előállíthatóságát.

Abban a speciális esetben, amikor $f_{ij}(z) = b_{ij}z$, és a $\mathbf{B} = [b_{ij}]$ hermitikus mátrix spektrálfelbontása

$$(3.3.22) \quad \mathbf{B} = \mathbf{V} \langle b_l \rangle \mathbf{V}^H = \sum_{l=1}^n b_l \mathbf{v}_l \mathbf{v}_l^H,$$

akkor $[f_{ij}(\mathbf{A})] = [b_{ij}\mathbf{A}] = \mathbf{A} \otimes \mathbf{B}$, és így a direkt szorzat spektrálfelbontása a 3.3.3 tétel speciális eseteként automatikusan adódik. Itt azonban a \mathbf{V} modálmátrix független a k indextől, ugyanis

$$[f_{ij}(a_k)] = [b_{ij}a_k] = a_k[b_{ij}] = \mathbf{V}\langle a_k b_i \rangle \mathbf{V}^H.$$

A 3.3.3 tétel következményeként tehát kimondható, hogy az \mathbf{A} és \mathbf{B} mátrix direkt szorzatának sajátértékei az \mathbf{A} mátrix a_k sajátértékeinek és a \mathbf{B} mátrix b_l sajátértékeinek $a_k b_l$ szorzatai ($k = 1, 2, \dots, m$; $l = 1, 2, \dots, n$), sajátvektoraik pedig az \mathbf{A} mátrix \mathbf{u}_k sajátvektorainak és a \mathbf{B} mátrix \mathbf{v}_l sajátvektorainak $\mathbf{u}_k \otimes \mathbf{v}_l$ direkt szorzatai.

3.3.4 Kronecker-polinomok

A 3.3.3 tétel következménye általánosítható az ún. Kronecker-polinomok spektrálfelbontására [37].

3.3.4 definíció. Az \mathbf{A} és \mathbf{B} mátrix Kronecker-polinomján a

$$(3.3.23) \quad \mathbf{T} = \sum_{(\nu)} \sum_{(\mu)} c_{\mu\nu} \mathbf{A}^\mu \otimes \mathbf{B}^\nu$$

kifejezést értjük, ahol μ és ν véges sok nemnegatív egész értéket felvehet.

Az 3.3.3 tételből, ill. annak két mátrix direkt szorzatára vonatkozó speciális esetéből közvetlenül következik a Kronecker-polinomok spektrálfelbontására vonatkozó következő tétel.

3.3.4 tétel. A (3.3.23) Kronecker-polinom spektrálfelbontása a (3.3.15) és (3.3.22) jelölések segítségével

$$\sum_{(\nu)} \sum_{(\mu)} c_{\mu\nu} \mathbf{A}^\mu \otimes \mathbf{B}^\nu = \sum_{l=1}^n \sum_{k=1}^m \left(\sum_{(\nu)} \sum_{(\mu)} c_{\mu\nu} a_k^\mu b_l^\nu \right) (\mathbf{u}_k \otimes \mathbf{v}_l) (\mathbf{u}_k^H \otimes \mathbf{v}_l^H)$$

alakban állítható elő.

A mátrixfüggvények előállítására vonatkozó tételből természetesen következik az is, hogy egy Kronecker-polinom tetszőleges függvénye – ha az létezik – ugyancsak előállítható a fenti spektrálfelbontás segítségével.

3.3.5 tétel. Legyen $g(z)$ olyan függvény, amelyre a $g(\mathbf{T})$ mátrixfüggvény létezik, ahol \mathbf{T} a (3.3.23) alatti Kronecker-polinom. Ekkor a $g(\mathbf{T})$ függvény ij indexű blokkjának pq indexű eleme

$$(3.3.24) \quad \{g(\mathbf{T})\}_{ij,pq} = \sum_{l=1}^n \sum_{k=1}^m g \left(\sum_{(\nu)} \sum_{(\mu)} c_{\mu\nu} a_k^\mu b_l^\nu \right) u_{pk} v_{il} \bar{u}_{qk} \bar{v}_{jl}$$

alakban írható fel.

15. Példa. Határozzuk meg az

$$M = \begin{array}{c} \begin{array}{cccc} 1 & 2 & \dots & m \end{array} \\ \left[\begin{array}{ccc|cc|cc|cc} 4 & -1 & & & -1 & & & & \\ -1 & 4 & & & & -1 & & & \\ & & \ddots & & & & \ddots & & \\ & & & -1 & & & & -1 & \\ & & & -1 & 4 & & & & \\ \hline -1 & & & 4 & -1 & & & & \\ & -1 & & -1 & 4 & & & & \\ & & \ddots & & & \ddots & & -1 & \\ & & & -1 & & -1 & 4 & & \\ \hline & & & & & & & -1 & \\ & & & & & & & -1 & \\ & & & & & & & & \ddots \\ & & & & & & & & -1 \\ \hline & & & -1 & & & & 4 & -1 \\ & & & -1 & & & & -1 & 4 \\ & & & & \ddots & & & & \ddots & -1 \\ & & & & & -1 & & & -1 & 4 \end{array} \right] \begin{array}{c} (1) \\ (2) \\ \vdots \\ (n) \end{array} \end{array}$$

hipermátrix spektrálfelbontását!

Megoldás. Mindenekelőtt észre kell vennünk, hogy az

$$(3.3.25) \quad \mathbf{A}_m = \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & & \ddots & & \\ & & & \ddots & -1 \\ & & & -1 & 2 \end{bmatrix} \begin{array}{c} (1) \\ (2) \\ \vdots \\ (m) \end{array}$$

kontinuáns mátrix segítségével az adott hipermátrix a következő particionált alakban írható:

$$M = \begin{bmatrix} \mathbf{A} + 2\mathbf{E} & & -\mathbf{E} & & \\ & \ddots & & \ddots & \\ & & -\mathbf{E} & \mathbf{A} + 2\mathbf{E} & -\mathbf{E} \\ & & & -\mathbf{E} & \mathbf{A} + 2\mathbf{E} \\ & & & & \ddots & -\mathbf{E} \\ & & & & & -\mathbf{E} & \mathbf{A} + 2\mathbf{E} \end{bmatrix}.$$

Ha az így felírt hipermátrixot két hipermátrix összegeként írjuk fel:

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & & & & \\ & \mathbf{A} & & & \\ & & \mathbf{A} & & \\ & & & \ddots & \\ & & & & \mathbf{A} \end{bmatrix} + \begin{bmatrix} 2\mathbf{E} & -\mathbf{E} & & & \\ -\mathbf{E} & 2\mathbf{E} & -\mathbf{E} & & \\ & -\mathbf{E} & 2\mathbf{E} & & \\ & & & \ddots & -\mathbf{E} \\ & & & -\mathbf{E} & 2\mathbf{E} \end{bmatrix},$$

akkor a direkt szorzat definíciójából (a (3.3.9) és (3.3.10) példák alapján) látható, hogy az adott mátrix

$$(3.3.26) \quad \mathbf{M} = \mathbf{A}_m \otimes \mathbf{E}_n + \mathbf{E}_m \otimes \mathbf{A}_n$$

Kronecker-polinom alakjában írható fel. Tekintettel arra, hogy a (3.3.25) mátrix a jól ismert kontinuáns mátrix, amelynek sajátértékeit és sajátvektorait explicit alakban is ki tudjuk fejezni, az adott \mathbf{M} mátrix spektrálfelbontása is felírható explicit alakban. Figyelembe véve tehát, hogy \mathbf{A}_m sajátértékei

$$\lambda_k^{(m)} = 4 \sin^2 \frac{k\pi}{2(m+1)},$$

sajátvektorainak elemei pedig

$$u_{pk}^{(m)} = \sqrt{\frac{2}{m+1}} \sin \frac{pk\pi}{m+1},$$

az adott \mathbf{M} hipermátrix sajátértékeire

$$(3.3.27) \quad \lambda_k^{(m)} + \lambda_l^{(n)} = 4 \sin^2 \frac{k\pi}{2(m+1)} + 4 \sin^2 \frac{l\pi}{2(n+1)}$$

adódik, \mathbf{M} elemei pedig a spektrálfelbontás segítségével a következő alakban írhatók fel:

$$(3.3.28) \quad [\mathbf{M}]_{ij,pq} = \frac{4}{(m+1)(n+1)} \sum_{l=1}^n \sum_{k=1}^m \left(4 \sin^2 \frac{k\pi}{2(m+1)} + \right. \\ \left. + 4 \sin^2 \frac{l\pi}{2(n+1)} \right) \sin \frac{pk\pi}{m+1} \sin \frac{qk\pi}{m+1} \sin \frac{il\pi}{n+1} \sin \frac{jl\pi}{n+1}.$$

* * *

Megjegyzés. A (3.3.23) összefüggéssel definiált Kronecker-polinomnak azt a speciális esetét, amiben az együtthatók $c_{01} = c_{10} = 1$, a többi együttható pedig zérus, az \mathbf{A}_m és \mathbf{B}_n mátrix *Kronecker-féle összegének* nevezzük. A fenti példában vizsgált \mathbf{M} mátrix tehát az m -edrendű és az n -edrendű \mathbf{A}_m , ill. \mathbf{A}_n mátrix Kronecker-féle összege.

Az alábbiakban két mátrix Kronecker-féle összegének egy speciális függvényét, nevezetesen az exponenciális függvényét vizsgáljuk meg, és bebizonyítjuk azt a meglepő tulajdonságát, mely szerint egyenlő az egyes mátrixok exponenciális függvényének direkt szorzatával. Tételben kimondva (lásd [2]):

3.3.6 tétel. *Ha \mathbf{A}_m tetszőleges m -edrendű, \mathbf{B}_n pedig tetszőleges n -edrendű mátrix, akkor fennáll az*

$$(3.3.29) \quad e^{\mathbf{A}_m \otimes \mathbf{E}_n + \mathbf{E}_m \otimes \mathbf{B}_n} = e^{\mathbf{A}_m} \otimes e^{\mathbf{B}_n}$$

összefüggés.

Bizonyítás. Mindenekelőtt megmutatjuk, hogy a kitevőben szereplő mátrixok a szorzás műveletére nézve kommutatívak, amiből következően a kitevő hatványaira alkalmazható a binomiális tétel. Ugyanis az 3.3.2 tétel felhasználásával

$$(\mathbf{A}_m \otimes \mathbf{E}_n)(\mathbf{E}_m \otimes \mathbf{B}_n) = \mathbf{A}_m \otimes \mathbf{B}_n, \quad \text{ill.} \quad (\mathbf{E}_m \otimes \mathbf{B}_n)(\mathbf{A}_m \otimes \mathbf{E}_n) = \mathbf{A}_m \otimes \mathbf{B}_n,$$

tehát a (3.3.29) összefüggés bal oldalán lévő exponenciális függvény hatványsorának egyes tagjait felírhatjuk a binomiális tétel segítségével:

$$\begin{aligned} e^{\mathbf{A}_m \otimes \mathbf{E}_n + \mathbf{E}_m \otimes \mathbf{B}_n} &= \sum_{p=0}^{\infty} \frac{1}{p!} (\mathbf{A}_m \otimes \mathbf{E}_n + \mathbf{E}_m \otimes \mathbf{B}_n)^p = \\ &= \sum_{p=0}^{\infty} \sum_{k=0}^p \frac{1}{k!(p-k)!} (\mathbf{A}_m \otimes \mathbf{E}_n)^k (\mathbf{E}_m \otimes \mathbf{B}_n)^{p-k} = \\ &= \sum_{p=0}^{\infty} \sum_{k=0}^p \frac{1}{k!(p-k)!} \mathbf{A}_m^k \otimes \mathbf{B}_n^{p-k}. \end{aligned}$$

A (3.3.29) összefüggés jobb oldalán álló függvények hatványsorának direkt szorzatából adódóan pedig

$$e^{\mathbf{A}_m} \otimes e^{\mathbf{B}_n} = \sum_{i=0}^{\infty} \frac{1}{i!} \mathbf{A}_m^i \otimes \sum_{j=0}^{\infty} \frac{1}{j!} \mathbf{B}_n^j = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{1}{i!j!} \mathbf{A}_m^i \otimes \mathbf{B}_n^j.$$

Ha most úgy rendezzük át ezt a kettős szummát, hogy rendre azokat a hatványokat gyűjtjük össze, amelyekre a kitevők összege állandó, tehát

$$i + j = p \quad (p = 0, 1, 2, \dots),$$

akkor $i = k$, $j = p - k$ helyettesítéssel az alábbi adódik:

$$\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{1}{i!j!} \mathbf{A}_m^i \otimes \mathbf{B}_n^j = \sum_{p=0}^{\infty} \sum_{k=0}^p \frac{1}{k!(p-k)!} \mathbf{A}_m^k \otimes \mathbf{B}_n^{p-k},$$

és ezzel a tételt bebizonyítottuk. ■

3.4 MÁTRIXFÜGGVÉNY ELŐÁLLÍTÁSA A MINIMÁLPOLINOM TÖBBSZÖRÖS GYÖKEI ESETÉN

Mielőtt rátérnénk az általános tárgyalásra, egyszerű példát mutatunk olyan mátrixra, amelynek minimálpolinomja többszörös gyököt is tartalmaz.

16. Példa. Határozzuk meg az $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ mátrix minimálpolinomjának gyökeit!

Megoldás. Az \mathbf{A} mátrix karakterisztikus polinomja

$$D(\lambda) = |\lambda \mathbf{E} - \mathbf{A}| = \begin{vmatrix} \lambda - 1 & -1 \\ 0 & \lambda - 1 \end{vmatrix} = (\lambda - 1)^2,$$

a karakterisztikus mátrix adjungáltja pedig $\text{adj}(\lambda \mathbf{E} - \mathbf{A}) = \begin{bmatrix} \lambda - 1 & 1 \\ 0 & \lambda - 1 \end{bmatrix}$. Az adjungált mátrix elemeinek legnagyobb közös osztója 1, ezért a minimálpolinom megegyezik a karakterisztikus polinommal: $\Delta(\lambda) = D(\lambda) = (\lambda - 1)^2$. Az adott másodrendű \mathbf{A} mátrix minimálpolinomjának tehát $\lambda = 1$ kétszeres gyöke.

* * *

3.4.1 Mátrixfüggvény előállítása Hermite-féle mátrixpolinomok segítségével

A minimálegyenlet egyszeres gyökei esetén láttuk, hogy a Lagrange-féle mátrixpolinomok segítségével sikerült tetszőleges fokszámú mátrixpolinomot – és ezáltal mátrixfüggvényt is – a minimálpolinomnál alacsonyabb fokszámú mátrixpolinomra redukálni. Most azt mutatjuk meg, hogy ha a minimálpolinomnak többszörös gyöke is van, akkor az ún. *Hermite-féle mátrixpolinomok* teszik lehetővé ezt a redukción.

Legyen az n -edrendű \mathbf{A} mátrix karakterisztikus polinomja

$$D(\lambda) = |\lambda \mathbf{E} - \mathbf{A}| = \prod_{k=1}^s (\lambda - \lambda_k)^{\alpha_k}, \quad \text{ahol} \quad \sum_{k=1}^s \alpha_k = n,$$

minimálpolinomja pedig

$$\Delta(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{\gamma_k}, \quad \text{ahol} \quad s < \sum_{k=1}^s \gamma_k = m \leq n,$$

és tegyük fel, hogy a λ_k sajátértékek az $f(z) = \sum_{k=0}^{\infty} c_k z^k$ hatványsorral definiált $f(z)$ függvény konvergenciakörének belsejébe esnek. Feladatunk az $f(\mathbf{A})$ mátrixfüggvény előállítás. A 3.2.2 pontban megadott értelmezés szerint:

$$f(\mathbf{A}) = \lim_{N \rightarrow \infty} S_N(\mathbf{A}) = \lim_{N \rightarrow \infty} \sum_{k=0}^N c_k \mathbf{A}^k,$$

vagyis az $f(\mathbf{A})$ mátrixfüggvényt mátrixpolinomok sorozatának határértékeként definiáljuk – amennyiben ez a határérték létezik. Az $f(\mathbf{A})$ mátrixfüggvény előállítását ez esetben is az teszi lehetővé, hogy az $S_N(\mathbf{A})$ részletösszegeket egy, a minimálpolinomnál alacsonyabb fokszámú polinomra redukáljuk, és így a mátrixpolinom-sorozat határértékét visszavezetjük számsorozat határértékére. Képezzük e célból ismét az $S_N(z)$ polinomnak ($N > m$) a $\Delta(z)$ minimálpolinomra vett, ezúttal legfeljebb $(m-1)$ -edfokú $R_N(z)$ osztási maradékát:

$$(3.4.1) \quad S_N(z) \equiv \Delta(z)q(z) + R_N(z).$$

Az $R_N(z)$ polinom meghatározásához most m feltétel szükséges; tehát – szemben a minimálpolinom egyszeres gyökeinek esetével – nem elegendő annak az ismerete, hogy a λ_k helyeken az értéke megegyezik az $S_N(z)$ polinomnak ott felvett értékével, mert ez mindössze $s < m$ feltételt ad. Mivel jelen esetben a minimálegyenlet λ_k gyökeinek a multiplicitása γ_k (és a γ_k értékek közül legalább az egyik nagyobb, mint 1, különben a már tárgyalt csupa egyszeres gyök esetével állnánk szemben), ezért a $\Delta(z)$ minimálpolinom első, második, \dots , $(\gamma_k - 1)$ -edik deriváltja a λ_k helyen zérus. Emiatt a (3.4.1) azonosságból következik nemcsak az, hogy $R_N(\lambda_k) = S_N(\lambda_k)$, hanem

$$(3.4.2) \quad \begin{aligned} R'_N(\lambda_k) &= S'_N(\lambda_k) \\ R''_N(\lambda_k) &= S''_N(\lambda_k) \\ &\dots\dots\dots \\ R_N^{(\gamma_k-1)}(\lambda_k) &= S_N^{(\gamma_k-1)}(\lambda_k) \quad (k = 1, 2, \dots, s). \end{aligned}$$

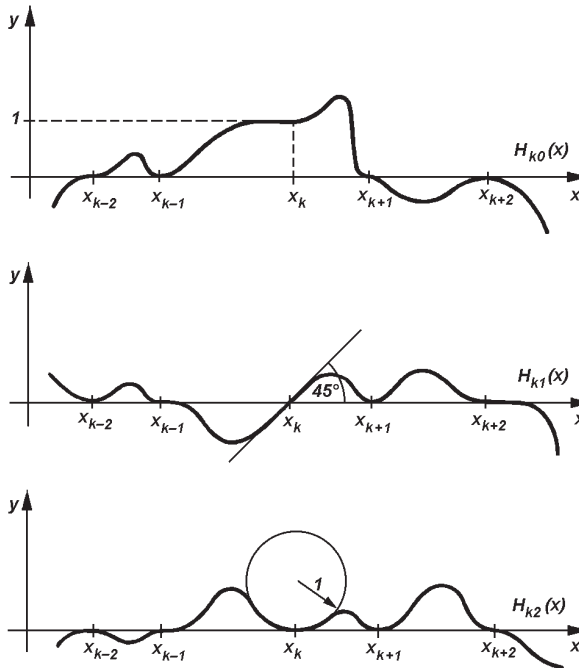
Ez összesen éppen $\sum_{k=1}^s \gamma_k = m$ feltétel, amelyek az $R_N(z)$ legfeljebb $(m-1)$ -edfokú polinomot egyértelműen meghatározzák. Azt az eljárást, amellyel ez a polinom meghatározható, *Hermite-féle interpolációnak* (bővebben lásd [18]) nevezzük. Az Hermite-féle interpoláció lényegében a Lagrange-féle interpoláció általánosítása arra az esetre, amikor az x_1, x_2, \dots, x_s alappontokban nem csak az y_1, y_2, \dots, y_s függvényértékek adottak, hanem az $y'_k, y''_k, \dots, y_k^{(\gamma_k-1)}$

szukcesszív deriváltak értékei is. Ha $\sum_{k=1}^s \gamma_k = m$, akkor m feltételünk van,

3.4 ... ELŐÁLLÍTÁSA A MINIMÁLPOLINOM TÖBBSZÖRÖS GYÖKEI ESETÉN 313

és ezekből m számú együtthatót határozhatunk meg egyértelműen, vagyis a legalacsonyabb fokú $T(x)$ racionális egész függvény legfeljebb $(m-1)$ -edfokú polinom. Most is ahhoz hasonlóan fogunk eljárni, mint ahogyan azt a Lagrange-féle interpoláció esetében tettük. Bevezetjük a kettős indexű $H_{k\nu}(x)$ ún. *Hermite-féle interpolációs alappolinomokat*, amelyek a következő tulajdonságúak:

$$(3.4.3) \quad H_{k\nu}^{(\mu)}(x_l) = \delta_{kl}\delta_{\nu\mu} \quad (k, l = 1, 2, \dots, s; \quad \nu, \mu = 0, 1, \dots, \gamma_k - 1).$$



A $k\nu$ indexű Hermite-féle alappolinomnak tehát az a tulajdonsága, hogy az x_k helyen a ν -edik deriváltja 1, de ezen a helyen a többi deriváltja a (γ_k-1) -edik deriváltig bezárólag (0-dik deriváltnak tekintve magát a függvényt), valamint az összes többi helyeken az összes előfordulható deriváltja 0 (lásd az ábrát). Az így definiált $H_{k\nu}(x)$ alappolinomok segítségével a keresett $T(x)$ függvény a következő alakban adódik:

$$T(x) = \sum_{k=1}^s \{y_k H_{k0}(x) + y'_k H_{k1}(x) + \dots + y_k^{(\gamma_k-1)} H_{k, \gamma_k-1}(x)\}.$$

A fentiekből következik, hogy az Hermite-féle interpolációs alappolinomok segítségével – a (3.2.23) képlet általánosításaként – bármely, legfeljebb $(m-1)$ -edfokú polinom egyértelműen előállítható.

Állítsuk elő eszerint a „legegyszerűbb polinomokat” – az $1, x, \frac{x^2}{2!}, \dots, \frac{x^p}{p!}$ ($p \leq m-1$) hatványokat!

Az $y \equiv 1$ függvény értéke bármely helyen 1, deriváltjainak az értéke pedig 0, tehát

$$(3.4.4) \quad 1 \equiv \sum_{k=1}^s H_{k0}(x).$$

Az $y \equiv x$ függvény értéke az x_k helyen x_k , deriváltja 1, tehát

$$x \equiv \sum_{k=1}^s \{x_k H_{k0}(x) + H_{k1}(x)\}.$$

Hasonló megfontolás alapján

$$\begin{aligned} \frac{x^2}{2!} &\equiv \sum_{k=1}^s \left\{ \frac{x_k^2}{2!} H_{k0}(x) + x_k H_{k1}(x) + H_{k2}(x) \right\}, \\ \frac{x^3}{3!} &\equiv \sum_{k=1}^s \left\{ \frac{x_k^3}{3!} H_{k0}(x) + \frac{x_k^2}{2!} H_{k1}(x) + x_k H_{k2}(x) + H_{k3}(x) \right\}, \end{aligned}$$

és általánosan

$$(3.4.5) \quad \frac{x^p}{p!} \equiv \sum_{k=1}^s \left\{ \frac{x_k^p}{p!} H_{k0}(x) + \frac{x_k^{p-1}}{(p-1)!} H_{k1}(x) + \dots + \frac{x_k^{p-q_k}}{(p-q_k)!} H_{kq_k}(x) \right\},$$

ahol $q_k = \begin{cases} \gamma_k - 1, & \text{ha } \gamma_k - 1 \leq p, \\ p, & \text{ha } \gamma_k - 1 \geq p. \end{cases}$ Ily módon tehát x p -edik hatványa ($p \leq m-1$) előállítható a minimálpolinom gyökhelyeihez tartozó Hermite-féle interpolációs alappolinomok segítségével.

17. Példa. Határozzuk meg azt a legalacsonyabb fokú polinomot, amely áthalad a $P_1(2, 4)$, $P_2(6, 2)$ pontokon, első deriváltja az $x_1 = 2$ helyen $y'_1 = \frac{1}{2}$, az $x_2 = 6$ helyen $y'_2 = \frac{1}{4}$.

Megoldás. Először az alappolinomokat számítjuk ki. Mivel négy feltételünk van, az alappolinomok harmadfokúak lesznek. A (3.4.3) definíció értelmében $H_{10}(x)$ olyan polinom, amely deriváltjával együtt zérus az $x_2 = 6$ helyen, tehát $x = 6$ kétszeres gyöke. Ezért a következő alakban kereshetők:

$$H_{10}(x) = (Ax + B)(x - 6)^2;$$

3.4 ... ELŐÁLLÍTÁSA A MINIMÁLPOLINOM TÖBBSZÖRÖS GYÖKEI ESETÉN 315

az A és B együtthatókat a $H_{10}(2) = 1$, $H'_{10}(2) = 0$ feltételekből határozzuk meg. Mivel $H'_{10}(x) = A(x-6)^2 + 2(Ax+B)(x-6)$, A és B számára a következő egyenletrendszert kapjuk:

$$(2A+B)(2-6)^2 = 1, \quad A(2-6)^2 + 2(2A+B)(2-6) = 0,$$

azaz

$$2A+B = \frac{1}{16}, \quad -8B = 0,$$

ahonnan $B = 0$, $A = \frac{1}{32}$. Tehát $H_{10}(x) = \frac{1}{32}x(x-6)^2$. A $H_{11}(x)$ alappolinom alakja ugyancsak a feltételek alapján – mivel az $x_1 = 2$ helyen a függvényérték, továbbá az $x_2 = 6$ helyen a függvényérték és első deriváltja zérus – $H_{11}(x) = C(x-2)(x-6)^2$, ahol C értékét a még fel nem használt $H'_{11}(2) = 1$ feltételből határozzuk meg, ahonnan $C = \frac{1}{16}$. Tehát $H_{11}(x) = \frac{1}{16}(x-2)(x-6)^2$. Hasonlóképpen határozhatók meg a $H_{20}(x)$ és $H_{21}(x)$ alappolinomok is: $H_{20}(x) = (Dx+E)(x-2)^2$. A feltételek:

$$H_{20}(6) = 1, \quad H'_{20}(2) = 0,$$

$$D = -\frac{1}{32}, \quad E = \frac{1}{4},$$

tehát $H_{20}(x) = -\frac{1}{32}(x-8)(x-2)^2$.

Végül $H_{21}(x) = F(x-6)(x-2)^2$.

A feltétel:

$$H'_{21}(6) = 1, \quad F = \frac{1}{16},$$

tehát $H_{21}(x) = \frac{1}{16}(x-6)(x-2)^2$.

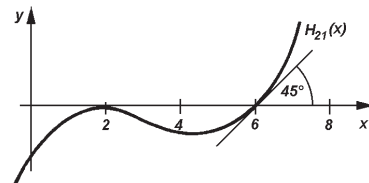
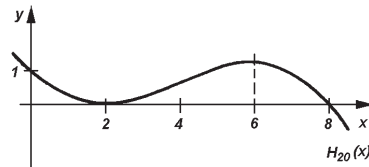
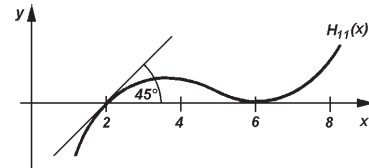
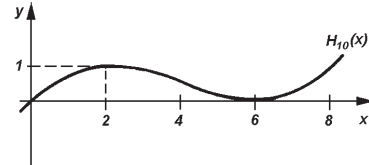
A keresett négy alappolinom tehát a következő (lásd az ábrát):

$$H_{10}(x) = \frac{1}{32}x(x-6)^2,$$

$$H_{11}(x) = \frac{1}{16}(x-2)(x-6)^2,$$

$$H_{20}(x) = -\frac{1}{32}(x-8)(x-2)^2,$$

$$H_{21}(x) = \frac{1}{16}(x-6)(x-2)^2.$$



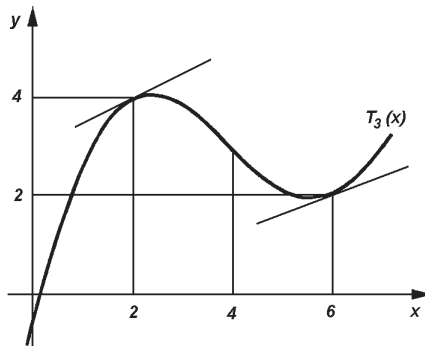
Ellenőrzésképpen győződjünk meg arról, hogy teljesül-e a (3.4.4) azonosság:

$$H_{10}(x) + H_{20}(x) = \frac{1}{32} \{ (x-6)^2 x - (x-2)^2 (x-8) \} \equiv 1.$$

A keresett polinomra tehát a következő adódik:

$$\begin{aligned} T(x) &= y_1 H_{10}(x) + y'_1 H_{11}(x) + y_2 H_{20}(x) + y'_2 H_{21}(x) = \\ &= \frac{1}{8} x(x-6)^2 + \frac{1}{32} (x-2)(x-6)^2 - \frac{1}{16} (x-8)(x-2)^2 + \\ &\quad + \frac{1}{64} (x-6)(x-2)^2, \end{aligned}$$

ill. összevonás után $T(x) = \frac{7}{64}x^3 - \frac{43}{32}x^2 + \frac{73}{16}x - \frac{5}{8}$ (lásd az ábrát).



* * *

Térjünk vissza eredeti feladatunkra, az $R_N(z)$ legfeljebb $(m-1)$ -edfokú polinom meghatározására. Mivel a (3.4.2) összefüggések alapján összesen m feltételünk van (ismerjük ugyanis s különböző helyen a polinomnak és egymás után következő deriváltjainak összesen m számú értékét), az e helyekhez tartozó Hermite-féle alappolinomok segítségével a

$$\begin{aligned} (3.4.6) \quad R_N(z) &= \sum_{k=1}^s [S_N(\lambda_k) H_{k0}(z) + S'_N(\lambda_k) H_{k1}(z) + \cdots + \\ &\quad + S_N^{(\gamma_k-1)}(\lambda_k) H_{k, \gamma_k-1}(z)] \end{aligned}$$

alakban írhatjuk fel az $R_N(z)$ polinomot. Ha most z helyére az \mathbf{A} mátrixot helyettesítjük, akkor a 3.2.2 tétel alapján a (3.4.1) összefüggésből következik,

3.4 ... ELŐÁLLÍTÁSA A MINIMÁLPOLINOM TÖBBSZÖRÖS GYÖKEI ESETÉN 317

hogy $S_N(\mathbf{A}) = R_N(\mathbf{A})$, amiből (3.4.6) behelyettesítésével adódik:

$$S_N(\mathbf{A}) = \sum_{k=1}^s \{ S_N(\lambda_k) H_{k0}(\mathbf{A}) + S'_N(\lambda_k) H_{k1}(\mathbf{A}) + \cdots + \\ + S_N^{(\gamma_k-1)}(\lambda_k) H_{k, \gamma_k-1}(\mathbf{A}) \}.$$

Végül az $N \rightarrow \infty$ határátmenetet elvégezve, a keresett $f(\mathbf{A})$ mátrixfüggvényt a következő alakban nyerjük:

$$(3.4.7) \quad f(\mathbf{A}) = \sum_{k=1}^s \sum_{\nu=0}^{\gamma_k-1} f^{(\nu)}(\lambda_k) H_{k\nu}(\mathbf{A}).$$

A mátrixfüggvények előállítására vonatkozó eredményt tétel formájában is kimondjuk:

3.4.1 tétel. *Ha az \mathbf{A} mátrix karakterisztikus polinomja*

$$D(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{\alpha_k}, \quad \text{ahol} \quad \sum_{k=1}^s \alpha_k = n,$$

minimálpolinomja pedig

$$\Delta(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{\gamma_k}, \quad \text{ahol} \quad s < \sum_{k=1}^s \gamma_k = m \leq n,$$

és a λ_k sajátértékek az $f(z) = \sum_{k=0}^{\infty} c_k z^k$ hatványsorral definiált függvény konvergenciakörének belsejébe esnek, akkor az $f(\mathbf{A})$ mátrixfüggvény a λ_k helyekhez tartozó, a

$$H_{k\nu}^{(\mu)}(\lambda_l) = \delta_{kl} \delta_{\mu\nu} \quad (k, l = 1, 2, \dots, s; \quad \mu, \nu = 0, 1, 2, \dots, \gamma_k - 1)$$

összefüggésekkel definiált $H_{k\nu}(z)$ Hermite-féle interpolációs alappolinomok segítségével előállítható az

$$f(\mathbf{A}) = \sum_{k=1}^s \sum_{\nu=0}^{\gamma_k-1} f^{(\nu)}(\lambda_k) H_{k\nu}(\mathbf{A})$$

alakban.

Az előállításban előforduló $H_{k\nu}(\mathbf{A})$ mátrixpolinomokat *Hermite-féle mátrixpolinomoknak* nevezzük. A következőkben ezek tulajdonságaival foglalkozunk.

3.4.2 Az Hermite-féle mátrixpolinomok tulajdonságai

(a) Először a $H_{k0}(\mathbf{A})$ mátrixpolinomot vizsgáljuk.

3.4.2 tétel. *A $H_{k0}(\mathbf{A})$ Hermite-féle mátrixpolinom α_k rangú projektor, és érvényes a*

$$(3.4.8) \quad H_{k0}(\mathbf{A})H_{l0}(\mathbf{A}) = \delta_{kl}H_{k0}(\mathbf{A})$$

összefüggés.

Bizonyítás. A $H_{k0}(z)$ alappolinom definíciójából következik, hogy a $H_{k0}(z)H_{l0}(z)$ szorzat $k \neq l$ esetén osztható a minimálpolinommal:

$$(3.4.9) \quad H_{k0}(z)H_{l0}(z) = \Delta(z)q_1(z), \quad \text{ha} \quad k \neq l;$$

tehát z helyére az \mathbf{A} mátrixot helyettesítve, a 3.2.2 tétel alapján $k \neq l$ esetén

$$(3.4.10) \quad H_{k0}(\mathbf{A})H_{l0}(\mathbf{A}) = \mathbf{0}.$$

Másrészt a (3.4.4) azonosságban x helyére az \mathbf{A} mátrixot írva,

$$(3.4.11) \quad \sum_{k=1}^s H_{k0}(\mathbf{A}) = \mathbf{E}$$

adódik. Ha ennek az egyenlőségnek mindkét oldalát a $H_{l0}(\mathbf{A})$ mátrixszal szorozzuk, akkor (3.4.10) felhasználásával a (3.4.8) összefüggést kapjuk, $H_{k0}(\mathbf{A})$ tehát valóban projektor. Be kell még látni, hogy $\varrho(H_{k0}(\mathbf{A})) = \alpha_k$. Ennek bizonyításához felhasználjuk a $H_{k0}(z)$ polinom definíciójából következő

$$(z - \lambda_k)^{\gamma_k} H_{k0}(z) = \Delta(z)q_2(z)$$

azonosságot. Ha itt z helyére az \mathbf{A} mátrixot helyettesítjük, akkor ismét a 3.2.2 tétel alapján

$$(3.4.12) \quad (\mathbf{A} - \lambda_k \mathbf{E})^{\gamma_k} H_{k0}(\mathbf{A}) = \mathbf{0}.$$

Az $(\mathbf{A} - \lambda_k \mathbf{E})^{\gamma_k}$ tényező rangja legalább $n - \alpha_k$. Ugyanis, ha a 3.1.4 tételt alkalmazzuk, és az \mathbf{A} mátrixot alkalmas unitér transzformációval háromszögmátrixra transzformáljuk, akkor a 2.6.2 tétel értelmében sajátértékei nem változnak, tehát a háromszögmátrixra transzformált $\mathbf{A} - \lambda_k \mathbf{E}$ mátrixnak, sőt az $(\mathbf{A} - \lambda_k \mathbf{E})^{\gamma_k}$ mátrixnak is van zérustól különböző, $(n - \alpha_k)$ -rendű minora, így a rangja valóban legalább $n - \alpha_k$. Az 1.3.13 tétel értelmében tehát a (3.4.12) szorzat csak akkor lehet zérus, ha a tényezők rangjának összege nem nagyobb, mint n , azaz ha $H_{k0}(\mathbf{A})$ rangja legfeljebb α_k :

$$(3.4.13) \quad \varrho(H_{k0}(\mathbf{A})) \leq \alpha_k.$$

3.4 ... ELŐÁLLÍTÁSA A MINIMÁLPOLINOM TÖBBSZÖRÖS GYÖKEI ESETÉN 319

Másrészt a (3.4.11) egyenlőségre alkalmazva az 1.3.5 tételt, mely szerint mátrixok összegének a rangja nem lehet nagyobb az egyes tagok rangjának összegénél, a következő becslést kapjuk:

$$(3.4.14) \quad n = \varrho(\mathbf{E}) = \varrho\left(\sum_{k=1}^s H_{k0}(\mathbf{A})\right) \leq \sum_{k=1}^s \varrho(H_{k0}(\mathbf{A})).$$

Ha tehát a (3.4.13) összefüggésben az egyenlőtlenség jele valamelyik k esetén fennállna, akkor ellentmondásra jutnánk a (3.4.14) összefüggésben. Ebből következik, hogy

$$(3.4.15) \quad \varrho(H_{k0}(\mathbf{A})) = \alpha_k$$

valóban fennáll. ■

(b) Ezek után a $H_{k\nu}(\mathbf{A})$ ($\nu > 0$) mátrixpolinom és a $H_{k0}(\mathbf{A})$ mátrixpolinom közötti kapcsolatot vizsgáljuk. Először megmutatjuk, hogy a $H_{k\nu}(\mathbf{A})$ mátrixpolinomok $\nu > 0$ esetén kifejezhetők a $H_{k0}(\mathbf{A})$ mátrixpolinom segítségével.

Tekintsük először a

$$(3.4.16) \quad H_{k1}(z) - (z - \lambda_k)H_{k0}(z)$$

különbséget. Ez a λ_k helyen zérus, első deriváltja,

$$H'_{k1}(z) - H_{k0}(z) - (z - \lambda_k)H'_{k0}(z),$$

a λ_k helyen szintén zérus (az első két tag $H_{k\nu}(z)$ definíciója értelmében 1–1); a magasabb deriváltak a $(\gamma_k - 1)$ -edikig bezárólag, az Hermite-féle alappolinomok definíciója alapján, ugyancsak zérussal egyenlők. A (3.4.16) különbség tehát osztható a minimálpolinommal:

$$H_{k1}(z) - (z - \lambda_k)H_{k0}(z) = \Delta(z)q_3(z).$$

Behelyettesítve z helyére az \mathbf{A} mátrixot, a 3.2.2 tétel alapján az alábbi összefüggést kapjuk:

$$H_{k1}(\mathbf{A}) = (\mathbf{A} - \lambda_k \mathbf{E})H_{k0}(\mathbf{A}).$$

Hasonló megfontolással belátható, hogy a

$$H_{k\nu}(z) - \frac{(z - \lambda_k)^\nu}{\nu!} H_{k0}(z) \quad (\nu = 1, 2, \dots, \gamma_k - 1)$$

különbség szintén osztható a minimálpolinommal (ennek a deriváltjai is $(\gamma_k - 1)$ -edrendig bezárólag zérust adnak a λ_k helyeken), tehát z helyére az \mathbf{A} mátrixot helyettesítve,

$$(3.4.17) \quad H_{k\nu}(\mathbf{A}) = \frac{1}{\nu!} (\mathbf{A} - \lambda_k \mathbf{E})^\nu H_{k0}(\mathbf{A}) \quad (\nu = 1, 2, \dots, \gamma_k - 1).$$

(c) Most az összes $H_{k\nu}(\mathbf{A})$ mátrixpolinom közötti összefüggést mutatjuk meg. Ennek segítségével is meghatározhatók a mátrixpolinomok. Ha a (3.4.5) összefüggésben x helyére az \mathbf{A} mátrixot helyettesítjük, akkor az alábbi egyenlőséget nyerjük:

$$(3.4.18) \quad \frac{1}{p!} \mathbf{A}^p = \sum_{k=1}^s \left\{ \frac{\lambda_k^p}{p!} H_{k0}(\mathbf{A}) + \frac{\lambda_k^{p-1}}{(p-1)!} H_{k1}(\mathbf{A}) + \cdots + \frac{\lambda_k^{p-q_k}}{(p-q_k)!} H_{kq_k}(\mathbf{A}) \right\},$$

$$\text{ahol} \quad q_k = \begin{cases} \gamma_k - 1, & \text{ha } \gamma_k - 1 \leq p, \\ p, & \text{ha } \gamma_k - 1 \geq p. \end{cases}$$

Ha a (3.4.18) egyenlőséget a $p = 0, 1, \dots, m-1$ értékekre felírjuk, akkor az m számú $H_{k\nu}(\mathbf{A})$ mátrixpolinomra olyan lineáris egyenletrendszert kapunk, amelyből ezek kiszámíthatók.

(d) A $H_{k\nu}$ mátrixpolinom fontos tulajdonságát bizonyítjuk be a következő tételben.

3.4.3 tétel. A $H_{k\nu}(\mathbf{A})$ mátrixpolinom $\nu > 0$ esetén nilpotens mátrix.

Bizonyítás. Minden $k\nu$ indexpárhoz ($\nu > 0$) található olyan q pozitív egész szám, amellyel $q\nu \geq \gamma_k$. Ha a (3.4.17) egyenlőség mindkét oldalát erre a q hatványra emeljük, akkor a (3.4.12) összefüggés felhasználásával $H_{k\nu}^q(\mathbf{A}) = 0$ ($\nu > 0$) adódik, tehát $H_{k\nu}(\mathbf{A})$ valóban nilpotens mátrix. ■

(e) Abban a speciális esetben, amikor a minimálpolinomnak pl. λ_1 egyszeres gyöke, a hozzá tartozó $H_{10}(\mathbf{A})$ mátrixpolinom a 3.2.9 tételben szereplő (3.2.53) képlethez hasonló módon számítható. Erre vonatkozik az alábbi tétel.

3.4.4 tétel. Legyen az \mathbf{A} mátrix karakterisztikus polinomja

$$D(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{\alpha_k}, \quad \text{ahol} \quad \sum_{k=1}^s \alpha_k = n,$$

minimálpolinomja

$$\Delta(\lambda) = (\lambda - \lambda_1) \prod_{k=2}^s (\lambda - \lambda_k)^{\gamma_k}, \quad \text{ahol} \quad \sum_{k=2}^s \gamma_k = m - 1,$$

és az $\text{adj}(\lambda \mathbf{E} - \mathbf{A})$ mátrix elemeinek legnagyobb közös osztója $\theta(\lambda)$. Az $\mathbf{F}(\lambda) = \frac{\text{adj}(\lambda \mathbf{E} - \mathbf{A})}{\theta(\lambda)}$ redukált adjungált segítségével a minimálpolinom egyszeres λ_1 gyökéhez tartozó $H_{10}(\mathbf{A})$ mátrixpolinom a

$$(3.4.19) \quad H_{10}(\mathbf{A}) = \frac{1}{\Delta'(\lambda_1)} \mathbf{F}(\lambda_1)$$

képlettel számítható.

3.4 ... ELŐÁLLÍTÁSA A MINIMÁLPOLINOM TÖBBSZÖRÖS GYÖKEI ESETÉN 321

Bizonyítás. Alkalmazzuk ismét a 3.2.9 tétel bizonyításának a gondolatmenetét, és állítsuk elő a (3.2.55) függvényt – amely most y -ban legfeljebb $(m-1)$ -edfokú racionális egész függvény – a megfelelő Hermite-féle alappolinomok segítségével:

$$\frac{\Delta(x) - \Delta(y)}{x - y} \equiv \frac{\Delta(x) - \Delta(\lambda_1)}{x - \lambda_1} H_{10}(y) + \sum_{k=2}^s \sum_{\nu=0}^{\gamma_k-1} \left(\frac{\Delta(x) - \Delta(y)}{x - y} \right)_{y=\lambda_k}^{(\nu)} H_{k\nu}(y).$$

Vegyük figyelembe, hogy $H_{k\nu}(y)$ együtthatói – mint y polinomjának a deriváltjai – maguk is polinomok, amelyek $y = \lambda_k$ helyettesítés után csak x -től függnek. A szorzat deriválási szabályának alkalmazásával egyszerűen igazolható, hogy

$$\left(\frac{\Delta(x) - \Delta(y)}{x - y} \right)^{(\nu)} \equiv \frac{\nu! \Delta(x)}{(x - y)^{\nu+1}} - \sum_{i=0}^{\nu} \frac{\nu!}{i!} \frac{\Delta^{(i)}(y)}{(x - y)^{\nu-i+1}}.$$

Ha most y helyére behelyettesítjük a λ_k gyököket, és figyelembe vesszük, hogy ezek multiplicitása γ_k , azaz $\Delta^{(i)}(\lambda_k) = 0$ ($i = 0, 1, 2, \dots, \gamma_k - 1$, $k = 2, 3, \dots, s$), akkor

$$\left(\frac{\Delta(x) - \Delta(y)}{x - y} \right)_{y=\lambda_k}^{(\nu)} = \frac{\nu! \Delta(x)}{(x - \lambda_k)^{\nu+1}}$$

írható. Mivel $\Delta(\lambda_1) = 0$ (de $\Delta'(\lambda_1) \neq 0$), továbbá $H_{10}(x) = \frac{\Delta(x)}{\Delta'(\lambda_1)(x - \lambda_1)}$, ezek behelyettesítésével a fenti azonosságra

$$\frac{\Delta(x) - \Delta(y)}{x - y} \equiv \Delta'(\lambda_1) H_{10}(x) H_{10}(y) + \sum_{k=2}^s \sum_{\nu=0}^{\gamma_k-1} \frac{\nu! \Delta(x)}{(x - \lambda_k)^{\nu+1}} H_{k\nu}(y)$$

adódik. Ha y helyére az \mathbf{A} és x helyére a $\lambda \mathbf{E}$ mátrixot írjuk, és figyelembe vesszük a $(\lambda \mathbf{E} - \mathbf{A})^{-1}$ reciprokmátrix definícióját, valamint a 3.2.2 tételt, akkor a következő összefüggést nyerjük:

$$\frac{\text{adj}(\lambda \mathbf{E} - \mathbf{A})}{\theta(\lambda)} = \Delta'(\lambda_1) H_{10}(\lambda) H_{10}(\mathbf{A}) + \sum_{k=2}^s \sum_{\nu=0}^{\gamma_k-1} \frac{\nu! \Delta(\lambda)}{(\lambda - \lambda_k)^{\nu+1}} H_{k\nu}(\mathbf{A}).$$

Ha most λ helyére a minimálpolinom egyszeres λ_1 gyökét behelyettesítjük, és tekintetbe vesszük azt, hogy a definíció alapján $H_{10}(\lambda_1) = 1$ és

$$\left\{ \frac{\Delta(\lambda)}{(\lambda - \lambda_k)^{\nu+1}} \right\}_{\lambda=\lambda_1} = 0 \quad (\nu = 0, 1, 2, \dots, \gamma_k - 1, \quad k = 2, 3, \dots, s),$$

akkor innen (3.4.19) következik. ■

3.4.3 Mátrixok kvázidiagonalizálása

Az eddigi eredményeinket a következőkben az $f(z) \equiv z$ függvényre alkalmazzuk. Ekkor ugyanis (3.4.18) szerint az \mathbf{A} mátrixot a $H_{k\nu}(\mathbf{A})$ Hermite-féle mátrixpolinomok lineáris kombinációjaként állítjuk elő. Az alábbiakban ennek segítségével megmutatjuk, hogy a minimálpolinom többszörös gyökei esetén a mátrix nem diagonalizálható, viszont *kvázidiagonalizálható* – vagyis hasonlósági transzformációval olyan particionált alakra hozható, amelynek csak diagonálblokkjai (főátlóban álló blokkjai) különböznek zérustól.

Itt ki kell majd használnunk a nilpotens mátrixnak azt a tulajdonságát, hogy hasonlósági transzformációval nem diagonalizálható. Először bizonyítjuk a következő tételt.

3.4.5 tétel. *A nilpotens mátrixok valamennyi sajátértéke zérussal egyenlő.*

Bizonyítás. Ha a nilpotens mátrixot uniter transzformációval háromszög-mátrixra transzformáljuk és hatványozzuk, akkor a főátlóban elhelyezkedő sajátértékek hatványozódnak. Mivel a mátrix valamely egész kitevős hatványa 0, ebből következik, hogy valamennyi sajátértéke 0. ■

A tételből következik, hogy ha a nilpotens mátrix hasonlósági transzformációval diagonalizálható lenne, akkor a zérusmátrixhoz lenne hasonló, ami azonban lehetetlen.

Tekintsük most a (3.4.18) összefüggést $p = 1$ esetén:

$$(3.4.20) \quad \mathbf{A} = \sum_{k=1}^s \{ \lambda_k H_{k0}(\mathbf{A}) + H_{k1}(\mathbf{A}) \}.$$

A 3.4.2 és a 3.4.3 tétel szerint az \mathbf{A} mátrixot ezzel projektorok (idempotens mátrixok) és nilpotens mátrixok lineáris kombinációjaként állítottuk elő. Mivel a nilpotens mátrixok nem diagonalizálhatók, ebből már lehet sejteni, hogy az \mathbf{A} mátrix sem diagonalizálható hasonlósági transzformációval. Miután belátjuk, hogy ez valóban így van, a továbbiakban annak vizsgálatával foglalkozunk, hogy ilyen esetben az \mathbf{A} mátrixot milyen legegyszerűbb alakra lehet transzformálni.

Helyettesítsük be a $H_{k1}(\mathbf{A})$ mátrixpolinom (3.4.17) alatti kifejezését – amelyet $\nu = 1$ esetén kapunk – a (3.4.20) egyenlőségbe:

$$(3.4.21) \quad \mathbf{A} = \sum_{k=1}^s \{ \lambda_k H_{k0}(\mathbf{A}) + (\mathbf{A} - \lambda_k \mathbf{E}) H_{k0}(\mathbf{A}) \}.$$

Vegyük figyelembe azt, hogy $H_{k0}(\mathbf{A})$ projektor, tehát a helyére $H_{k0}^2(\mathbf{A})$ is írható, és azt, hogy \mathbf{A} polinomjai felcserélhetőek. Így a fenti összefüggés új

3.4 ... ELŐÁLLÍTÁSA A MINIMÁLPOLINOM TÖBBSZÖRÖS GYÖKEI ESETÉN 323

alakja:

$$(3.4.22) \quad \mathbf{A} = \sum_{k=1}^s \{ \lambda_k H_{k0}(\mathbf{A}) + H_{k0}(\mathbf{A})(\mathbf{A} - \lambda_k \mathbf{E}) H_{k0}(\mathbf{A}) \}.$$

Legyen az α_k rangú $H_{k0}(\mathbf{A})$ projektor egy minimális diadikus felbontása

$$(3.4.23) \quad H_{k0}(\mathbf{A}) = \mathbf{U}_k \mathbf{V}_k^T.$$

Az 1.6.3 tétel értelmében \mathbf{U}_k oszlopvektorai és \mathbf{V}_k^T sorvektorai nemteljes biortogonális vektorrendszert alkotnak. Ha most figyelembe vesszük a (3.4.10) összefüggést, akkor onnan $\mathbf{U}_k \mathbf{V}_k^T \mathbf{U}_l \mathbf{V}_l^T = \mathbf{0}$ ($k \neq l$), vagyis az 1.3.10. tétel alapján

$$(3.4.24) \quad \mathbf{V}_k^T \mathbf{U}_l = \mathbf{0} \quad (k \neq l),$$

a (3.4.11) összefüggésből pedig az $\mathbf{U} = [\mathbf{U}_1 \mathbf{U}_2 \dots \mathbf{U}_s]$, $\mathbf{V}^T = \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \\ \vdots \\ \mathbf{V}_s^T \end{bmatrix}$ jelöléssel

$\sum_{k=1}^s \mathbf{U}_k \mathbf{V}_k^T = \mathbf{U} \mathbf{V}^T = \mathbf{E}$ következik. Tehát

$$(3.4.25) \quad \mathbf{V}^T = \mathbf{U}^{-1},$$

és így az összes $H_{k0}(\mathbf{A})$ mátrix diadikus felbontásával nyert \mathbf{u}_ν és \mathbf{v}_ν^T vektorok teljes biortogonális vektorrendszert alkotnak. Ha most a (3.4.23) felbontást behelyettesítjük a (3.4.22) összefüggésbe, és a kapcsos zárójelből balra kiemeljük az \mathbf{U}_k tényezőt, jobbra pedig a \mathbf{V}_k^T tényezőt, akkor

$$(3.4.26) \quad \mathbf{A} = \sum_{k=1}^s \mathbf{U}_k \{ \lambda_k \mathbf{E}_{\alpha_k} + \mathbf{V}_k^T (\mathbf{A} - \lambda_k \mathbf{E}) \mathbf{U}_k \} \mathbf{V}_k^T.$$

(Itt és a továbbiakban az n -től különböző rendszámú egységmátrix rendjét alsó indexben feltüntetjük.) Felhasználva a (3.4.24) ortogonalitási relációt, az alábbi összefüggéseket kapjuk:

$$(3.4.27) \quad \mathbf{A} \mathbf{U}_p = \mathbf{U}_p \{ \lambda_p \mathbf{E}_{\alpha_p} + \mathbf{V}_p^T (\mathbf{A} - \lambda_p \mathbf{E}) \mathbf{U}_p \},$$

valamint

$$(3.4.28) \quad \mathbf{V}_q^T \mathbf{A} = \{ \lambda_q \mathbf{E}_{\alpha_q} + \mathbf{V}_q^T (\mathbf{A} - \lambda_q \mathbf{E}) \mathbf{U}_q \} \mathbf{V}_q^T,$$

végül

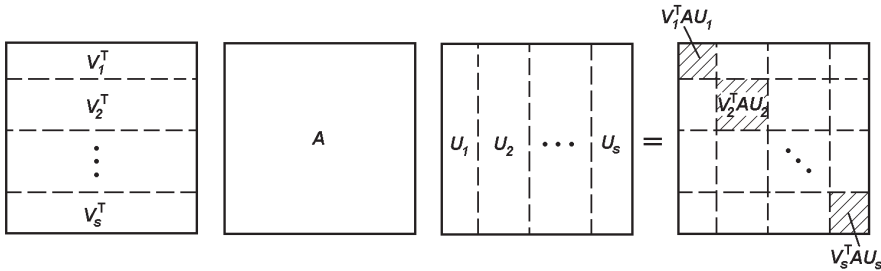
$$(3.4.29) \quad \mathbf{V}_q^T \mathbf{A} \mathbf{U}_p = \delta_{pq} \{ \lambda_p \mathbf{E}_{\alpha_p} + \mathbf{V}_p^T (\mathbf{A} - \lambda_p \mathbf{E}) \mathbf{U}_p \}.$$

Vázlatosan szemléltetve:

Innen kiolvasható, hogy \mathbf{U}_p oszlopvektorai – mint egy adott bázishoz tartozó koordinátavektorok – olyan α_p -dimenziós alteret generálnak, amely az \mathbf{A} mátrix által meghatározott transzformációval szemben invariáns. A \mathbf{V}_p^T sorvektorai pedig – mint ugyanehhez a bázishoz tartozó koordinátavektorok – az \mathbf{A}^T transzponált mátrix által meghatározott transzformációval szemben invariáns alteret generálnak. Ebből az is következik, hogy ha az \mathbf{A} mátrixot balról a \mathbf{V}^T , jobbról pedig az \mathbf{U} mátrixszal szorozzuk, vagyis ha a (3.4.25) összefüggés alapján az \mathbf{U} mátrixszal végzett hasonlósági transzformációnak vetjük alá, akkor olyan particionált mátrixot kapunk, amelynek csak diagonálblokkjaiban találhatók zérustól különböző elemek: az ilyen mátrixot *kvázidiagonál-mátrixnak* nevezzük. A zérustól különböző blokkok rendszáma rendre $\alpha_1, \alpha_2, \dots, \alpha_s$, tehát megegyezik a sajátértékeknek a karakterisztikus polinombeli multiplicitásával.

3.4 ... ELŐÁLLÍTÁSA A MINIMÁLPOLINOM TÖBBSZÖRÖS GYÖKEI ESETÉN 325

A mondottakat összefoglalva és szemléltetve, a következőket állapíthatjuk meg. Ha az n -edrendű \mathbf{A} mátrix λ_k sajátértékeinek ($k = 1, 2, \dots, s$) multiplicitása a karakterisztikus polinomban α_k , a minimálpolinom fokszáma pedig m (ahol $s < m \leq n$), akkor az α_k rangú és $(m-1)$ -edfokú $H_{k0}(\mathbf{A})$ mátrixpolinomok minimális diadikus felbontásával nyert $\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_s$, ill. $\mathbf{V}_1^T, \mathbf{V}_2^T, \dots, \mathbf{V}_s^T$ tényezőkből összeállított \mathbf{U} és \mathbf{V}^T mátrixok segítségével végrehajtott hasonlósági transzformáció az \mathbf{A} mátrixot kvázidiagonál-mátrix alakba viszi át:



Találtunk tehát egy olyan hasonlósági transzformációt, amellyel az adott mátrix kvázidiagonál-mátrixszá transzformálható; ennek diagonálblokkjai (3.4.29) szerint $\mathbf{V}_p^T \mathbf{A} \mathbf{U}_p = \lambda_p \mathbf{E}_{\alpha_p} + \mathbf{V}_p^T (\mathbf{A} - \lambda_p \mathbf{E}) \mathbf{U}_p$ ($p = 1, 2, \dots, s$).

Ezek vizsgálata azt mutatja, hogy ha a minimálegyenletnek többszörös gyöke is van, akkor nem létezik olyan hasonlósági transzformáció, amellyel ezek a blokkok diagonalizálhatók, és így az adott mátrix sem diagonalizálható. Tekintsük ugyanis a főatlóban levő

$$(3.4.30) \quad \lambda_p \mathbf{E}_{\alpha_p} + \mathbf{V}_p^T (\mathbf{A} - \lambda_p \mathbf{E}) \mathbf{U}_p$$

blokkokat. Az első tag az α_p rendű egységmátrix többszöröse, a második tagról pedig könnyen belátható, hogy nilpotens mátrix. Képezzük ebből a célból a második tag γ_p -edik hatványát, és vegyük figyelembe, hogy $\mathbf{U}_p \mathbf{V}_p^T = H_{p0}(\mathbf{A})$. Így módon az alábbi összefüggést kapjuk:

$$\begin{aligned} & \overset{1}{\mathbf{V}_p^T (\mathbf{A} - \lambda_p \mathbf{E}) \mathbf{U}_p} \overset{2}{\mathbf{V}_p^T (\mathbf{A} - \lambda_p \mathbf{E}) \mathbf{U}_p} \dots \overset{\gamma_p}{\mathbf{V}_p^T (\mathbf{A} - \lambda_p \mathbf{E}) \mathbf{U}_p} = \\ & = \mathbf{V}_p^T (\mathbf{A} - \lambda_p \mathbf{E}) H_{p0}(\mathbf{A}) (\mathbf{A} - \lambda_p \mathbf{E}) H_{p0}(\mathbf{A}) \dots H_{p0}(\mathbf{A}) (\mathbf{A} - \lambda_p \mathbf{E}) \mathbf{U}_p. \end{aligned}$$

Tekintettel arra, hogy \mathbf{A} polinomjai felcserélhetőek és $H_{p0}(\mathbf{A})$ projektor, tehát minden hatványa önmaga, a fenti kifejezésre a következő adódik:

$$\mathbf{V}_p^T (\mathbf{A} - \lambda_p \mathbf{E})^{\gamma_p} H_{p0}(\mathbf{A}) \mathbf{U}_p;$$

ez a (3.4.12) összefüggés alapján zérussal egyenlő. A (3.4.30) kifejezés második tagja tehát valóban nilpotens mátrix. A nilpotens mátrixok azonban –

mint azt a 3.4.5 tétel következményeként megmutattuk – hasonlósági transzformációval nem diagonalizálhatók, és mivel az egységmátrix és ezáltal a (3.4.30) kifejezés első tagja hasonlósági transzformáció esetén nem változik, tehát a (3.4.30) kifejezés, vagyis a kvázidiagonál-mátrix alakra transzformált \mathbf{A} mátrix főátlójában levő blokkok hasonlósági transzformációval valóban nem diagonalizálhatók.

Abban az esetben, amikor a minimálegyenletnek csupa egyszeres gyöke van, a (3.4.20) egyenlőség jobb oldalán $H_{k1}(\mathbf{A})$ nem szerepel, következésképpen a (3.4.30) kifejezés második tagja is hiányzik, és így a kvázidiagonál-mátrix diagonálmátrixszá redukálódik. Vagyis speciális esetként megkapjuk az egyszerű struktúrájú mátrixok diagonálalakra való transzformációját.

Ezzel tehát megmutattuk, hogy *ha egy mátrix minimálegyenletének van többszörös gyöke, akkor a mátrix hasonlósági transzformációval nem diagonalizálható, tehát a mátrix diagonalizálhatóságának nemcsak elégséges, hanem szükséges feltétele is, hogy a minimálegyenletnek csupa egyszeres gyöke legyen.*

3.4.4 Nilpotens mátrixok transzformációja Jordan-féle normálalakra

A következőkben azzal a kérdéssel foglalkozunk, hogy hasonlósági transzformációval milyen legegyszerűbb alakra transzformálhatók a nem diagonalizálható – nem egyszerű struktúrájú – mátrixok. Mivel azt már beláttuk, hogyan lehet a nem egyszerű struktúrájú mátrixokat hasonlósági transzformációval olyan kvázidiagonál-mátrix alakra transzformálni, amelynek minden (főátlóban levő) blokkja az egységmátrix többszörösének és egy nilpotens mátrixnak az összege, ezért a továbbiakban elegendő a nilpotens mátrixok redukciójával foglalkozni (lásd [28]).

Már az első fejezetben megismerkedtünk a nilpotens mátrixok definíciójával. Mielőtt a nilpotens mátrixok elméletébe mélyebb bepillantást nyerénénk, szükségünk van néhány újabb fogalom bevezetésére.

3.4.1 definíció. Az n -edrendű \mathbf{N} mátrixot α indexű nilpotens mátrixnak nevezzük, ha

$$(3.4.31) \quad \mathbf{N}^{\alpha-1} \neq \mathbf{0} \quad \text{de} \quad \mathbf{N}^{\alpha} = \mathbf{0}.$$

3.4.2 definíció. Ha egy nilpotens mátrix indexe megegyezik a rendszámával, akkor azt **nemderogatórius nilpotens mátrixnak** nevezzük.

3.4.3 definíció. Ha egy nilpotens mátrix indexe kisebb, mint a rendszáma, akkor azt **derogatórius nilpotens mátrixnak** nevezzük.

Megjegyzés. A nilpotens mátrix indexe nem lehet a rendszámánál nagyobb. Ugyanis a 3.4.4 tétel miatt a nilpotens mátrix valamennyi sajátértéke 0, tehát karakterisztikus egyenlete

3.4 ... ELŐÁLLÍTÁSA A MINIMÁLPOLINOM TÖBBSZÖRÖS GYÖKEI ESETÉN 327

$\lambda^n = 0$. A Cayley–Hamilton-tétel értelmében karakterisztikus egyenletét kielégíti, tehát n -edik hatványa biztosan a zérusmátrix.

(a) Először tekintsük a *nemderogatórius nilpotens* mátrixokat és vizsgáljuk meg, milyen legegyszerűbb alakra transzformálhatók alkalmasan választott hasonlósági transzformációval. A „legegyszerűbb” alakon azt értjük, hogy a lehető legkevesebb elem legyen zérustól különböző. Mivel a hasonlósági transzformáció tulajdonképpen a bázis transzformációját (a koordináta-rendszer transzformációját) jelenti, feladatunk a következőképpen is megfogalmazható: adott nilpotens lineáris transzformáció mátrixa milyen bázis választása esetén a legegyszerűbb. Látni fogjuk, hogy a nilpotens transzformációk mátrixának legegyszerűbb alakja az ún. Jordan-blokkokból álló kvázidiagonál-mátrix.

3.4.4 definíció. *Tetszőleges értékű λ mellett a*

$$\mathbf{J}_\alpha = \begin{bmatrix} \lambda & 0 & \dots & 0 \\ 1 & \lambda & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \vdots \\ 0 & 0 & \dots & 1 & \lambda \end{bmatrix} \begin{matrix} (1) \\ (2) \\ \\ (\alpha) \end{matrix}$$

alakú mátrixot α rendű alsó Jordan-blokknak nevezzük, az ilyen diagonálblokkokból álló kvázidiagonál-mátrixot pedig alsó Jordan-féle normálalakú mátrixnak.*

Ezek után bebizonyítjuk a következő tételt.

3.4.6 tétel. *Minden α rendű nemderogatórius nilpotens mátrix hasonlósági transzformációval alsó Jordan-féle normálalakra hozható; a normálalak mátrixa egyetlen α rendű alsó Jordan-blokk.*

Bizonyítás. Ha az α rendű \mathbf{N} nilpotens mátrix nemderogatórius, akkor indexe is α , tehát érvényes rá a (3.4.31) összefüggés. Tegyük most fel, hogy található olyan \mathbf{x} és \mathbf{y} vektor, amelyek kielégítik az alábbi összefüggéseket:

$$(3.4.32) \quad \mathbf{y}^\top \mathbf{x} = \mathbf{y}^\top \mathbf{N} \mathbf{x} = \mathbf{y}^\top \mathbf{N}^2 \mathbf{x} = \dots = \mathbf{y}^\top \mathbf{N}^{\alpha-2} \mathbf{x} = 0; \quad \mathbf{y}^\top \mathbf{N}^{\alpha-1} \mathbf{x} = 1.$$

Képezzük a segítségükkel a következő vektorrendszert:

$$(3.4.33) \quad \begin{matrix} \mathbf{y}^\top \mathbf{N}^{\alpha-1} \\ \mathbf{y}^\top \mathbf{N}^{\alpha-2} \\ \vdots \\ \mathbf{y}^\top \end{matrix}, \quad \mathbf{x}, \mathbf{N} \mathbf{x}, \mathbf{N}^2 \mathbf{x}, \dots, \mathbf{N}^{\alpha-1} \mathbf{x}.$$

*C. Jordan (1838–1921) francia matematikus.

Mindenekelőtt megmutatjuk, hogy ez a vektorrendszer teljes biortogonális vektorrendszer. Ha ugyanis a (3.4.33) vektorrendszer vektorait egy-egy mátrixba rendezzük és e mátrixokat összeszorozzuk, akkor egyrészt a (3.4.32) feltételek miatt, másrészt a nilpotens mátrix azon tulajdonsága miatt, hogy minden, az α -nál magasabb hatványa is zérus, a szorzat az α rendű egység-mátrixot adja:

$$\begin{aligned}
 (3.4.34) \quad & \begin{bmatrix} \mathbf{y}^T \mathbf{N}^{\alpha-1} \\ \mathbf{y}^T \mathbf{N}^{\alpha-2} \\ \vdots \\ \mathbf{y}^T \mathbf{N} \\ \mathbf{y}^T \end{bmatrix} \begin{bmatrix} \mathbf{x} & \mathbf{N}\mathbf{x} & \mathbf{N}^2\mathbf{x} \dots \mathbf{N}^{\alpha-1}\mathbf{x} \end{bmatrix} = \\
 & = \begin{bmatrix} \mathbf{y}^T \mathbf{N}^{\alpha-1}\mathbf{x} & \mathbf{y}^T \mathbf{N}^{\alpha}\mathbf{x} & \dots \mathbf{y}^T \mathbf{N}^{2\alpha-2}\mathbf{x} \\ \mathbf{y}^T \mathbf{N}^{\alpha-2}\mathbf{x} & \mathbf{y}^T \mathbf{N}^{\alpha-1}\mathbf{x} \dots \mathbf{y}^T \mathbf{N}^{2\alpha-3}\mathbf{x} \\ \dots & \dots & \dots \\ \mathbf{y}^T \mathbf{x} & \mathbf{y}^T \mathbf{N}\mathbf{x} & \dots \mathbf{y}^T \mathbf{N}^{\alpha-1}\mathbf{x} \end{bmatrix} = \mathbf{E}_{\alpha}.
 \end{aligned}$$

A (3.4.33) vektorrendszer tehát valóban teljes biortogonális vektorrendszer; vessük alá most az adott nilpotens mátrixot az ennek segítségével végzett hasonlósági transzformációnak. A beszorzást elvégezve, a (3.4.32) feltételek figyelembevételével azonnal látható, hogy a keresett Jordan-féle normálalakot kapjuk:

$$\begin{aligned}
 (3.4.35) \quad & \begin{bmatrix} \mathbf{y}^T \mathbf{N}^{\alpha-1} \\ \mathbf{y}^T \mathbf{N}^{\alpha-2} \\ \vdots \\ \mathbf{y}^T \mathbf{N} \\ \mathbf{y}^T \end{bmatrix} \begin{bmatrix} & & & & \\ & & & & \\ & & \mathbf{N} & & \\ & & & & \\ & & & & \end{bmatrix} \begin{bmatrix} \mathbf{x} & \mathbf{N}\mathbf{x} & \mathbf{N}^2\mathbf{x} \dots \mathbf{N}^{\alpha-1}\mathbf{x} \end{bmatrix} = \\
 & = \begin{bmatrix} \mathbf{y}^T \mathbf{N}^{\alpha}\mathbf{x} & \mathbf{y}^T \mathbf{N}^{\alpha+1}\mathbf{x} \dots \mathbf{y}^T \mathbf{N}^{2\alpha-1}\mathbf{x} \\ \mathbf{y}^T \mathbf{N}^{\alpha-1}\mathbf{x} & \mathbf{y}^T \mathbf{N}^{\alpha}\mathbf{x} \dots \mathbf{y}^T \mathbf{N}^{2\alpha-2}\mathbf{x} \\ \dots & \dots & \dots \\ \mathbf{y}^T \mathbf{N}\mathbf{x} & \mathbf{y}^T \mathbf{N}^2\mathbf{x} \dots \mathbf{y}^T \mathbf{N}^{\alpha}\mathbf{x} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \dots & 0 \\ 1 & 0 & 0 \dots & 0 \\ 0 & 1 & 0 \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix} = \mathbf{J}_{\alpha}.
 \end{aligned}$$

Most megmutatjuk, hogy a (3.4.32) feltételeket kielégítő \mathbf{x} és \mathbf{y} vektor valóban létezik. Ehhez először határozzunk meg olyan \mathbf{y} és \mathbf{z} vektort, amelyek kielégítik az

$$(3.4.36) \quad \mathbf{y}^T \mathbf{N}^{\alpha-1} \mathbf{z} = 1$$

feltételt. Ilyen \mathbf{y} és \mathbf{z} vektor biztosan létezik és könnyen található; ez következik a (3.4.31) feltételből, mely szerint $\mathbf{N}^{\alpha-1} \neq \mathbf{0}$, tehát van legalább egy

3.4 ... ELŐÁLLÍTÁSA A MINIMÁLPOLINOM TÖBBSZÖRÖS GYÖKEI ESETÉN 329

zérustól különböző eleme. Legyen egy ilyen eleme például az ij indexű elem: $\mathbf{e}_i^T \mathbf{N}^{\alpha-1} \mathbf{e}_j \neq 0$. Ekkor tehát pl. az $\mathbf{y} = \mathbf{e}_i$, $\mathbf{z} = \frac{1}{\mathbf{e}_i^T \mathbf{N}^{\alpha-1} \mathbf{e}_j} \mathbf{e}_j$ választással a (3.4.36) feltétel teljesül. Ennek alapján, a (3.4.32) feltételekben szereplő \mathbf{x} vektort az

$$(3.4.37) \quad \mathbf{x} = (\mathbf{E} + c_1 \mathbf{N} + c_2 \mathbf{N}^2 + \dots + c_{\alpha-1} \mathbf{N}^{\alpha-1}) \mathbf{z}$$

alakban kereshetjük. Az itt szereplő c_1, c_2, \dots együtthatók a (3.4.32) feltételekből rekurzív úton egyszerűen meghatározhatók. Ha ugyanis a (3.4.37) összefüggést balról rendre az $\mathbf{y}^T \mathbf{N}^{\alpha-1}$, $\mathbf{y}^T \mathbf{N}^{\alpha-2}$, \dots , $\mathbf{y}^T \mathbf{N}$, \mathbf{y}^T vektorokkal megszorozzuk, akkor a következő egyenleteket kapjuk:

$$\begin{aligned} \mathbf{y}^T \mathbf{N}^{\alpha-1} \mathbf{x} &= \mathbf{y}^T \mathbf{N}^{\alpha-1} \mathbf{z} = 1 \\ \mathbf{y}^T \mathbf{N}^{\alpha-2} \mathbf{x} &= \mathbf{y}^T \mathbf{N}^{\alpha-2} \mathbf{z} + \mathbf{y}^T \mathbf{N}^{\alpha-1} \mathbf{z} c_1 = 0 \\ \mathbf{y}^T \mathbf{N}^{\alpha-3} \mathbf{x} &= \mathbf{y}^T \mathbf{N}^{\alpha-3} \mathbf{z} + \mathbf{y}^T \mathbf{N}^{\alpha-2} \mathbf{z} c_1 + \mathbf{y}^T \mathbf{N}^{\alpha-1} \mathbf{z} c_2 = 0 \\ &\dots \dots \dots \\ \mathbf{y}^T \mathbf{x} &= \mathbf{y}^T \mathbf{z} + \mathbf{y}^T \mathbf{N} \mathbf{z} c_1 + \dots + \mathbf{y}^T \mathbf{N}^{\alpha-1} \mathbf{z} c_{\alpha-1} = 0. \end{aligned}$$

A (3.4.36) feltétel miatt az első összefüggés automatikusan teljesül; ezt behelyettesítve az összes többi egyenletbe, a másodikból közvetlenül adódik c_1 értéke; ezt behelyettesítve az összes többi egyenletbe, a harmadik egyenletből nyerjük c_2 értékét, és így tovább. Vagyis

$$\begin{aligned} c_1 &= -\mathbf{y}^T \mathbf{N}^{\alpha-2} \mathbf{z} \\ c_2 &= -\mathbf{y}^T \mathbf{N}^{\alpha-3} \mathbf{z} + (\mathbf{y}^T \mathbf{N}^{\alpha-2} \mathbf{z})^2 \\ c_3 &= -\mathbf{y}^T \mathbf{N}^{\alpha-4} \mathbf{z} + 2(\mathbf{y}^T \mathbf{N}^{\alpha-3} \mathbf{z})(\mathbf{y}^T \mathbf{N}^{\alpha-2} \mathbf{z}) - (\mathbf{y}^T \mathbf{N}^{\alpha-2} \mathbf{z})^3 \\ &\dots \dots \dots \end{aligned}$$

Ezzel a 3.4.6 tételt bebizonyítottuk. ■

A bizonyítás során felhasznált transzformáló mátrixok oszlopvektoraira bevezetjük a fővektorok fogalmát:

3.4.5 definíció. *Ha egy nemderogatórius nilpotens mátrix egy \mathbf{U} mátrixszal végzett hasonlósági transzformációval Jordan-féle normálalakra hozható, akkor az \mathbf{U} mátrix oszlopvektorait a nilpotens mátrix **fővektorainak** nevezzük.*

A (3.4.33) vektorrendszer alapján azt is mondhatjuk, hogy az \mathbf{x} vektor által generált fővektorlanc alkotja az \mathbf{U} transzformáló mátrix oszlopait, és – amint erről könnyen meggyőződhetünk – az $\mathbf{N}^{\alpha-1} \mathbf{x}$ vektor az adott nemderogatórius nilpotens mátrixnak egyetlen sajátvektora.

A fővektorok segítségével a 3.4.6 tétel még a következőképpen fogalmazható meg: minden nemderogatórius nilpotens transzformáció mátrixa a fővektorok által alkotott bázisban Jordan-féle normálalakú mátrix.

A 3.4.4 definícióval meghatározott Jordan-blokkot azért nevezzük *alsó* Jordan-blokknak, mivel a zérustól különböző elemek a főátló alatt helyezkednek el. Könnyen belátható, hogy a fővektorok sorrendjének megfordításával olyan új bázisra térhetünk át, amelyben az adott nilpotens transzformáció mátrixa az alsó Jordan-blokk transzponáltja – és mivel itt a zérustól különböző elemek a főátló felett helyezkednek el, ezt a mátrixot *felső Jordan-blokknak* nevezzük. Ugyanis vezessük be az alábbi, α rendű involutórius permutáló mátrixot:

$$(3.4.38) \quad \mathbf{I} = \begin{bmatrix} 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 1 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

Egyszerűen igazolható, hogy

$$(3.4.39) \quad \mathbf{I}\mathbf{J}_\alpha\mathbf{I} = \mathbf{J}_\alpha^\top, \quad \mathbf{I}^2 = \mathbf{E}.$$

Tehát ha a (3.4.35) egyenlet mindkét oldalát balról és jobbról az \mathbf{I} involutórius mátrixszal megszorozzuk, akkor innen közvetlenül kiolvasható a következő tétel.

3.4.7 tétel. *Legyen \mathbf{N} a rendezett szám- α -sok terében értelmezett α rendű, nemderogatórius nilpotens transzformáció mátrixa valamilyen adott bázisban. Ha új bázisként az*

$$(3.4.40) \quad \mathbf{x}, \mathbf{N}\mathbf{x}, \mathbf{N}^2\mathbf{x}, \dots, \mathbf{N}^{\alpha-1}\mathbf{x}$$

fővektorokat választjuk, akkor a transzformáció mátrixa alsó Jordan-féle normálalakú lesz, ha pedig a fővektorokat az

$$(3.4.41) \quad \mathbf{N}^{\alpha-1}\mathbf{x}, \mathbf{N}^{\alpha-2}\mathbf{x}, \dots, \mathbf{N}^2\mathbf{x}, \mathbf{N}\mathbf{x}, \mathbf{x}$$

sorrendben választjuk a bázis vektorainak, akkor a transzformáció mátrixa felső Jordan-féle normálalakú mátrix lesz.

A (3.4.34) összefüggés szerint a (3.4.40) alatti fővektorokkal biortogonális vektorrendszert képeznek az alábbi sorvektorok:

$$(3.4.42) \quad \begin{array}{c} \mathbf{y}^\top \mathbf{N}^{\alpha-1} \\ \mathbf{y}^\top \mathbf{N}^{\alpha-2} \\ \vdots \\ \mathbf{y}^\top \mathbf{N} \\ \mathbf{y}^\top \end{array}$$

3.4 ...ELŐÁLLÍTÁSA A MINIMÁLPOLINOM TÖBBSZÖRÖS GYÖKEI ESETÉN 331

ebből következik, hogy a (3.4.41) alatti fővektorokkal az

$$(3.4.43) \quad \begin{pmatrix} \mathbf{y}^\top \\ \mathbf{y}^\top \mathbf{N} \\ \mathbf{y}^\top \mathbf{N}^2 \\ \vdots \\ \mathbf{y}^\top \mathbf{N}^{\alpha-1} \end{pmatrix}$$

sorvektorok alkotnak biortogonális vektorrendszert. Tehát a felső Jordan-féle normálalakhoz az alábbi transzformációval jutunk:

$$\begin{aligned}
 (3.4.44) \quad & \left[\begin{array}{c} \mathbf{y}^\top \\ \mathbf{y}^\top \mathbf{N} \\ \mathbf{y}^\top \mathbf{N}^2 \\ \vdots \\ \mathbf{y}^\top \mathbf{N}^{\alpha-1} \end{array} \right] \left[\begin{array}{c} \\ \\ \mathbf{N} \\ \\ \end{array} \right] \left[\begin{array}{cccc} \mathbf{N}^{\alpha-1} \mathbf{x} & \mathbf{N}^{\alpha-2} \mathbf{x} \dots \mathbf{N} \mathbf{x} & \mathbf{x} & \end{array} \right] = \\
 &= \left[\begin{array}{cccc} 0 & 1 & 0 \dots 0 \\ 0 & 0 & 1 \dots 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 \dots 1 \\ 0 & 0 & 0 \dots 0 \end{array} \right] = \mathbf{J}_\alpha^\top.
 \end{aligned}$$

(b) Most térjünk rá a *derogatórius* nilpotens mátrixok Jordan-féle normálalakra transzformálásának vizsgálatára. Legyen az α rendű nilpotens mátrix indexe $\beta < \alpha$. A következő tételben megmutatjuk, hogy a derogatórius mátrixok is Jordan-féle normálalakra transzformálhatók.

Megjegyzés. Felhívjuk az Olvasó figyelmét arra, hogy a (3.4.38) alatti **I** involutórius mátrixszal végzett **IAI** hasonlósági transzformáció az **A** mátrix elemeit annak „középpontjára” tükrözi, vagyis az ij indexű elemét az $(n+1-i, n+1-j)$ indexű elemével felcseréli. Tehát a \mathbf{J}_α -mátrixot is a „középpontjára” tükrözi. Ez ugyan alakilag megegyezik \mathbf{J}_α transzponáltjával, de nem szabad arra a nyilvánvalóan téves következtetésre jutni, mintha a fenti involutórius transzformációval általában egy transzponálást lehetne végrehajtani.

3.4.8 tétel. Minden derogatórius nilpotens mátrix alkalmasan választott hasonlósági transzformáció segítségével Jordan-féle normálalakra hozható.

Bizonyítás. Mivel az \mathbf{N} nilpotens mátrix indexe most $\beta < \alpha$, ezért $\mathbf{N}^{\beta-1} \neq \mathbf{0}$, de $\mathbf{N}^\beta = \mathbf{0}$. A 3.4.6 tétel bizonyítása során megismert gondolatmenetet követve ismét olyan \mathbf{x} és \mathbf{y}^\top vektort választunk, amelyek kielégítik az alábbi feltételeket:

$$(3.4.45) \quad \mathbf{y}^\top \mathbf{x} = \mathbf{y}^\top \mathbf{N} \mathbf{x} = \mathbf{y}^\top \mathbf{N}^2 \mathbf{x} = \cdots = \mathbf{y}^\top \mathbf{N}^{\beta-2} \mathbf{x} = 0, \quad \mathbf{y}^\top \mathbf{N}^{\beta-1} \mathbf{x} = 1$$

(lásd (3.4.32)). Ezek segítségével elkészítjük a következő vektorrendszert:

$$(3.4.46) \quad \begin{array}{c} \mathbf{y}^\top \mathbf{N}^{\beta-1} \\ \mathbf{y}^\top \mathbf{N}^{\beta-2} \\ \vdots \\ \mathbf{y}^\top \end{array} \quad \text{és} \quad \mathbf{x}, \mathbf{N}\mathbf{x}, \mathbf{N}^2\mathbf{x}, \dots, \mathbf{N}^{\beta-1}\mathbf{x}.$$

Így ismét biortogonális vektorrendszert kapunk, mivel azonban β kisebb a vektorok dimenziójánál, a vektorrendszer nem teljes. Annak érdekében, hogy ismét hasonlósági transzformációt alkalmazhassunk, a (3.4.46) nemteljes biortogonális vektorrendszert teljessé kell tennünk. Vezessük be a (3.4.46) rendszer vektoraiból alkotott mátrixokat;

$$(3.4.47) \quad \mathbf{V}_\beta^\top = \begin{bmatrix} \mathbf{y}^\top \mathbf{N}^{\beta-1} \\ \mathbf{y}^\top \mathbf{N}^{\beta-2} \\ \vdots \\ \mathbf{y}^\top \end{bmatrix}, \quad \mathbf{U}_\beta = [\mathbf{x} \quad \mathbf{N}\mathbf{x} \quad \mathbf{N}^2\mathbf{x} \dots \mathbf{N}^{\beta-1}\mathbf{x}].$$

A (3.4.45) feltételekből következik, hogy

$$(3.4.48) \quad \mathbf{V}_\beta^\top \mathbf{U}_\beta = \mathbf{E}_\beta.$$

A 3.1.3 tétel következményeként megmutattuk, hogy nemteljes biortogonális vektorrendszert a következőképpen egészíthetünk ki teljessé: a (3.4.48) definiáló összefüggés bal oldalán álló szorzat tényezőit fordított sorrendben összeszorozva, β rangú projektort kapunk. Ezután vesszük a komplementer projektor tetszőleges minimális diadikus felbontását, és az ebben található oszlop-, ill. sorvektorokkal kiegészítve az adott vektorrendszert, azt teljessé tesszük. Esetünkben legyen ez a felbontás

$$(3.4.49) \quad \begin{aligned} & [\mathbf{x} \quad \mathbf{N}\mathbf{x} \dots \mathbf{N}^{\beta-1}\mathbf{x}] \begin{bmatrix} \mathbf{y}^\top \mathbf{N}^{\beta-1} \\ \mathbf{y}^\top \mathbf{N}^{\beta-2} \\ \vdots \\ \mathbf{y}^\top \end{bmatrix} = \\ & = \mathbf{E} - \mathbf{x}\mathbf{y}^\top \mathbf{N}^{\beta-1} - \mathbf{N}\mathbf{x}\mathbf{y}^\top \mathbf{N}^{\beta-2} - \dots - \mathbf{N}^{\beta-1}\mathbf{x}\mathbf{y}^\top = \sum_{k=1}^{\alpha-\beta} \mathbf{w}_k \mathbf{z}_k^\top. \end{aligned}$$

Az így nyert $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{\alpha-\beta}$, ill. $\mathbf{z}_1^\top, \mathbf{z}_2^\top, \dots, \mathbf{z}_{\alpha-\beta}^\top$ vektorokkal kiegészítve a (3.4.46) fővektorokat, a kapott teljes biortogonális vektorrendszer segítségével végezzünk hasonlósági transzformációt az adott \mathbf{N} nilpotens mátrixon:

$$(3.4.50) \quad \begin{bmatrix} \mathbf{y}^\top \mathbf{N}^{\beta-1} \\ \mathbf{y}^\top \mathbf{N}^{\beta-2} \\ \vdots \\ \mathbf{y}^\top \\ \hline \mathbf{z}_1^\top \\ \mathbf{z}_2^\top \\ \vdots \\ \mathbf{z}_{\alpha-\beta}^\top \end{bmatrix} \begin{bmatrix} \mathbf{x} & \mathbf{N}\mathbf{x} & \mathbf{N}^2\mathbf{x} \dots \mathbf{N}^{\beta-1}\mathbf{x} | \mathbf{w}_1 \dots \mathbf{w}_{\alpha-\beta} \end{bmatrix} = \begin{bmatrix} \mathbf{J}_\beta & \mathbf{0} \\ \hline \mathbf{0} & \tilde{\mathbf{N}}_{\alpha-\beta} \end{bmatrix}.$$

A transzformációval olyan kvázidiagonál-mátrixot kapunk, amelynek a főátlójában egy β rendű Jordan-blokk és egy $(\alpha - \beta)$ rendű nilpotens mátrix áll. (A zérus elemek a biortogonalitási feltételekből adódnak, ill. annak következményei, hogy a nilpotens mátrix β -nál magasabb hatványai zérussal egyenlők.) Ha ezek után ugyanezt az eljárást megismételjük az $(\alpha - \beta)$ rendű $\tilde{\mathbf{N}}_{\alpha-\beta}$ nilpotens mátrixszal, akkor látható, hogy véges számú lépés után – vagyis véges sok hasonlósági transzformáció alkalmazásával – elérhető, hogy a főátlóban csupa Jordan-blokk álljon. Megjegyezzük még, hogy az egymás után következő Jordan-blokkok rendszáma (más szóval az egymás után redukálható nilpotens mátrixok indexe) monoton nemnövekvő sorozatot alkot. Ha ugyanis pl. $\tilde{\mathbf{N}}_{\alpha-\beta}$ indexe $\gamma > \beta$ volna, akkor az eredetileg adott \mathbf{N} nilpotens mátrix indexe is γ volna, amiből az következne, hogy az első lépésben éppen egy γ rendű Jordan-blokkot választottunk volna le. Ebből a megfontolásból egyébként az is látszik, hogy két, egymás után következő Jordan-blokk rendszáma megegyezhet. ■

18. Példa. Transzformáljuk az

$$\mathbf{N} = \begin{bmatrix} 2 & -1 & 1 & -1 \\ -3 & 4 & -5 & 4 \\ 8 & -4 & 4 & -4 \\ 15 & -10 & 11 & -10 \end{bmatrix}$$

negyedrendű mátrixot Jordan-féle normálalakra.

Megoldás. Hatványozással meggyőződhetünk arról, hogy

$$\mathbf{N}^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 2 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ -2 & 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 1 & -1 \end{bmatrix},$$

$$\mathbf{N}^3 = \mathbf{0},$$

tehát az adott mátrix derogatórius nilpotens mátrix, amelynek rendje 4, indexe pedig $\beta = 3$. A fővektorok megkonstruálása céljából induljunk ki az $\mathbf{N}^2 \neq \mathbf{0}$ mátrix második sorának zérustól különböző harmadik eleméből, vagyis legyen $\mathbf{y}^\top = \mathbf{e}_2^\top$ és $\mathbf{z} = \mathbf{e}_3$, mivel így $\mathbf{y}^\top \mathbf{N}^2 \mathbf{z} = \mathbf{e}_2^\top \mathbf{N}^2 \mathbf{e}_3 = 1 \neq 0$. Az \mathbf{x} vektort az alábbi alakban keressük:

$$\mathbf{x} = (\mathbf{E} + c_1 \mathbf{N} + c_2 \mathbf{N}^2) \mathbf{z}.$$

A 3.4.6 tétel bizonyításában fellépő összefüggések felhasználásával a c_1, c_2 együtthatókra adódó egyenletek

$$c_1 = -\mathbf{y}^\top \mathbf{N} \mathbf{z} = -\mathbf{e}_2^\top \mathbf{N} \mathbf{e}_3 = 5; \quad c_2 = -\mathbf{y}^\top \mathbf{z} + (\mathbf{y}^\top \mathbf{N} \mathbf{z})^2 = 25,$$

$$\text{vagyis } \mathbf{x} = (\mathbf{E} + 5\mathbf{N} + 25\mathbf{N}^2) \mathbf{e}_3 = \begin{bmatrix} 5 \\ 0 \\ 21 \\ 30 \end{bmatrix}. \text{ Ezzel a fővektorok nemteljes}$$

biortogonális rendszere:

$$\begin{bmatrix} \mathbf{y}^\top \mathbf{N}^2 \\ \mathbf{y}^\top \mathbf{N} \\ \mathbf{y}^\top \end{bmatrix} = \begin{bmatrix} 2 & -1 & 1 & -1 \\ -3 & 4 & -5 & 4 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

$$[\mathbf{x} \quad \mathbf{N}\mathbf{x} \quad \mathbf{N}^2\mathbf{x}] = \begin{bmatrix} 5 & 1 & 0 \\ 0 & 0 & 1 \\ 21 & 4 & 0 \\ 30 & 6 & -1 \end{bmatrix}.$$

Ennek a biortogonális rendszernek teljessé tétele érdekében képezzük a (3.4.49) komplementer projektor diadikus felbontását:

$$\mathbf{E} - \begin{bmatrix} \mathbf{x} & \mathbf{N}\mathbf{x} & \mathbf{N}^2\mathbf{x} \end{bmatrix} \begin{bmatrix} \mathbf{y}^\top \mathbf{N}^2 \\ \mathbf{y}^\top \mathbf{N} \\ \mathbf{y}^\top \end{bmatrix} = \begin{bmatrix} -6 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ -30 & 5 & 0 & 5 \\ -42 & 7 & 0 & 7 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 5 \\ 7 \end{bmatrix} \begin{bmatrix} -6 & 1 & 0 & 1 \end{bmatrix}.$$

Az így nyert oszlop-, ill. sorvektorral kiegészítve a fenti nemteljes biortogonális vektorrendszert, hajtsuk végre az adott mátrix hasonlósági transzformációját:

$$\underbrace{\begin{bmatrix} 2 & -1 & 1 & -1 \\ -3 & 4 & -5 & 4 \\ 0 & 1 & 0 & 0 \\ -6 & 1 & 0 & 1 \end{bmatrix}}_{\mathbf{V}^\top} \underbrace{\begin{bmatrix} 2 & -1 & 1 & -1 \\ -3 & 4 & -5 & 4 \\ 8 & -4 & 4 & -4 \\ 15 & -10 & 11 & -10 \end{bmatrix}}_{\mathbf{N}} \underbrace{\begin{bmatrix} 5 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 21 & 4 & 0 & 5 \\ 30 & 6 & -1 & 7 \end{bmatrix}}_{\mathbf{U}} =$$

3.4 ... ELŐÁLLÍTÁSA A MINIMÁLPOLINOM TÖBBSZÖRÖS GYÖKEI ESETÉN 335

$$= \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] = \left[\begin{array}{ccc|ccc} & & & & & \\ & \mathbf{J}_3 & & & \mathbf{0} & \\ & & & & & \\ \hline & & & \mathbf{0} & & \tilde{\mathbf{N}} \end{array} \right].$$

Látható, hogy a fenti transzformációval az adott $\beta = 3$ indexű derogatórius nilpotens mátrixot egy hamadrendű Jordan-blokkra és egy $\alpha - \beta = 1$ rendű nilpotens blokkra redukáltuk, mivel azonban az egyetlen „elsőrendű” nilpotens mátrix a zérus elem, a redukciót ezzel teljesen végrehajtottuk.

* * *

19. Példa. Hozzuk Jordan-féle normálalakra az

$$\mathbf{A} = \begin{bmatrix} 3 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 0 & 5 & -3 \\ 4 & -1 & 3 & -1 \end{bmatrix}$$

mátrixot.

Megoldás. Mivel a mátrix karakterisztikus egyenlete

$$D(\lambda) \equiv |\lambda \mathbf{E} - \mathbf{A}| = (\lambda - 2)^4 = 0,$$

a mátrix egyetlen négyszeres sajátértéke 2.

Írjuk fel az \mathbf{A} mátrixot a (3.4.21) képlet segítségével. Mivel a minimál-egyenletnek egyetlen gyöke van, ezért $H_{10}(\mathbf{A}) = \mathbf{E}$, és így

$$\mathbf{A} = 2\mathbf{E} + (\mathbf{A} - 2\mathbf{E}) = \begin{bmatrix} 2 & & & \\ & 2 & & \\ & & 2 & \\ & & & 2 \end{bmatrix} + \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 3 & 0 & 3 & -3 \\ 4 & -1 & 3 & -3 \end{bmatrix}.$$

Jelölje \mathbf{N} az $\mathbf{A} - 2\mathbf{E}$ nilpotens mátrixot. Mivel $\mathbf{N}^2 = \mathbf{0}$, ezért \mathbf{N} egy 2 indexű, negyedrendű nilpotens mátrix. A fővektorok meghatározása céljából induljunk ki az $\mathbf{N} \neq \mathbf{0}$ mátrixból. Ennek bal felső sarokeleme zérustól különbözik, ezért választható

$$\mathbf{y}^\top = \mathbf{e}_1^\top \text{ és } \mathbf{z} = \mathbf{e}_1 \text{ és így } \mathbf{y}^\top \mathbf{N} \mathbf{z} = \mathbf{e}_1^\top \mathbf{N} \mathbf{e}_1 = 1.$$

Keressük az \mathbf{x} vektort a következő alakban: $\mathbf{x} = (\mathbf{E} + c_1 \mathbf{N}) \mathbf{z}$. Mivel $\mathbf{y}^\top \mathbf{x} = 0 = \mathbf{y}^\top \mathbf{z} + c_1 \mathbf{y}^\top \mathbf{N} \mathbf{z}$, innen $c_1 = -1$, és így

$$\mathbf{x} = (\mathbf{E} - \mathbf{N}) \mathbf{e}_1 = \begin{bmatrix} 0 \\ -1 \\ -3 \\ -4 \end{bmatrix}.$$

A fővektorok nemteljes biortogonális rendszere:

$$[\mathbf{x} \quad \mathbf{Nx}] = \begin{bmatrix} 0 & 1 \\ -1 & 1 \\ -3 & 3 \\ -4 & 4 \end{bmatrix}, \quad \begin{bmatrix} \mathbf{y}^\top \mathbf{N} \\ \mathbf{y}^\top \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

A biortogonális rendszert teljessé tesszük az $\mathbf{E} - [\mathbf{x} \quad \mathbf{Nx}] \begin{bmatrix} \mathbf{y}^\top \mathbf{N} \\ \mathbf{y}^\top \end{bmatrix}$ projektor minimális diadikus felbontásával:

$$\begin{aligned} \mathbf{E} - [\mathbf{x} \quad \mathbf{Nx}] \begin{bmatrix} \mathbf{y}^\top \mathbf{N} \\ \mathbf{y}^\top \end{bmatrix} &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -3 & 1 & 0 \\ 0 & -4 & 0 & 1 \end{bmatrix} = \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -4 & 0 & 1 \\ 0 & -3 & 1 & 0 \end{bmatrix} = [\mathbf{w}_1 \quad \mathbf{w}_2] \begin{bmatrix} \mathbf{z}_1^\top \\ \mathbf{z}_2^\top \end{bmatrix}. \end{aligned}$$

A teljes biortogonális vektorrendszerrel hajtsuk végre az \mathbf{N} mátrix hasonlósági transzformációját:

$$\begin{aligned} \begin{bmatrix} \mathbf{y}^\top \mathbf{N} \\ \mathbf{y}^\top \\ \mathbf{z}_1^\top \\ \mathbf{z}_2^\top \end{bmatrix} \mathbf{N} [\mathbf{x} \quad \mathbf{Nx} \quad \mathbf{w}_1 \quad \mathbf{w}_2] &= \\ &= \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -4 & 0 & 1 \\ 0 & -3 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 3 & 0 & 3 & -3 \\ 4 & -1 & 3 & -3 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -3 & 3 & 0 & 1 \\ -4 & 4 & 1 & 0 \end{bmatrix} = \\ &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -3 & 3 \\ 0 & 0 & -3 & 3 \end{bmatrix} = \begin{bmatrix} \mathbf{J}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{N}_1 \end{bmatrix}, \end{aligned}$$

ahol $\mathbf{N}_1 = \begin{bmatrix} -3 & 3 \\ -3 & 3 \end{bmatrix}$. Mivel $\mathbf{N}_1^2 = \mathbf{0}$, ezért \mathbf{N}_1 másodrendű, nemderogatórius nilpotens mátrix. A fővektorok meghatározása céljából legyen most

$$\mathbf{y}^\top = \mathbf{e}_1^\top \quad \text{és} \quad \mathbf{z} = \frac{1}{3} \mathbf{e}_2 \quad \text{és így} \quad \mathbf{y}^\top \mathbf{N}_1 \mathbf{z} = 1.$$

3.4 ...ELŐÁLLÍTÁSA A MINIMÁLPOLINOM TÖBBSZÖRÖS GYÖKEI ESETÉN 337

Az $\mathbf{x} = (\mathbf{E} + c_1 \mathbf{N}_1) \mathbf{z}$ összefüggésben szereplő c_1 együtthatóra $c_1 = 0$ adódik, tehát $\mathbf{x} = \frac{1}{3} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. A fővektorok biortogonális rendszere

$$[\mathbf{x} \quad \mathbf{N}_1 \mathbf{x}] = \begin{bmatrix} 0 & 1 \\ \frac{1}{3} & 1 \end{bmatrix}; \quad \begin{bmatrix} \mathbf{y}^\top \mathbf{N}_1 \\ \mathbf{y}^\top \end{bmatrix} = \begin{bmatrix} -3 & 3 \\ 1 & 0 \end{bmatrix},$$

és innen megkapjuk \mathbf{N}_1 Jordan-féle normálalakját:

$$\begin{bmatrix} \mathbf{y}^\top \mathbf{N}_1 \\ \mathbf{y}^\top \end{bmatrix} \mathbf{N}_1 [\mathbf{x} \quad \mathbf{N}_1 \mathbf{x}] = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Ezt felhasználva,

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -3 & 3 \\ 0 & 0 & 1 & 0 \end{bmatrix} \left[\begin{array}{c|c} \mathbf{J}_1 & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{N}_1 \end{array} \right] \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{1}{3} & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

adódik. A kapott mátrix az $\mathbf{N} = \mathbf{A} - 2\mathbf{E}$ mátrix Jordan-féle normálalakja. A transzformáció mátrixát megadja a két fenti transzformációs mátrix szorzata:

$$\mathbf{V}^\top = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -3 & 3 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -4 & 0 & 1 \\ 0 & -3 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 3 & 3 & -3 \\ 0 & -4 & 0 & 1 \end{bmatrix},$$

$$\mathbf{U} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -3 & 3 & 0 & 1 \\ -4 & 4 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{1}{3} & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -3 & 3 & \frac{1}{3} & 1 \\ -4 & 4 & 0 & 1 \end{bmatrix},$$

ahonnan

$$\mathbf{N} = \mathbf{U} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{V}^\top, \quad \text{illetve} \quad \mathbf{A} = \mathbf{U} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ \hline 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix} \mathbf{V}^\top.$$

* * *

20. Példa. Határozzuk meg az

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

mátrix Jordan-féle normálalakját!

Megoldás. Mivel a mátrix karakterisztikus egyenlete

$$D(\lambda) \equiv |\lambda \mathbf{E} - \mathbf{A}| = (\lambda - 1)^6 = 0,$$

ezért $\lambda = 1$ hatszoros sajátérték. Az \mathbf{A} mátrix felbontása az Hermite-féle mátrixpolinomokkal: $\mathbf{A} = \mathbf{E} + (\mathbf{A} - \mathbf{E})$. Itt $\mathbf{A} - \mathbf{E} = \mathbf{N}$ hatodrendű, 3 indexű nilpotens mátrix, ahol

$$\mathbf{N} = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{N}^2 = \begin{bmatrix} 0 & 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

A fővektorok meghatározásához induljunk ki az $\mathbf{N}^2 \neq \mathbf{0}$ mátrix első sorának negyedik eleméből:

$$\mathbf{y}^\top = \mathbf{e}_1^\top, \quad \mathbf{z} = \frac{1}{2}\mathbf{e}_4 \quad \text{és így} \quad \mathbf{y}^\top \mathbf{N}^2 \mathbf{z} = 1.$$

Keressük az \mathbf{x} vektort az $\mathbf{x} = (\mathbf{E} + c_1 \mathbf{N} + c_2 \mathbf{N}^2)\mathbf{z}$ alakban és a c_1, c_2 együtt-hatók értékeit az $\mathbf{y}^\top \mathbf{N} \mathbf{z} = 0$ és $\mathbf{y}^\top \mathbf{x} = 0$ feltételekből határozhatjuk meg. Ezekből $c_1 = -\frac{1}{2}$, $c_2 = \frac{1}{4}$ adódik, és így az \mathbf{x} vektor:

$$\mathbf{x} = \begin{bmatrix} 0 \\ \frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 0 \\ 0 \end{bmatrix}.$$

A fővektorok nemteljes biortogonális rendszere ezúttal:

$$[\mathbf{N}^2 \mathbf{x} \quad \mathbf{N} \mathbf{x} \quad \mathbf{x}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \quad \begin{bmatrix} \mathbf{y}^\top \\ \mathbf{y}^\top \mathbf{N} \\ \mathbf{y}^\top \mathbf{N}^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 & 2 & 2 \end{bmatrix}.$$

3.4 ... ELŐÁLLÍTÁSA A MINIMÁLPOLINOM TÖBBSZÖRÖS GYÖKEI ESETÉN 339

A biortogonális vektorrendszert teljessé tesszük az

$$\mathbf{E} - [\mathbf{N}^2 \mathbf{x} \quad \mathbf{N} \mathbf{x} \quad \mathbf{x}] \begin{bmatrix} \mathbf{y}^\top \\ \mathbf{y}^\top \mathbf{N} \\ \mathbf{y}^\top \mathbf{N}^2 \end{bmatrix}$$

projektor minimális diadikus felbontásával:

$$\begin{aligned} \mathbf{E} - [\mathbf{N}^2 \mathbf{x} \quad \mathbf{N} \mathbf{x} \quad \mathbf{x}] \begin{bmatrix} \mathbf{y}^\top \\ \mathbf{y}^\top \mathbf{N} \\ \mathbf{y}^\top \mathbf{N}^2 \end{bmatrix} &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \\ &= \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

A transzformációt most az

$$\mathbf{U}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -\frac{1}{2} & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{V}_1^\top = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

mátrixok segítségével hajtjuk végre, így

$$\mathbf{V}_1^\top \mathbf{N} \mathbf{U}_1 = \left[\begin{array}{ccc|ccc} 0 & 1 & 0 & & & \\ 0 & 0 & 1 & & \mathbf{0} & \\ 0 & 0 & 0 & & & \\ \hline & & & 0 & 1 & 1 \\ & \mathbf{0} & & 0 & 0 & 0 \\ & & & 0 & 0 & 0 \end{array} \right] = \left[\begin{array}{ccc|ccc} & & & \mathbf{J}_1 & & \mathbf{0} \\ \hline & & & \mathbf{0} & & \mathbf{N}_1 \end{array} \right]$$

adódik, ahol $\mathbf{N}_1 = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $\mathbf{N}_1^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ egy harmadrendű, 2 indexű nilpotens mátrix.

Fővektorainak meghatározásához legyen most $\mathbf{y}^T = \mathbf{e}_1^T$ és $\mathbf{z} = \mathbf{e}_2$, így $\mathbf{y}^T \mathbf{N}_1 \mathbf{z} = 1$. Az $\mathbf{x} = (\mathbf{E} + c_1 \mathbf{N}_1) \mathbf{z}$ egyenletből $c_1 = 0$ adódik, tehát $\mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.

A fővektorok nemteljes biortogonális rendszere:

$$[\mathbf{N}_1 \mathbf{x} \quad \mathbf{x}] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} \mathbf{y}^T \\ \mathbf{y}^T \mathbf{N}_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

Ezt tegyük teljessé az

$$\mathbf{E} - [\mathbf{N}_1 \mathbf{x} \quad \mathbf{x}] \begin{bmatrix} \mathbf{y}^T \\ \mathbf{y}^T \mathbf{N}_1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$

minimális diadikus felbontással. Ennek felhasználásával hajtsuk végre a transzformációt:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{N}_1 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

így megkaptuk az \mathbf{N}_1 mátrix Jordan-féle normálalakját. Az adott mátrixon tehát az alábbi szorzatok segítségével végezhető el a transzformáció:

$$\mathbf{V}^T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{V}_1^T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{U} = \mathbf{U}_1 \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -\frac{1}{2} & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

ahonnan

3.4 ... ELŐÁLLÍTÁSA A MINIMÁLPOLINOM TÖBBSZÖRÖS GYÖKEI ESETÉN 341

$$\mathbf{N} = \mathbf{U} \left[\begin{array}{ccc|cc} 0 & 1 & 0 & & \\ 0 & 0 & 1 & & \\ 0 & 0 & 0 & & \\ \hline & & & 0 & 1 \\ & & & 0 & 0 \\ \hline & & & & 0 \end{array} \right] \mathbf{V}^T, \text{ azaz } \mathbf{A} = \mathbf{U} \left[\begin{array}{ccc|cc} 2 & 1 & 0 & & \\ 0 & 2 & 1 & & \\ 0 & 0 & 2 & & \\ \hline & & & 2 & 1 \\ & & & 0 & 2 \\ \hline & & & & 2 \end{array} \right] \mathbf{V}^T$$

adódik. Látható, hogy ebben az esetben a mátrix Jordan-féle normálalakja három Jordan-blokból áll (három Jordan-blokk „direkt összege”), ezek rendszáma monoton nemnövekvő sorrendben 3, 2 és 1.

* * *

3.4.5 Mátrixfüggvények kanonikus előállítása

Tekintsünk egyszerűség kedvéért egy α rendű nemderogatórius \mathbf{N} nilpotens mátrixot, és állítsuk elő ennek $f(\mathbf{N})$ függvényét, ahol $f(z)$ az $f(z) = \sum_{k=0}^{\infty} c_k z^k$ hatványsorral van definiálva. Tekintettel arra, hogy a $z = 0$ helyen ez a hatványsor mindig konvergens, és hogy a 3.4.5 tétel szerint nilpotens mátrix valamennyi sajátértéke 0, a $\sum_{k=1}^{\infty} c_k \mathbf{N}^k$ hatványsor konvergenciája, vagyis az $f(\mathbf{N})$ mátrixfüggvény létezése minden esetben biztosítva van.

Az $f(\mathbf{N})$ függvény előállításához felhasználjuk a (3.4.7) összefüggést. A minimálpolinom ebben az esetben

$$(3.4.51) \quad \Delta(\lambda) \equiv \lambda^\alpha,$$

ennek egyetlen (α multiplicitású) gyöke $\lambda_k = 0$; az egyetlen $H_{k0}(\mathbf{N})$ mátrixpolinom tehát az α rendű egységmátrix, a $H_{k\nu}(\mathbf{N})$ mátrixpolinomok pedig a (3.4.17) összefüggések alapján a következő egyszerű alakban adódnak:

$$(3.4.52) \quad H_{k\nu}(\mathbf{N}) = \frac{(\mathbf{N} - \lambda_k \mathbf{E})^\nu}{\nu!} H_{k0}(\mathbf{N}) = \frac{1}{\nu!} \mathbf{N}^\nu \quad (\nu = 1, \dots, \alpha - 1).$$

Behelyettesítve ezt a (3.4.7) képletbe, a keresett függvény az alábbi alakban írható fel:

$$(3.4.53) \quad f(\mathbf{N}) = f(0)\mathbf{E} + \frac{f'(0)}{1!} \mathbf{N} + \frac{f''(0)}{2!} \mathbf{N}^2 + \dots + \frac{f^{(\alpha-1)}(0)}{(\alpha-1)!} \mathbf{N}^{\alpha-1}.$$

Így a hatványsorral definiált $f(\mathbf{N})$ mátrixfüggvényt sikerült egy $(\alpha-1)$ -edfokú polinomra redukálnunk. Vegyük most figyelembe, hogy a 3.4.6 tétel alapján,

minden nemderogatórius α rendű nilpotens mátrix a fővektorok segítségével α rendű felső (vagy alsó) Jordan-blokkra transzformálható. Ebből következik, hogy a nilpotens mátrix bármely hatványa – tehát polinomja is – ugyanezzel a transzformációval a felső (vagy alsó) Jordan-blokk megfelelő hatványára, ill. polinomjára transzformálható. Legyen \mathbf{T} az adott nilpotens mátrix fővektoraiból alkotott mátrix: mivel $\mathbf{N} = \mathbf{T}\mathbf{J}\mathbf{T}^{-1}$, tehát (3.4.53) alapján

$$(3.4.54) \quad f(\mathbf{N}) = \mathbf{T} \left\{ f(0)\mathbf{E} + \frac{f'(0)}{1!}\mathbf{J} + \frac{f''(0)}{2!}\mathbf{J}^2 + \dots + \frac{f^{(\alpha-1)}(0)}{(\alpha-1)!}\mathbf{J}^{\alpha-1} \right\} \mathbf{T}^{-1}.$$

A kapcsos zárójelben álló kifejezés a Jordan-blokk egy polinomja (csupán érdekesség kedvéért jegyezzük meg, hogy ennek a polinomnak a szerkezete emlékeztet a Taylor-polinom szerkezetére), ilyen polinomokkal pedig találkozunk már az első fejezet példáiban (lásd (1.4.7)). Felhasználva az ott kapott eredményeket, az $f(\mathbf{N})$ függvényt most már a fővektorok segítségével az alábbi, *kanonikus alakban* írhatjuk fel:

$$(3.4.55) \quad f(\mathbf{N}) = \mathbf{T}f(\mathbf{J})\mathbf{T}^{-1},$$

ahol $f(\mathbf{J})$ a következő:

$$(3.4.56) \quad f(\mathbf{J}) = \begin{bmatrix} f(0) & f'(0) & \frac{1}{2!}f''(0) & \frac{1}{3!}f'''(0) \dots \frac{1}{(\alpha-1)!}f^{(\alpha-1)}(0) \\ 0 & f(0) & f'(0) & \frac{1}{2!}f''(0) \dots \frac{1}{(\alpha-2)!}f^{(\alpha-2)}(0) \\ 0 & 0 & . & . & \dots & . & . \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & . & . & \dots & f(0) & f'(0) \\ 0 & 0 & . & . & \dots & 0 & f(0) \end{bmatrix},$$

és \mathbf{T} oszlopai az adott \mathbf{N} nilpotens mátrix fővektorai. Ha az adott \mathbf{N} nilpotens mátrix derogatórius, akkor elvileg ugyanúgy járunk el, mint a nemderogatórius esetben. Ekkor azonban figyelembe kell venni, hogy a mátrix normálatalkja – amire véges sok egymás után alkalmazott hasonlósági transzformációval jutunk – Jordan-blokkokból felépített kvázidiagonál-mátrix. A (3.4.55) előállításban szereplő $f(\mathbf{J})$ polinom tehát nem egyetlen α rendű háromszögmátrix lesz, hanem az egyes Jordan-blokkok rendszámával megegyező rendszámú háromszögmátrixokból felépített kvázidiagonálmátrix.

Megjegyzés. A (3.4.56) alakú $f(\mathbf{J})$ háromszögmátrixok újabb hasonlósági transzformációval természetesen ismét Jordan-féle normálatalkra hozhatók és így megkaphatjuk a nilpotens mátrix függvényének Jordan-féle normálatalkját. Erre itt nem térünk ki.

Térjünk most vissza a legáltalánosabb típusú \mathbf{A} mátrixok $f(\mathbf{A})$ függvényének (3.4.7) alakjára. Feltesszük, hogy az ott megadott konvergenciafeltételek

3.4 ... ELŐÁLLÍTÁSA A MINIMÁLPOLINOM TÖBBSZÖRÖS GYÖKEI ESETÉN 343

teljesülnek, vagyis hogy az \mathbf{A} mátrix sajátértékei az $f(z)$ függvény konvergenciakörének belsejébe esnek. Fejezzük ki a (3.4.7)-ben szereplő $H_{k\nu}(\mathbf{A})$ Hermite-féle mátrixpolinomokat minden $\nu \geq 1$ -re a (3.4.17) képlet alapján a $H_{k0}(\mathbf{A})$ polinomok segítségével, és helyettesítsük be a (3.4.7) egyenletbe:

$$(3.4.57) \quad f(\mathbf{A}) = \sum_{k=1}^s \left\{ f(\lambda_k) \mathbf{E} + f'(\lambda_k)(\mathbf{A} - \lambda_k \mathbf{E}) + \dots + \frac{f^{(\gamma_k-1)}(\lambda_k)}{(\gamma_k-1)!} (\mathbf{A} - \lambda_k \mathbf{E})^{\gamma_k-1} \right\} H_{k0}(\mathbf{A}).$$

Megjegyezzük, hogy ebből a kifejezésből $f(z) \equiv z$ esetén közvetlenül adódik az \mathbf{A} mátrix (3.4.21) alatti felbontása. Ahhoz hasonlóan, ahogy ott a (3.4.22) majd a (3.4.26) összefüggések levezetése során tettük, szorozzunk be most is a $H_{k0}(\mathbf{A})$ projektorokkal és vegyük figyelembe, hogy ezek idempotens mátrixok. Így módon a következő kifejezésre jutunk:

$$f(\mathbf{A}) = \sum_{k=1}^s \left\{ f(\lambda_k) H_{k0}(\mathbf{A}) + H_{k0}(\mathbf{A}) f'(\lambda_k)(\mathbf{A} - \lambda_k \mathbf{E}) H_{k0}(\mathbf{A}) + H_{k0}(\mathbf{A}) \frac{f''(\lambda_k)}{2!} (\mathbf{A} - \lambda_k \mathbf{E})^2 H_{k0}^2(\mathbf{A}) + \dots + H_{k0}(\mathbf{A}) \frac{f^{(\gamma_k-1)}(\lambda_k)}{(\gamma_k-1)!} (\mathbf{A} - \lambda_k \mathbf{E})^{\gamma_k-1} H_{k0}^{\gamma_k-1}(\mathbf{A}) \right\}.$$

Helyettesítsük be a $H_{k0}(\mathbf{A})$ projektorok helyére azok (3.4.23) alatti diadikus felbontását, és emeljük ki balra az \mathbf{U}_k , jobbra pedig a \mathbf{V}_k^\top tényezőket:

$$(3.4.58) \quad f(\mathbf{A}) = \sum_{k=1}^s \mathbf{U}_k \left\{ f(\lambda_k) \mathbf{E}_{\alpha_k} + f'(\lambda_k) \mathbf{V}_k^\top (\mathbf{A} - \lambda_k \mathbf{E}) \mathbf{U}_k + \frac{1}{2!} f''(\lambda_k) \mathbf{V}_k^\top (\mathbf{A} - \lambda_k \mathbf{E})^2 H_{k0}(\mathbf{A}) \mathbf{U}_k + \dots + \frac{f^{(\gamma_k-1)}(\lambda_k)}{(\gamma_k-1)!} \mathbf{V}_k^\top (\mathbf{A} - \lambda_k \mathbf{E})^{\gamma_k-1} H_{k0}^{\gamma_k-2}(\mathbf{A}) \mathbf{U}_k \right\} \mathbf{V}_k^\top = \sum_{k=1}^s \mathbf{U}_k \mathbf{Q}_k \mathbf{V}_k^\top.$$

A 3.4.2 tétel következtében az összes \mathbf{U}_k oszlopvektoraiból alkotott \mathbf{U} mátrixszal végzett hasonlósági transzformáció az $f(\mathbf{A})$ mátrixfüggvényt olyan kvázidiagonálmátrixra transzformálja, amelynek főátlójában a (3.4.58)-ban szereplő \mathbf{Q}_k blokkok állnak. Nem nehéz belátni, hogy ezek a blokkok az

$$(3.4.59) \quad \mathbf{N}_k = \mathbf{V}_k^\top (\mathbf{A} - \lambda_k \mathbf{E}) \mathbf{U}_k$$

nilpotens mátrix polinomjai. Ennek hatványai ugyanis:

$$\mathbf{N}_k^2 = \mathbf{V}_k^\top (\mathbf{A} - \lambda_k \mathbf{E}) \mathbf{U}_k \mathbf{V}_k^\top (\mathbf{A} - \lambda_k \mathbf{E}) \mathbf{U}_k = \mathbf{V}_k^\top (\mathbf{A} - \lambda_k \mathbf{E})^2 H_{k0}(\mathbf{A}) \mathbf{U}_k;$$

hasonlóan adódik $\mathbf{N}_k^3 = \mathbf{V}_k^\top (\mathbf{A} - \lambda_k \mathbf{E})^3 H_{k0}^2(\mathbf{A}) \mathbf{U}_k$ stb., és (3.4.12) miatt $\mathbf{N}_k^{\gamma_k} = \mathbf{0}$. Ha $\gamma_k = \alpha_k$, akkor \mathbf{N}_k nemderogatórius, ha $\gamma_k < \alpha_k$, akkor \mathbf{N}_k derogatórius nilpotens mátrix. A 3.4.8 tétel értelmében tehát található olyan \mathbf{T}_k mátrix, amellyel végzett hasonlósági transzformációval az \mathbf{N}_k nilpotens mátrix Jordan-féle normálalakra hozható:

$$(3.4.60) \quad \mathbf{N}_k = \mathbf{V}_k^\top (\mathbf{A} - \lambda_k \mathbf{E}) \mathbf{U}_k = \mathbf{T}_k \mathbf{J}_k \mathbf{T}_k^{-1}.$$

Ezzel a \mathbf{T}_k mátrixszal a (3.4.58) összefüggésben szereplő \mathbf{Q}_k blokkok a következőképpen transzformálhatók:

$$(3.4.61) \quad \mathbf{Q}_k = \mathbf{T}_k \sum_{\nu=0}^{\gamma_k-1} \left(\frac{f^{(\nu)}(\lambda_k)}{\nu!} \mathbf{J}_k^\nu \right) \mathbf{T}_k^{-1}.$$

Jelölje \mathbf{T} az α_k rendű \mathbf{T}_k mátrixokból alkotott kvázidiagonál-mátrixot. Behelyettesítve a (3.4.58) összefüggésbe, megkapjuk az $f(\mathbf{A})$ mátrixfüggvény kanonikus előállítását:

$$(3.4.62) \quad f(\mathbf{A}) = \mathbf{U} \mathbf{T} \left\langle \sum_{\nu=0}^{\gamma_k-1} \frac{f^{(\nu)}(\lambda_k)}{\nu!} \mathbf{J}_k^\nu \right\rangle \mathbf{T}^{-1} \mathbf{U}^{-1}.$$

Az $\mathbf{U} \mathbf{T}$ mátrixszorzat oszlopvektorai szolgáltatják tehát azt a bázist, amelyben az adott transzformáció $f(\mathbf{A})$ függvényének mátrixa Jordan-blokkok (3.4.56) alakú függvényeiből mint háromszögmátrixokból alkotott kvázidiagonál-mátrix.

Arra a kérdésre, hogy adott λ_k sajátértékhez tartozó nilpotens mátrixpolinom hány Jordan-blokból álló kvázidiagonál-mátrixként állítható elő, az elemi osztók elmélete ad választ. Ezzel majd a következő szakaszban foglalkozunk.

21. Példa. Legyen

$$\mathbf{A} = \begin{bmatrix} 3 & -1 & -4 & 2 \\ 2 & 3 & -2 & -4 \\ 2 & -1 & -3 & 2 \\ 1 & 2 & -1 & -3 \end{bmatrix}.$$

Határozzuk meg $\sin \mathbf{A}$ kanonikus előállítását!

Megoldás. Egyszerű számolás után meggyőződhetünk arról, hogy \mathbf{A} karakterisztikus egyenlete megegyezik minimálegyenletével, mégpedig

$$D(\lambda) \equiv |\lambda \mathbf{E} - \mathbf{A}| \equiv \Delta(\lambda) \equiv (\lambda - 1)^2 (\lambda + 1)^2 = 0.$$

Vagyis a sajátértékek $\lambda_1 = 1$ és $\lambda_2 = -1$, és mindegyik kétszeres sajátérték. A $\lambda_1 = 1$ és $\lambda_2 = -1$ helyhez tartozó Hermite-féle alappolinomoknak a következő feltételeket kell kielégíteni:

3.4 ... ELŐÁLLÍTÁSA A MINIMÁLPOLINOM TÖBBSZÖRÖS GYÖKEI ESETÉN 345

$$\begin{aligned} H_{10}(1) &= 1, & H'_{10}(1) &= H_{10}(-1) = H'_{10}(-1) = 0, \\ H_{20}(-1) &= 1, & H'_{20}(-1) &= H_{20}(1) = H'_{20}(1) = 0. \end{aligned}$$

Ezekből adódik

$$H_{10}(x) = \frac{1}{4}(2 + 3x - x^3), \quad H_{20}(x) = \frac{1}{4}(2 - 3x + x^3).$$

Az x változó helyére az \mathbf{A} mátrixot írva, a $H_{k0}(\mathbf{A})$ megfelelő mátrixpolinomat kapjuk, amelyek szükségképpen másodrangú projektorok lesznek:

$$\begin{aligned} H_{10}(\mathbf{A}) &= \frac{1}{4}(2\mathbf{E} + 3\mathbf{A} - \mathbf{A}^3) = \\ &= \begin{bmatrix} 2 & 0 & -2 & 0 \\ 0 & 2 & 0 & -2 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} = \mathbf{U}_1 \mathbf{V}_1^T, \end{aligned}$$

$$\begin{aligned} H_{20}(\mathbf{A}) &= \frac{1}{4}(2\mathbf{E} - 3\mathbf{A} + \mathbf{A}^3) = \\ &= \begin{bmatrix} -1 & 0 & 2 & 0 \\ 0 & -1 & 0 & 2 \\ -1 & 0 & 2 & 0 \\ 0 & -1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 2 & 0 \\ 0 & -1 & 0 & 2 \end{bmatrix} = \mathbf{U}_2 \mathbf{V}_2^T. \end{aligned}$$

A diadikus felbontással kapott \mathbf{U}_1 , \mathbf{U}_2 , ill. \mathbf{V}_1^T , \mathbf{V}_2^T mátrixok segítségével meghatározható az a hasonlósági transzformáció, amely az adott mátrixot másodrendű blokkokból álló kvázidiagonál-mátrix alakba viszi át:

$$\begin{aligned} \mathbf{U} &= \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \quad \mathbf{V}^T = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 2 & 0 \\ 0 & -1 & 0 & 2 \end{bmatrix}, \\ \mathbf{V}^T \mathbf{A} \mathbf{U} &= \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 2 & 0 \\ 0 & -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 3 & -1 & -4 & 2 \\ 2 & 3 & -2 & -4 \\ 2 & -1 & -3 & 2 \\ 1 & 2 & -1 & -3 \end{bmatrix} \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} = \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \end{aligned}$$

A kapott kvázidiagonál-mátrix bal felső blokkja már alsó Jordan-féle normálalakú, a jobb alsó blokkot pedig még egy hasonlósági transzformációval alsó Jordan-féle normálalakra kell hozni.

A (3.4.38) és (3.4.39) összefüggések segítségével ez az $\mathbf{I} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ involutó-
rius permutáló mátrixszal végezhető el: $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$.

Az egyes nilpotens blokkokat a fentiek alapján a

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

mátrixszal végzett hasonlósági transzformáció redukálja Jordan-féle normálalakra, az adott mátrixot pedig az

$$\mathbf{UT} = \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 2 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

mátrixszal végzett hasonlósági transzformáció, ahol

$$(\mathbf{UT})^{-1} = \mathbf{TV}^T = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & -1 & 0 & 2 \\ -1 & 0 & 2 & 0 \end{bmatrix}.$$

Ezzel \mathbf{A} kanonikus előállítása az alábbi alakban adódik:

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 2 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ \hline 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{array} \right] \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & -1 & 0 & 2 \\ -1 & 0 & 2 & 0 \end{bmatrix},$$

ahonnan

$$\sin \mathbf{A} = \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 2 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \left[\begin{array}{cc|cc} \sin 1 & 0 & 0 & 0 \\ \cos 1 & \sin 1 & 0 & 0 \\ \hline 0 & 0 & \sin(-1) & 0 \\ 0 & 0 & \cos(-1) & \sin(-1) \end{array} \right] \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & -1 & 0 & 2 \\ -1 & 0 & 2 & 0 \end{bmatrix}.$$

* * *

3.5 ELEMI OSZTÓK ELMÉLETE

A lineáris transzformációk elméletéből már tudjuk, hogy a bázis megváltoztatása esetén a lineáris transzformáció mátrixán hasonlósági transzformációt kell végrehajtani. Ezért adott lineáris transzformáció jellemző adatai a transzformáció mátrixának csak olyan adatai lehetnek, amelyek a mátrix hasonlósági transzformációjával szemben invariánsak. Ebben a pontban megmutatjuk, hogy a legáltalánosabb lineáris transzformáció, amely az n -dimenziós lineáris teret önmagára képezi le, *elemi osztóival* jellemezhető egyértelműen. Mint látni fogjuk, a transzformáció mátrixának Jordan-féle normálalakjából ezek egyszerűen kiolvashatók, és a mátrix hasonlósági transzformációjával szemben valóban invariánsak.

3.5.1 A determinánsosztó invarianciája

E fejezet előző pontjaiban megismerkedtünk mátrixok Jordan-féle normálalakjának fogalmával és megmutattuk, hogy hasonlósági transzformációval bármely mátrix Jordan-blokkokból álló kvázidiagonál-mátrix alakjára hozható. Láttuk, hogy a mátrix minden sajátértékéhez *legalább* egy Jordan-blokk tartozik és *legfeljebb* annyi, amennyi a sajátérték multiplicitása a karakterisztikus egyenletben. Ez utóbbiból következik, hogy az „elsőrendű” Jordan-blokkok egyetlen elemmé fajulnak, és speciális esetként az egyszerű struktúrájú, vagyis diagonalizálható mátrixokhoz jutunk. Ha a minimálpolinom megegyezik a karakterisztikus polinommal, akkor a mátrix Jordan-féle normálalakja éppen annyi Jordan-blokkból áll, amennyi a különböző sajátértékek száma, és egy-egy Jordan-blokk rendszáma megegyezik a megfelelő sajátérték multiplicitásával. Az alábbiakban annak az esetnek a vizsgálatával foglalkozunk részletesebben, amikor a minimálpolinom fokszáma kisebb, mint a karakterisztikus polinomé, de gyökei nem mind egyszeresek.

Legyen az n -edrendű \mathbf{A} mátrix karakterisztikus polinomja

$$|\lambda \mathbf{E} - \mathbf{A}| \equiv D(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{\alpha_k}, \text{ ahol } \sum_{k=1}^s \alpha_k = n.$$

Vezessük be a következő fogalmakat.

3.5.1 definíció. A $\lambda \mathbf{E} - \mathbf{A}$ karakterisztikus mátrix ν -edrendű minorainak legnagyobb közös osztóját (amely egy állandó szorzótól eltekintve van csak meghatározva, és ezt úgy választhatjuk meg, hogy a legmagasabb fokú tag együtt-hatója 1 legyen) az \mathbf{A} mátrix ν -edik **determinánsosztójának** nevezzük és $D_\nu(\lambda)$ -val jelöljük.

Mivel egy mátrix tetszőleges, ν -edrendű minora $(\nu - 1)$ -edrendű minorok lineáris kombinációjaként fejezhető ki, ezért bármely ν -edrendű minor osztó

ható a $(\nu - 1)$ -edrendű minorok bármely közös osztójával. Ebből következik, hogy $D_\nu(\lambda)$ osztható a $D_{\nu-1}(\lambda)$ polinommal $(\nu = 1, 2, \dots, n)$.

3.5.2 definíció. Az \mathbf{A} mátrix bármely két, egymás utáni $D_\nu(\lambda)$ és $D_{\nu-1}(\lambda)$ determinánsosztójának

$$(3.5.1) \quad E_\nu(\lambda) = \frac{D_\nu(\lambda)}{D_{\nu-1}(\lambda)} \quad (\nu = 1, 2, \dots, n)$$

hányadosát az \mathbf{A} mátrix **invariáns faktorának** nevezzük.

E definíciók értelmében $D_n(\lambda)$ megegyezik a karakterisztikus polinommal, az $E_n(\lambda)$ invariáns faktor pedig a minimálpolinommal. Írjuk fel az $E_\nu(\lambda)$ invariáns faktorokat gyöktényezős alakban:

$$(3.5.2) \quad E_\nu(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{e_{\nu k}} \quad (\nu = 1, 2, \dots, n),$$

ahol $e_{nk} = \gamma_k$ a minimálpolinom gyökeinek multiplicitása.

3.5.3 definíció. Az invariáns faktorok (3.5.2) alatti előállításában az állandótól különböző $(\lambda - \lambda_k)^{e_{\nu k}}$ hatványokat a $\lambda \mathbf{E} - \mathbf{A}$ karakterisztikus mátrix (komplex számtest felett értelmezett) **elemi osztóinak** nevezzük.

3.5.1 tétel. Legyen az \mathbf{A} lineáris transzformáció mátrixa egy adott bázisban \mathbf{A} . A $\lambda \mathbf{E} - \mathbf{A}$ karakterisztikus mátrix k -adrendű minorainak $D_k(\lambda)$ legnagyobb közös osztója nem függ a bázis megválasztásától, tehát a koordináta-rendszer transzformációjával szemben invariáns.

Bizonyítás. Szorozzuk meg a $\lambda \mathbf{E} - \mathbf{A}$ karakterisztikus mátrixot tetszőleges \mathbf{C} , állandó elemű, nonszinguláris mátrixszal és tekintsük a $(\lambda \mathbf{E} - \mathbf{A})\mathbf{C}$ szorzat egy tetszőleges $i_1 i_2 \dots i_r$ indexű soraiból és a $j_1 j_2 \dots j_r$ indexű oszlopaiból alkotott r -edrendű minormátrixát. Az (1.2.6) összefüggés alapján

$$\{(\lambda \mathbf{E} - \mathbf{A})\mathbf{C}\}_{j_1 j_2 \dots j_r}^{i_1 i_2 \dots i_r} = (\lambda \mathbf{E} - \mathbf{A})_{j_1 j_2 \dots j_r}^{i_1 i_2 \dots i_r} \mathbf{C}_{j_1 j_2 \dots j_r}^{1 \ 2 \dots n}$$

adódik, vagyis az így kapott mátrix egy-egy oszlopa a $\lambda \mathbf{E} - \mathbf{A}$ mátrix *valamennyi* oszlopa $i_1 i_2 \dots i_r$ indexű eleméből alkotott oszlopoknak a lineáris kombinációja. A minormátrix determinánsa tehát az ezen oszlopokból összes lehetséges módon kiválasztható r oszlop által alkotott determinánsok lineáris kombinációja. Ez azt jelenti, hogy a $\lambda \mathbf{E} - \mathbf{A}$ mátrix valamennyi r -edrendű minorának bármelyik közös osztója (tehát a legnagyobb közös osztója is) egyúttal osztója a $(\lambda \mathbf{E} - \mathbf{A})\mathbf{C}$ tetszőlegesen kiválasztott r -edrendű minorának. Mivel pedig a $(\lambda \mathbf{E} - \mathbf{A})\mathbf{C}$ mátrixból a $\lambda \mathbf{E} - \mathbf{A}$ mátrixot megkapjuk, ha a \mathbf{C}^{-1} reciprokmátrixszal szorozzuk, fordítva is igaz: a $(\lambda \mathbf{E} - \mathbf{A})\mathbf{C}$ mátrix k -adrendű minorainak bármely osztója egyúttal osztója a $\lambda \mathbf{E} - \mathbf{A}$ karakterisztikus mátrix

k -adrendű minorainak is. Tehát a $\lambda \mathbf{E} - \mathbf{A}$ és a $(\lambda \mathbf{E} - \mathbf{A})\mathbf{C}$ mátrix k -adrendű minorainak közös osztói megegyeznek. Ugyanez érvényes a $(\lambda \mathbf{E} - \mathbf{A})\mathbf{C}$ és a $\mathbf{C}^{-1}(\lambda \mathbf{E} - \mathbf{A})\mathbf{C}$ mátrixokra, vagyis a $\lambda \mathbf{E} - \mathbf{A}$, valamint a $\mathbf{C}^{-1}(\lambda \mathbf{E} - \mathbf{A})\mathbf{C}$ mátrix k -adrendű minorainak közös osztói, és így a legnagyobb közös osztói is megegyeznek. Ez azt jelenti, hogy az \mathbf{A} mátrixhoz tartozó karakterisztikus mátrix, valamint a hasonlósági transzformációval nyert $\tilde{\mathbf{A}} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$ mátrixhoz tartozó $\lambda \mathbf{E} - \tilde{\mathbf{A}}$ karakterisztikus mátrix k -adrendű minorainak legnagyobb közös osztói megegyeznek. Mivel pedig a hasonlósági transzformáció a bázis lineáris transzformációját jelenti, ezért a $D_k(\lambda)$ legnagyobb közös osztó valóban független a bázis megválasztásától, tehát a koordináta-rendszer transzformációjával szemben invariáns. ■

Egy lineáris transzformációhoz tartozó $D_k(\lambda)$ determinánsosztók meghatározása céljából tehát olyan bázist célszerű választani, amelyben a transzformáció mátrixa minél egyszerűbb „szerkezetű”. Ezért úgy választjuk meg a bázist, hogy a lineáris transzformáció mátrixa e bázisban Jordan-féle normálalakú legyen.

3.5.2 A determinánsosztó invarianciája speciális esetben

Először határozzuk meg egyetlen Jordan-blokkhoz tartozó $D_k(\lambda)$ determinánsosztók sorozatát. Jelölje \mathbf{J}_n azt az egyetlen Jordan-blokkot tartalmazó n -edrendű Jordan-féle normálalakú mátrixot, amelynek n -szeres sajátértéke λ_0 :

$$\mathbf{J}_n = \begin{bmatrix} \lambda_0 & 0 & 0 \dots & 0 \\ 1 & \lambda_0 & 0 \dots & 0 \\ 0 & 1 & \lambda_0 \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 \dots 1 & \lambda_0 \end{bmatrix};$$

a karakterisztikus polinom nyilvánvalóan $|\lambda \mathbf{E} - \mathbf{J}_n| \equiv D_n(\lambda) = (\lambda - \lambda_0)^n$. A $\lambda \mathbf{E} - \mathbf{J}_n$ karakterisztikus mátrix első sorának utolsó eleméhez tartozó $(n-1)$ -edrendű minor előjeltől eltekintve 1, ezért az $(n-1)$ -edrendű minorok legnagyobb közös osztója is 1, tehát

$$(3.5.3) \quad D_{n-1}(\lambda) \equiv 1.$$

Ebből már következik, hogy az összes többi $D_k(\lambda)$ polinom is $(k < n-1)$ azonosan 1, hiszen mindegyik $D_k(\lambda)$ polinom osztója a nagyobb indexű megfelelő polinomnak. Tehát egyetlen Jordan-blokkra a $D_k(\lambda)$ determinánsosztók sorozata a következő:

$$(3.5.4) \quad D_n(\lambda) = (\lambda - \lambda_0)^n, \quad D_{n-1}(\lambda) = D_{n-2}(\lambda) = \dots = D_1(\lambda) = 1.$$

Tekintsünk most egy olyan \mathbf{J}_n mátrixot, amelynek egyetlen n -szeres – sajátértéke van, de amelynek a Jordan-féle normálalakja p számú Jordan-

blokkból áll; ezek rendszáma monoton nemnövekvő sorrendben legyen

$$(3.5.5) \quad e_n \geq e_{n-1} \geq \cdots \geq e_{n-p+1} \neq 0, \quad e_{n-p} = e_{n-p-1} = \cdots = e_1 = 0.$$

vagyis

$$(3.5.6) \quad \mathbf{J}_n = \begin{array}{c} \begin{array}{cccc} \overbrace{\hspace{1.5cm}}^{e_n} & \overbrace{\hspace{1.5cm}}^{e_{n-1}} & \cdots & \overbrace{\hspace{1.5cm}}^{e_{n-p+1}} \end{array} \\ \left[\begin{array}{cccc} \lambda_0 & & & \\ 1 & \lambda_0 & & \\ & 1 & \lambda_0 & \\ & & \ddots & \lambda_0 \\ & & & \lambda_0 \end{array} \right] & \begin{array}{c} (1) \\ \\ \\ \end{array} \\ \left[\begin{array}{cccc} & \lambda_0 & & \\ & 1 & \lambda_0 & \\ & & 1 & \lambda_0 \\ & & & \ddots \\ & & & \lambda_0 \end{array} \right] & \begin{array}{c} (2) \\ \\ \\ \end{array} \\ \vdots & \vdots \\ \left[\begin{array}{cccc} & & & \lambda_0 \\ & & & 1 \\ & & & \lambda_0 \end{array} \right] & \begin{array}{c} \\ \\ (p) \end{array} \end{array}$$

A karakterisztikus polinom ezúttal is $|\lambda \mathbf{E} - \mathbf{J}_n| \equiv D_n(\lambda) = (\lambda - \lambda_0)^n$. Mivel a karakterisztikus mátrix egy $(n-1)$ -edrendű minorát úgy kapjuk meg, hogy a $(\lambda \mathbf{E} - \mathbf{J}_n)$ mátrix valamely sorát és oszlopát töröljük, ezért az egyetlen Jordan-blokknál követett meggondolás alapján látható, hogy a (3.5.6) mátrixhoz tartozó karakterisztikus mátrix $(n-1)$ -edrendű minorai közül az első sor e_n -edik eleméhez tartozó minor tartalmazza a $\lambda - \lambda_0$ gyöktényezőt a legalacsonyabb hatványon, tehát ez lesz a minorok legnagyobb közös osztója. A (3.5.4) összefüggést most a bal felső, e_n -edrendű blokkra alkalmazva, azt kapjuk, hogy

$$(3.5.7) \quad D_{n-1}(\lambda) = (\lambda - \lambda_0)^{n-e_n}.$$

Ugyanezt a meggondolást tovább folytatva, meghatározható a többi determinánsosztó is. Így a $D_{n-2}(\lambda)$ polinomhoz úgy jutunk, hogy a $(\lambda \mathbf{E} - \mathbf{J}_n)$ karakterisztikus mátrixban töröljük az első és az (e_n+1) -edik sort, valamint az e_n -edik és az (e_n+e_{n-1}) -edik oszlopot. Ismét (3.5.4) miatt az $(n-2)$ -edrendű minorok közül ez fogja legalacsonyabb hatványon tartalmazni a $\lambda - \lambda_0$ gyöktényezőt, tehát ez lesz az $(n-2)$ -edrendű minorok legnagyobb közös osztója:

$$(3.5.8) \quad D_{n-2}(\lambda) = (\lambda - \lambda_0)^{n-e_n-e_{n-1}}.$$

Az eljárást folytatva, megkapjuk az összes többi $D_\nu(\lambda)$ polinomot is:

$$\begin{aligned} D_{n-3}(\lambda) &= (\lambda - \lambda_0)^{n-e_n-e_{n-1}-e_{n-2}}, \\ &\dots\dots\dots \\ D_{n-p}(\lambda) &= 1. \end{aligned}$$

3.5.3 A determinánsosztó invarianciája általános esetben

Tekintsük most az általános esetet, amikor a lineáris transzformációnak s különböző sajátértéke van, és a transzformáció mátrixának Jordan-féle normálalakjában az α_k multiplicitású λ_k sajátértékhez tartozó p_k darab Jordan-blokk $e_{\nu k}$ rendszámai, ismét monoton nemnövekvő sorrendben,

$$(3.5.9) \quad \begin{aligned} e_{nk} &\geq e_{n-1,k} \geq \dots \geq e_{n-p_k+1,k} \neq 0, \\ e_{n-p_k,k} &= e_{n-p_k-1,k} = \dots = e_{1k} = 0. \end{aligned}$$

(Félreérthető jelölések elkerülése végett itt és a továbbiakban a kettős indexben vesszővel választjuk el a két indexet, ha legalább az egyik kéttagú kifejezés.) A fenti gondolatmenet segítségével könnyen belátható, hogy a ν -edrendű minorok legnagyobb közös osztója ebben az esetben az egyes sajátértékekhez tartozó Jordan-blokkokból külön-külön meghatározható ν -edrendű minorok legnagyobb közös osztóinak szorzata lesz, azaz

$$\begin{aligned} D_n(\lambda) &= \prod_{k=1}^s (\lambda - \lambda_k)^{\alpha_k}, \quad \text{ahol} \quad \sum_{k=1}^s \alpha_k = n, \\ D_{n-1}(\lambda) &= \prod_{k=1}^s (\lambda - \lambda_k)^{\alpha_k - e_{nk}}, \\ (3.5.10) \quad D_{n-2}(\lambda) &= \prod_{k=1}^s (\lambda - \lambda_k)^{\alpha_k - e_{nk} - e_{n-1,k}}, \\ D_{n-3}(\lambda) &= \prod_{k=1}^s (\lambda - \lambda_k)^{\alpha_k - e_{nk} - e_{n-1,k} - e_{n-2,k}} \quad \text{stb.} \end{aligned}$$

A 3.5.2 definíció felhasználásával határozzuk meg az adott lineáris transzformáció invariáns faktorait. Képezve a (3.5.10) képlettel megadott determinánsosztók hányadosát, az invariáns faktorok (3.5.1) alapján

$$(3.5.11) \quad E_\nu(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{e_{\nu k}}$$

alakban írhatók fel. A kapott eredményt a következő tételben foglaljuk össze.

3.5.2 tétel. Tegyük fel, hogy egy Jordan-féle normálalakú mátrixban a λ_k sajátértékhez p_k számú Jordan-blokk tartozik, amelyek rendszáma monoton nemnövekvő sorrendben

$$(3.5.12) \quad \begin{aligned} e_{nk} &\geq e_{n-1,k} \geq \cdots \geq e_{n-p_k+1,k} \neq 0, \\ e_{n-p_k,k} &= e_{n-p_k-1,k} = \cdots = e_{1k} = 0 \quad (k = 1, 2, \dots, s), \end{aligned}$$

azaz $\mathbf{J} = \langle \mathbf{J}_1 \quad \mathbf{J}_2 \dots \mathbf{J}_s \rangle$, ahol

$$(3.5.13) \quad \mathbf{J}_k = \begin{array}{c} \begin{array}{cccc} \overbrace{\lambda_k}^{e_{nk}} & \overbrace{\lambda_k}^{e_{n-1,k}} & \cdots & \overbrace{\lambda_k}^{e_{n-p_k+1,k}} \\ \begin{array}{c} \lambda_k \\ 1 \quad \lambda_k \\ \cdots \\ 1 \quad \lambda_k \end{array} & & & \\ \hline & \begin{array}{c} \lambda_k \\ 1 \quad \lambda_k \\ \cdots \\ 1 \quad \lambda_k \end{array} & & \\ \hline & & \ddots & \\ \hline & & & \begin{array}{c} \lambda_k \\ 1 \quad \lambda_k \end{array} \\ \hline & & & \lambda_k \end{array} \begin{array}{l} (1) \\ (2) \\ \vdots \\ (p_k) \end{array} \end{array}$$

Az ehhez tartozó invariáns faktorokra ekkor a következő adódik:

$$(3.5.14) \quad E_n(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{e_{nk}},$$

ahol

$$(3.5.15) \quad \sum_{k=1}^s e_{nk} = m$$

az $E_n(\lambda) \equiv \Delta(\lambda)$ minimálpolinom fokszáma, továbbá

$$(3.5.16) \quad E_\nu(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{e_{\nu k}} \quad (\nu = n-1, n-2, \dots, 2, 1),$$

ahol

$$(3.5.17) \quad \sum_{\nu=1}^n e_{\nu k} = \alpha_k \quad (k = 1, 2, \dots, s).$$

A (3.5.12) összefüggés alapján a kitevőkre fennáll:

$$e_{1k} = e_{2k} = \dots = e_{n-p_k, k} = 0,$$

tehát (3.5.17) tulajdonképpen egy p_k tagú összeg:

$$(3.5.18) \quad \sum_{\nu=n-p_k+1}^n e_{\nu k} = \alpha_k \quad (k = 1, 2, \dots, s).$$

3.5.4 Az elemi osztók és a Jordan-féle normálalak

Látható tehát, hogy az invariáns faktorok sorozata egyértelműen meghatározza a mátrix Jordan-féle normálalakját. Mivel továbbá az invariáns faktorok tényezői a 3.5.3 definícióval bevezetett elemi osztók, azt is mondhatjuk, hogy az elemi osztók egyértelműen meghatározzák a mátrix Jordan-féle normálalakját. Ha tehát ismerjük egy mátrix sajátértékeit, akkor az elemi osztók kitevőit táblázatba foglalva, pontos képet kaphatunk a Jordan-féle normálalakú mátrix szerkezetéről:

	λ_1	λ_2	\dots	λ_s	
n	e_{n1}	e_{n2}	\dots	e_{ns}	$\sum_{k=1}^s e_{nk} = m$
$n-1$	$e_{n-1,1}$	$e_{n-1,2}$	\dots	$e_{n-1,s}$	
$n-2$	$e_{n-2,1}$	$e_{n-2,2}$	\dots	$e_{n-2,s}$	
\vdots	\dots	\dots	\dots	\dots	
	e_{11}	e_{12}	\dots	e_{1s}	
	α_1	α_2	\dots	α_s	$\sum_{k=1}^s \alpha_k = n$
	p_1	p_2	\dots	p_s	

Az első oszlopban áll az invariáns faktor rendszáma; a táblázat belső oszlopaiban egy-egy sajátértéknek megfelelő kitevők állnak, melyek monoton nemnövekvő sorozatot alkotnak. Ezek összege a sajátérték α_k multiplicitását adja (utóbbiak összege pedig a mátrix n rendszáma). Az egyes sorokban álló kitevők összege a megfelelő invariáns faktorok fokszámával egyenlő; az első sorban álló kitevők a minimálpolinom gyökeinek a multiplicitását adják, ezek összege pedig a minimálpolinom m fokszáma. A táblázat elemei között természetesen zérus is lehet, ezeket nem szükséges kiírni. Az egyes oszlopokban álló, zérustól különböző kitevők *száma* a megfelelő sajátértékhez tartozó Jordan-blokkok p_k számát jelenti (ezt tüntettük fel a táblázat alján, az α_k multiplicitások alatt). Érdekes megfigyelni, hogy a p_k szám egyúttal a $\lambda_k \mathbf{E} - \mathbf{A}$ mátrix rangcsökkenését (defektusát, nullitását) is megadja, tehát p_k a λ_k sajátértékhez tartozó lineárisan független sajátvektorok száma. Ez

egyszerűen abból következik, hogy egyetlen Jordan-blokk, ha a főátlóban álló elemek zérussal egyenlők, nemderogatórius nilpotens mátrix, a rangja eggyel kisebb, mint rendszáma, vagyis a defektusa 1. Ha tehát \mathbf{A} Jordan-féle normálalakú mátrix és a $\lambda_k \mathbf{E} - \mathbf{A}$ mátrixban a λ_k sajátértékhez p_k Jordan-blokk tartozik, akkor mindegyik blokk defektusa 1, vagyis $\lambda_k \mathbf{E} - \mathbf{A}$ defektusa p_k . Ez azt jelenti, hogy a λ_k sajátértékhez tartozó \mathbf{u}_k sajátvektorokat meghatározó $(\lambda_k \mathbf{E} - \mathbf{A})\mathbf{u}_k = \mathbf{0}$ egyenletnek éppen p_k lineárisan független megoldása van.

A táblázatból az is kiolvasható, hogy egy mátrix akkor és csak akkor diagonalizálható, ha valamennyi elemi osztója elsőfokú, tehát ha a táblázat valamennyi, zérustól különböző eleme 1. Ekkor egyúttal minden k -ra $p_k = \alpha_k$, tehát minden sajátértékhez annyi lineárisan független sajátvektor tartozik, amennyi a multiplicitása.

A Jordan-féle normálalak struktúráját tehát az elemi osztók kitevőinek táblázata egyértelműen meghatározza. Ha a zérus elemeket elhagyjuk és az egymás alatt álló kitevőket zárójelbe foglalva, egymás mellé írjuk a különböző sajátértékekhez tartozó, zérustól különböző kitevők rendszerét, akkor ezzel a Jordan-féle normálalakot tömör írásmóddal jellemezhetjük. Ebből a célból bevezetjük a Segre-féle karakterisztika fogalmát.

3.5.4 definíció. A mátrix Segre*-féle karakterisztikájának nevezzük az elemi osztók kitevőinek

$$(3.5.19) \quad [(e_{n1} \ e_{n-1,1} \ \dots)(e_{n2} \ e_{n-1,2} \ \dots) \dots (e_{ns} \ e_{n-1,s} \ \dots)]$$

alakban megadott rendszerét.

22. Példa. Írjuk fel annak a mátrixnak a Jordan-féle normálalakját és invariáns faktorait, amelynek Segre-féle karakterisztikája

$$[(2 \ 2 \ 1)(3 \ 1)(2)(1)].$$

Megoldás. A Segre-féle karakterisztikából látható, hogy a keresett mátrixnak négy, különböző sajátértéke van. Az egyes sajátértékek legyenek $\lambda_1, \lambda_2, \lambda_3, \lambda_4$. Az elemi osztók kitevőit az előbb ismertetett felépítésű táblázatba foglaljuk. A mátrix rendszáma egyenlő a karakterisztikus polinom fokszámával, azaz a kitevők összegével – esetünkben 12-vel.

	λ_1	λ_2	λ_3	λ_4	
12	2	3	2	1	$m = 8$
11	2	1			
10	1				
α_k	5	4	2	1	$n = 12$
p_k	3	2	1	1	

*E. Segre (1903–1977) olasz matematikus.

Innen kiolvashatók az alábbi egyszerű összefüggések.

A karakterisztikus polinom: $D(\lambda) = (\lambda - \lambda_1)^5(\lambda - \lambda_2)^4(\lambda - \lambda_3)^2(\lambda - \lambda_4)$.

A $\Delta(\lambda)$ minimálpolinom (együttal az $E_{12}(\lambda)$ invariáns faktor):

$$\Delta(\lambda) \equiv E_{12}(\lambda) = (\lambda - \lambda_1)^2(\lambda - \lambda_2)^3(\lambda - \lambda_3)^2(\lambda - \lambda_4).$$

A többi invariáns faktor pedig

$$E_{11}(\lambda) = (\lambda - \lambda_1)^2(\lambda - \lambda_2);$$

$$E_{10}(\lambda) = \lambda - \lambda_1;$$

$$E_9(\lambda) = \cdots = E_1(\lambda) = 1.$$

Az egyes sajátértékekhez tartozó defektusok a következők: $p_1 = 3$; $p_2 = 2$; $p_3 = 1$; $p_4 = 1$. Ez annyit jelent, hogy az egyes sajátértékekhez tartozó lineárisan független sajátvektorok száma rendre 3, 2, 1, 1, vagyis a 12-edrendű mátrixhoz összesen 7 lineárisan független sajátvektor tartozik. A keresett Jordan-féle normálmátrix:

$$(3.5.20) \quad \begin{bmatrix} \lambda_1 & & & & & & \\ & 1 & \lambda_1 & & & & \\ & & & \lambda_1 & & & \\ & & & & 1 & \lambda_1 & \\ & & & & & & \lambda_1 \\ & & & & & & \\ \hline & & & \lambda_2 & & & \\ & & & & 1 & \lambda_2 & \\ & & & & & 1 & \lambda_2 \\ & & & & & & \lambda_2 \\ \hline & & & & & & \lambda_3 \\ & & & & & 1 & \lambda_3 \\ \hline & & & & & & \lambda_4 \end{bmatrix}.$$

* * *

Az elemi osztók kitevőit tartalmazó táblázat bevezetésekor már rámutattunk arra, hogy az egyes oszlopokban álló, zérustól különböző elemek száma egyenlő a λ_k sajátértékhez tartozó Jordan-blokkok p_k számával, és ugyancsak egyenlők a $\lambda_k \mathbf{E} - \mathbf{A}$ mátrix defektusával.

Vegyünk most figyelembe, hogy $\lambda_k \mathbf{E} - \mathbf{A}$ hatványozása esetén a λ_k sajátértékhez tartozó Jordan-blokkokban a -1 elemek egy hellyel balra tolódnak, és ezzel a mátrix defektusa minden egyes hatványozáskor annnyival nő, amennyi a blokkok száma. Így eljutunk egy olyan hatványhoz, amelyben a legkisebb rendszámú Jordan-blokk zérussá válik, ettől kezdve tehát a $\lambda_k \mathbf{E} - \mathbf{A}$ mátrix

defektusa kisebb számmal fog nőni (hiszen a Jordan-blokkok száma kevesebb lett). Jelölje d_{k1} a $\lambda_k \mathbf{E} - \mathbf{A}$ mátrix defektusát, $d_{k1} + d_{k2}$ a $(\lambda_k \mathbf{E} - \mathbf{A})^2$ mátrix defektusát, és így tovább. Végül $(\lambda_k \mathbf{E} - \mathbf{A})^{\beta_k}$ defektusa, $d_{k1} + d_{k2} + \dots + d_{k\beta_k}$, egyenlő lesz a λ_k sajátérték multiplicitásával, ha β_k a $\lambda_k \mathbf{E} - \mathbf{A}$ mátrix (λ_k -hoz tartozó) nilpotens blokkjának az indexe. A fenti megfontolásból következik, hogy a $d_{k\nu}$ számok monoton nemnövekvő sorozatot alkotnak:

$$(3.5.21) \quad p_k = d_{k1} \geq d_{k2} \geq \dots \geq d_{k\beta_k} \geq 1.$$

Az így nyert $d_{k\nu}$ számok és az elemi osztók $e_{\nu k}$ kitevői között kölcsönösen egyértelmű megfeleltetés létesíthető, tehát a $d_{k\nu}$ számok a mátrix Jordan-féle normálalakját ugyanúgy egyértelműen meghatározzák, mint az elemi osztók $e_{\nu k}$ kitevői. A $d_{k\nu}$ számokat a Segre-féle karakterisztikához hasonlóan csoportosíthatjuk, így kapjuk a mátrix Weyr-féle karakterisztikáját.

3.5.5 definíció. Jelölje λ_k ($k = 1, 2, \dots, s$) az \mathbf{A} mátrix különböző sajátértékeit, és képezzük a $(\lambda_k \mathbf{E} - \mathbf{A})$ karakterisztikus mátrix hatványainak $(\lambda_k \mathbf{E} - \mathbf{A})^\nu$ sorozatát ($\nu = 1, 2, \dots$). Jelölje e hatványok defektusát rendre

$$d_{k1}, \quad d_{k1} + d_{k2}, \quad \dots, \quad d_{k1} + d_{k2} + \dots + d_{k\nu}, \quad \dots$$

Az így meghatározott $d_{k\nu}$ számok monoton nemnövekvő sorozatát minden k értékre egy-egy csoportba foglalva, ezek

$$[(d_{11}, \quad d_{12}, \quad \dots)(d_{21}, \quad d_{22}, \quad \dots) \dots (d_{s1}, \quad d_{s2}, \quad \dots)]$$

alakban megadott rendszerét az \mathbf{A} mátrix **Weyr-féle karakterisztikájának** nevezzük.

A Segre-féle karakterisztikából úgy kapjuk meg a Weyr-féle karakterisztikát, hogy minden csoportba először a zárójelben álló elemek számát írjuk be, ezután a zárójelen belül minden számból levonunk 1-et, megszámloljuk a zérustól különböző elemeket és ezek számát az előző mellé írjuk. Az eljárást addig folytatjuk, amíg zérustól különböző számot találunk a Segre-féle karakterisztika tekintett csoportjában. Az így kapott számok alkotják a Weyr-féle karakterisztika megfelelő csoportját. Ezután rátérünk a következő zárójelre stb.

Abban az esetben, ha az elemi osztók mind lineárisak, vagyis ha

$$e_{\nu k} = 1 \quad (\nu = n, n-1, \dots, n-p_k+1; k = 1, 2, \dots, s),$$

és így $p_k = \alpha_k$, akkor a Segre-féle karakterisztika

$$[(\underbrace{1 \quad 1 \quad \dots \quad 1}_{\alpha_1})(\underbrace{1 \quad 1 \quad \dots \quad 1}_{\alpha_2}) \dots (\underbrace{1 \quad 1 \quad \dots \quad 1}_{\alpha_s})],$$

innen a Weyr-féle karakterisztika $[(\alpha_1) \ (\alpha_2) \ \dots (\alpha_s)]$. Ez azt jelenti, hogy egy diagonalizálható mátrix Weyr-féle karakterisztikáját a sajátértékek multiplicitásai alkotják: $d_{k1} = \alpha_k$, tehát a λ_k sajátértékhez α_k lineárisan független sajátvektor tartozik. Így a lineárisan független sajátvektorok száma megegyezik az α_k multiplicitások összegével, vagyis a mátrix rendszámával. Tekintettel arra, hogy a Weyr-féle karakterisztika ekkor egyetlen számot tartalmazó csoportokból áll, azaz $d_{k2} = 0$, ez azt jelenti, hogy $\lambda_k \mathbf{E} - \mathbf{A}$ hatványozásával annak defektusa nem nő. Ezzel gyakorlatilag is jól használható kritériumot találtunk annak eldöntésére, hogy egy mátrix diagonalizálható-e, vagyis hogy egyszerű struktúrájú-e. Eszerint *egy mátrix akkor és csak akkor egyszerű struktúrájú, ha karakterisztikus mátrixának rangja – bármelyik sajátérték behelyettesítésével – a hatványozással szemben invariáns*.

23. Példa. írjuk fel az előző példa Weyr-féle karakterisztikáját, és értékeljük jelentését.

Megoldás. A Segre-féle karakterisztika: $[(2 \ 2 \ 1)(3 \ 1)(2)(1)]$. Vonjunk le mindegyik számból 1-et addig, amíg mindegyik szám zérus nem lesz (ahol már zérust kaptunk, ott a továbbiakban pontot írunk)

$$\begin{aligned} &[(1 \ 1 \ 0)(2 \ 0)(1)(0)] \\ &[(0 \ 0 \ .)(1 \ .)(0)(.)] \\ &[(. \ . \ .)(0 \ .)(.)(.)]. \end{aligned}$$

Ezek alapján a Weyr-féle karakterisztikát – csoportonként haladva – a zérustól különböző elemek száma szolgáltatja: $[(3 \ 2)(2 \ 1 \ 1)(1 \ 1)(1)]$. A (3.5.20) mátrix szerkezetéből látható, hogy a Weyr-féle karakterisztika jelentése a következő:

$$\begin{aligned} \lambda_1 \mathbf{E} - \mathbf{A} &\quad \text{defektusa } 3 \\ (\lambda_1 \mathbf{E} - \mathbf{A})^2 &\quad \text{defektusa } 3 + 2 \\ \lambda_2 \mathbf{E} - \mathbf{A} &\quad \text{defektusa } 2 \\ (\lambda_2 \mathbf{E} - \mathbf{A})^2 &\quad \text{defektusa } 2 + 1 \\ (\lambda_2 \mathbf{E} - \mathbf{A})^3 &\quad \text{defektusa } 2 + 1 + 1 \\ (\lambda_3 \mathbf{E} - \mathbf{A}) &\quad \text{defektusa } 1 \\ (\lambda_3 \mathbf{E} - \mathbf{A})^2 &\quad \text{defektusa } 1 + 1 \\ (\lambda_4 \mathbf{E} - \mathbf{A}) &\quad \text{defektusa } 1. \end{aligned}$$

A defektusokat táblázatba foglaljuk:

	$p_k = d_{k1}$	d_{k2}	d_{k3}	$\sum_{\nu} d_{k\nu} = \alpha_k$
λ_1	3	2		$\alpha_1 = 5$
λ_2	2	1	1	$\alpha_2 = 4$
λ_3	1	1		$\alpha_3 = 2$
λ_4	1			$\alpha_4 = 1$
				$n = \sum_k \alpha_k = 12$
				* * *

A mátrixok Jordan-féle normálalakjának beható vizsgálata lehetővé teszi a következő két általános tétel bizonyítását.

3.5.3 tétel. *Két mátrix akkor és csak akkor hasonló, ha invariáns faktoraik – és így elemi osztóik is – megegyeznek.*

Bizonyítás. A 3.5.1 tételben bebizonyítottuk, hogy hasonló mátrixok determinánsosztói megegyeznek. A 3.5.2 definícióból következik, hogy akkor az ezek hányadosaiból képzett invariáns faktorok, tehát az elemi osztók is megegyeznek.

Megfordítva, tegyük fel, hogy az **A** és a **B** mátrix invariáns faktorai – és így elemi osztói is – megegyeznek. Mivel bármely mátrix hasonlósági transzformációval Jordan-féle normálalakra hozható, továbbá **A** és **B** invariáns faktorai megegyeznek, ezért megegyezik Jordan-féle normálalakjuk is. Tehát **A** és **B** ugyanahhoz a mátrixhoz hasonló, így **A** is hasonló **B**-hez. ■

3.5.4 tétel. *Adott lineáris transzformáció Jordan-féle normálalakját egyértelműen meghatározza.*

Bizonyítás. Adott lineáris transzformáció mátrixa egy – a bázis megválasztásától függő – hasonlósági transzformációtól eltekintve egyértelműen meghatározott. Mivel pedig hasonló mátrixok elemi osztói megegyeznek és ezek a Jordan-féle normálalakot egyértelműen meghatározzák, ezért az adott lineáris transzformáció Jordan-féle normálalakja egyértelműen meghatározott. ■

Minthogy hasonló mátrixok ugyanazt a lineáris transzformációt határozzák meg, így a bizonyított tételek alapján következik, hogy az invariáns faktorok és az elemi osztók – mint a bázis megválasztásától független mennyiségek – a lineáris transzformáció jellemző adatai.

3.6 LINEÁRIS DIFFERENCIÁLEGYENLET-RENDSZEREK

A következőkben a mátrixfüggvények legfőbb alkalmazási területével, a lineáris differenciálegyenlet-rendszerek elméletével és megoldási módszereivel

foglalkozunk. Feltételezzük, hogy az olvasó a differenciálegyenletekkel kapcsolatos alapvető ismereteknek birtokában van, ezért ezekre csupán utalunk. Azt kívánjuk megmutatni, hogy a mátrixelmélet milyen matematikai apparátust szolgáltat a lineáris differenciálegyenletek tárgyalásához. Elsőrendű lineáris differenciálegyenlet-rendszerekkel kezdjük vizsgálatainkat. Röviden ismertetjük a változó együtthatójú lineáris differenciálegyenlet-rendszereket, majd részletesen elemezzük az állandó együtthatójú rendszereket. Megmutatjuk, hogy utóbbiak megoldása mátrixokra vonatkozó sajátérték-feladatra vezet: az általános tárgyaláshoz a mátrixfüggvények definícióját és kanonikus előállítását használjuk fel.

Célul azt tűzzük ki, hogy az adott differenciálegyenletnek az előírt feltételeket kielégítő megoldását explicit alakban nyerjük. Ezt az teszi lehetővé, hogy az elsőrendű differenciálegyenlet-rendszer megoldásának feladata lényegében algebrai egyenletrendszer megoldását is magában foglalja.

Röviden foglalkozunk az alkalmazások szempontjából különösen fontos másodrendű rendszerekkel, és számos példa részletes kidolgozásával illusztráljuk a megoldási módszereket.

3.6.1 Explicit alakban megadott lineáris elsőrendű közönséges differenciálegyenlet-rendszerek

Ebben a pontban az alábbi alakú elsőrendű differenciálegyenlet-rendszer vizsgálatával foglalkozunk:

[illegible]

ahol $x_i = x_i(t)$ ($i = 1, 2, \dots, n$) jelöli a valós t változótól függő keresett függvényeket, $\dot{x}_i = \frac{dx_i(t)}{dt}$ ($i = 1, 2, \dots, n$) jelöli ezek t szerinti deriváltját, $f_i(t)$ és $a_{ij}(t)$ ($i, j = 1, 2, \dots, n$) pedig adott függvények, amelyekről feltesszük, hogy a t változó valamely T intervallumában mind folytonosak. (A t szerinti deriválást azért jelöljük itt ponttal, mert a vizsgált differenciálegyenlet-rendszerek legfontosabb alkalmazási területén az ismeretlen függvények pontrendszer mozgását határozzák meg, ahol a t változó rendszerint az időt jelenti és így paraméterként szerepel; az analízisben a paraméter szerinti deriválást általában ponttal jelöljük.)

Keressük a (3.6.1) rendszernek azt a megoldását, amely kielégíti a következő alakban adott kezdeti feltételeket:

$$(3.6.2) \quad x_1(t_0) = x_{10}, \quad x_2(t_0) = x_{20}, \quad \dots, \quad x_n(t_0) = x_{n0}.$$

Feltesszük tehát, hogy az egyenletek száma megegyezik az ismeretlen függvények számával és hogy az egyenletekből az első deriváltak egyértelműen kifejezhetők. A differenciálegyenletek elméletéből ismeretes, hogy ha t_0 benne van az $a_{ij}(t)$ és $f_i(t)$ függvények közös T folytonossági intervallumában, akkor a (3.6.1) egyenletrendszernek ebben az intervallumban egyetlen olyan megoldása van, amelyik kielégíti a (3.6.2) feltételeket. A következőkben megkíséreljük ennek a megoldásnak explicit alakban való előállítását. Látni fogjuk, hogy az általános esetben ezt csupán iterált integrálok konvergens végtelen sorával tudjuk kifejezni. Megvizsgáljuk azonban, hogy milyen feltételek mellett hozható ez a végtelen sor zárt alakra. Különösen egyszerűvé válik a megoldás az állandó együtthatójú rendszerek esetén.

Ahhoz, hogy a (3.6.1) differenciálegyenlet-rendszer vizsgálatára a mátrixelméletet alkalmazhassuk, néhány fogalmat kell bevezetni mátrixokra. Legyen $\mathbf{X}(t)$ egy olyan mátrix, amelynek $x_{ij}(t)$ elemei a t független változónak az $[a, b]$ zárt intervallumban differenciálható függvényei. Az $\mathbf{X}(t)$ mátrix t szerinti deriváltján ekkor az alábbi mátrixot értjük:

$$\frac{d\mathbf{X}(t)}{dt} = \left[\frac{dx_{ij}(t)}{dt} \right],$$

vagy rövidebb jelöléssel: $\dot{\mathbf{X}}(t) = [\dot{x}_{ij}(t)]$. Hasonlóképpen, ha $x_{ij}(t)$ primitív függvénye $y_{ij}(t)$ a $[t_0, t]$ intervallumban, azaz $\frac{dy_{ij}(t)}{dt} = x_{ij}(t)$, akkor az $\mathbf{X}(t)$ mátrix $[t_0, t]$ intervallumra vett határozott integrálján az alábbi mátrixot értjük:

$$\int_{t_0}^t \mathbf{X}(\tau) d\tau = \mathbf{Y}(t) - \mathbf{Y}(t_0) = [y_{ij}(t) - y_{ij}(t_0)].$$

Tehát röviden azt mondhatjuk, hogy mátrixot elemenként differenciálunk és integrálunk.

Tekintsük most a (3.6.1) differenciálegyenlet-rendszert a (3.6.2) kezdeti feltételekkel együtt, és vezessük be a következő jelöléseket. Legyen az ismeretlen függvényekből alkotott n elemű vektor \mathbf{x} , az $a_{ij}(t)$ együtthatófüggvényekből alkotott n -edrendű kvadratikus mátrix $\mathbf{A}(t)$, az $f_i(t)$ függvényekből alkotott n elemű vektor $\mathbf{f}(t)$ és az x_{i0} kezdeti értékekből alkotott n elemű vektor \mathbf{x}_0 . A (3.6.1) egyenlet és a (3.6.2) feltételek ezekkel a jelölésekkel az alábbi alakban írhatók fel:

$$(3.6.3) \quad \dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} + \mathbf{f}(t),$$

$$(3.6.4) \quad \mathbf{x}(t_0) = \mathbf{x}_0.$$

A (3.6.3) egyenletet $\mathbf{f}(t) \equiv 0$ esetén *homogén lineáris*, ellenkező esetben *inhomogén lineáris elsőrendű közönséges differenciálegyenlet-rendszernek* nevezzük. A (3.6.4) kezdeti feltételeket $\mathbf{x}_0 = 0$ esetén homogén, ellenkező esetben inhomogén kezdeti feltételeknek nevezzük.

Könnyen belátható a lineáris differenciálegyenletek elméletének azon alapvető tétele, amely szerint az inhomogén lineáris differenciálegyenlet általános megoldását az inhomogén egyenlet egy partikuláris megoldásának és a hozzá tartozó homogén egyenlet általános megoldásának az összege adja. Ha ugyanis \mathbf{x}_p a (3.6.3) egyenlet egy partikuláris megoldása, és az általános megoldást

$$(3.6.5) \quad \mathbf{x} = \boldsymbol{\xi} + \mathbf{x}_p$$

alakú összeg alakjában keressük, akkor behelyettesítve (3.6.5)-öt a (3.6.3) egyenletbe, azt kapjuk, hogy $\dot{\boldsymbol{\xi}} + \dot{\mathbf{x}}_p = \mathbf{A}(t)(\boldsymbol{\xi} + \mathbf{x}_p) + \mathbf{f}(t)$. A beszorzást elvégezve, és figyelembe véve, hogy \mathbf{x}_p kielégíti a (3.6.3) egyenletet, vagyis $\dot{\mathbf{x}}_p = \mathbf{A}(t)\mathbf{x}_p + \mathbf{f}(t)$, a $\boldsymbol{\xi}$ vektorra a következőt nyerjük:

$$(3.6.6) \quad \dot{\boldsymbol{\xi}} = \mathbf{A}(t)\boldsymbol{\xi}.$$

Tehát $\boldsymbol{\xi}$ a (3.6.3) egyenlethez tartozó homogén lineáris egyenlet általános megoldása. Ugyancsak a lineáris differenciálegyenletek általános elméletéből ismeretes, hogy a (3.6.6) alakú homogén lineáris differenciálegyenlet-rendszernek – az $a_{ij}(t)$ függvények közös folytonossági intervallumában – mindig van n lineárisan független megoldása, és ezek lineáris kombinációja adja az egyenlet általános megoldását. Az egyenlet n lineárisan független megoldásából alkotott rendszert *alaprendszernek* nevezzük. Ha az alaprendszert alkotó megoldásvektorokat úgy választjuk meg, hogy a $t = t_0$ helyen az elsőnek az első eleme 1, a másodiknak a második eleme 1, és így tovább, és valamennyi többi elem 0, akkor *normált alaprendszerrel* beszélünk. Jelölje $\mathbf{x}_1(t), \mathbf{x}_2(t), \dots, \mathbf{x}_n(t)$ a (3.6.6) egyenlet ilyen normált alaprendszert alkotó megoldásvektorait. Ha ezeket egy $\mathbf{X}(t)$ mátrix oszlopvektorainak tekintjük:

$$(3.6.7) \quad \mathbf{X}(t) = [\mathbf{x}_1(t) \quad \mathbf{x}_2(t) \dots \mathbf{x}_n(t)],$$

akkor a normált alaprendszer definíciójából következik, hogy

$$(3.6.8) \quad \mathbf{X}(t_0) = \mathbf{E}.$$

A normált alaprendszert alkotó vektorok (3.6.7) mátrixát a (3.6.3) egyenlet t_0 helyhez tartozó *rezolvensmátrixának* nevezzük. A (3.6.6) homogén lineáris egyenlet általános megoldását tehát a rezolvensmátrix oszlopainak lineáris kombinációja adja, vagyis az általános megoldás felírható mint

$$(3.6.9) \quad \boldsymbol{\xi}(t) = \mathbf{X}(t)\mathbf{c},$$

ahol a \mathbf{c} vektor elemei tetszőleges állandók.

3.6.2 A differenciálegyenlet-rendszer megoldása a rezolvensmátrix ismeretében

A (3.6.3) inhomogén lineáris differenciálegyenlet (3.6.4) inhomogén feltételeket kielégítő megoldását sokszor célszerű a következő megfontolások segítségével meghatározni. Meghatározzuk a (3.6.6) homogén lineáris egyenletnek azt a partikuláris megoldását, amely kielégíti a (3.6.4) inhomogén kezdeti feltételeket, a (3.6.3) inhomogén lineáris egyenletnek pedig azt a partikuláris megoldását, amely nem „rontja el” a már kielégített kezdeti feltételeket, vagyis amelynek minden eleme 0 a $t = t_0$ helyen, és ezt a két megoldást összegezzük:

inhomogén lineáris differenciálegyenlet inhomogén kezdeti feltételeket kielégítő megoldása =
= **homogén lineáris differenciálegyenlet inhomogén kezdeti feltételeket kielégítő megoldása** +
+ **inhomogén lineáris differenciálegyenlet homogén kezdeti feltételeket kielégítő megoldása**.

Ha tehát $\mathbf{x}(t; t_0, \mathbf{x}_0)$, ill. $\boldsymbol{\xi}(t; t_0, \mathbf{x}_0)$ jelöli az inhomogén, illetve homogén lineáris differenciálegyenletnek azt a megoldását, amely a t_0 helyen az \mathbf{x}_0 kezdeti értékeket veszi fel, akkor

$$(3.6.10) \quad \mathbf{x}(t; t_0, \mathbf{x}_0) = \boldsymbol{\xi}(t; t_0, \mathbf{x}_0) + \mathbf{x}(t; t_0, 0).$$

Mivel a (3.6.6) homogén lineáris egyenlet általános megoldását a rezolvensmátrix segítségével (3.6.9) alakban kaptuk, ezért ebből – (3.6.8) figyelembevételével – a (3.6.4) kezdeti feltételeket kielégítő partikuláris megoldásra $\boldsymbol{\xi}(t_0) = \mathbf{X}(t_0)\mathbf{c} = \mathbf{E}\mathbf{c} = \mathbf{x}_0$ adódik, vagyis $\mathbf{c} = \mathbf{x}_0$. Visszahelyettesítve ezt a (3.6.9) egyenletbe,

$$(3.6.11) \quad \boldsymbol{\xi}(t; t_0, \mathbf{x}_0) = \mathbf{X}(t)\mathbf{x}_0,$$

és ezzel megkaptuk a (3.6.10) megoldás első tagját – a homogén egyenletnek az inhomogén kezdeti feltételeket kielégítő partikuláris megoldását.

Az inhomogén lineáris differenciálegyenlet partikuláris megoldását az állandók variálásának a módszere szerint az

$$(3.6.12) \quad \mathbf{x}_p(t) = \mathbf{X}(t)\mathbf{c}(t)$$

szorzat alakjában keressük. Behelyettesítve ezt a (3.6.3) egyenletbe és figyelembe véve, hogy a szorzat deriválási szabálya mátrixokra is változatlanul érvényben van – azzal a megkötéssel, hogy a tényezők sorrendje nem felcserélhető –, az alábbi egyenletet kapjuk:

$$(3.6.13) \quad \dot{\mathbf{X}}(t)\mathbf{c}(t) + \mathbf{X}(t)\dot{\mathbf{c}}(t) = \mathbf{A}(t)\mathbf{X}(t)\mathbf{c}(t) + \mathbf{f}(t).$$

Mivel a rezolvensmátrix kielégíti a (3.6.6) homogén lineáris differenciálegyenletet (a definíciója értelmében ugyanis minden oszlopa a (3.6.6) egyenlet megoldása), ebből következik, hogy az egész mátrix is kielégíti a (3.6.6) egyenletet:

$$(3.6.14) \quad \dot{\mathbf{X}}(t) = \mathbf{A}(t)\mathbf{X}(t);$$

ezt behelyettesítve a (3.6.13) egyenletbe, összevonás után

$$(3.6.15) \quad \mathbf{X}(t)\dot{\mathbf{c}}(t) = \mathbf{f}(t)$$

adódik. A lineáris differenciálegyenletek elméletéből ismeretes, hogy ha a (3.6.6) egyenlet partikuláris megoldásai egy t_0 helyen lineárisan függetlenek, akkor az $a_{ij}(t)$ függvények közös folytonossági intervallumában mindenütt azok, ezért az $\mathbf{X}(t)$ rezolvensmátrix az egész vizsgált intervallumban nonszinguláris, tehát invertálható. Szorozzuk meg a (3.6.15) egyenletet balról a rezolvensmátrix inverzével: $\dot{\mathbf{c}}(t) = \mathbf{X}^{-1}(t)\mathbf{f}(t)$, és integráljunk tetszőleges t_1 alsó határtól a t felső határig:

$$\mathbf{c}(t) = \int_{t_1}^t \mathbf{X}^{-1}(\tau)\mathbf{f}(\tau) d\tau.$$

Visszahelyettesítve a (3.6.12) egyenletbe, megkapjuk a (3.6.3) inhomogén lineáris egyenlet egy partikuláris megoldását:

$$(3.6.16) \quad \mathbf{x}_p(t) = \mathbf{X}(t) \int_{t_1}^t \mathbf{X}^{-1}(\tau)\mathbf{f}(\tau) d\tau.$$

Mi azonban az inhomogén lineáris egyenletnek azt a partikuláris megoldását keressük, amely a homogén kezdeti feltételeket elégíti ki, vagyis azt a megoldásvektort, amelynek a $t = t_0$ helyen minden eleme 0. Ez elérhető, ha a (3.6.16) megoldásban az integrál t_1 alsó határául éppen a t_0 értéket választjuk. Így a keresett partikuláris megoldás:

$$(3.6.17) \quad \mathbf{x}(t; t_0, 0) = \mathbf{X}(t) \int_{t_0}^t \mathbf{X}^{-1}(\tau)\mathbf{f}(\tau) d\tau.$$

Ha tehát a (3.6.11) és a (3.6.17) kifejezést behelyettesítjük a (3.6.10) összefüggésbe, akkor a (3.6.3) egyenlet (3.6.4) feltételeket kielégítő megoldását kapjuk:

$$(3.6.18) \quad \mathbf{x}(t; t_0, \mathbf{x}_0) = \mathbf{X}(t)\mathbf{x}_0 + \mathbf{X}(t) \int_{t_0}^t \mathbf{X}^{-1}(\tau)\mathbf{f}(\tau) d\tau.$$

3.6.3 A rezolvensmátrix meghatározása

Mint látjuk, a (3.6.3) differenciálegyenlet megoldásának a kulcsa a rezolvensmátrix ismerete. Ennek meghatározása céljából a (3.6.3) egyenletet átírjuk a (3.6.4) kezdeti feltételeket is magában foglaló integrálegyenletté:

$$(3.6.19) \quad \mathbf{X}(t) = \mathbf{E} + \int_{t_0}^t \mathbf{A}(\tau) \mathbf{X}(\tau) d\tau.$$

Ennek megoldására a Picard*-féle szukcesszív approximáció módszerét alkalmazzuk. E szerint a (3.6.19) integrálegyenlet megoldása az

$$(3.6.20) \quad \mathbf{X}_{k+1}(t) = \mathbf{E} + \int_{t_0}^t \mathbf{A}(\tau) \mathbf{X}_k(\tau) d\tau; \quad \mathbf{X}_0 \equiv \mathbf{E} \quad (k = 0, 1, \dots)$$

rekurziós összefüggéssel definiált

$$(3.6.21) \quad \mathbf{X}_0(t), \quad \mathbf{X}_1(t), \quad \mathbf{X}_2(t), \dots$$

mátrixsorozat határértéke. A differenciálegyenletek elméletéből ismeretes, hogy az $a_{ij}(t)$ együtthatófüggvények közös folytonossági intervallumának belsőjében tetszőleges helyen felvett t_0 esetén ebben az intervallumban a (3.6.21) sorozat egyenletesen konvergens, és határértéke a (3.6.19) integrálegyenlet egyetlen megoldását adja. A (3.6.20) összefüggésből a (3.6.21) sorozat elemeire a következő kifejezéseket kapjuk:

$$\begin{aligned} \mathbf{X}_0(t) &= \mathbf{E}, \\ \mathbf{X}_1(t) &= \mathbf{E} + \int_{t_0}^t \mathbf{A}(\tau) d\tau, \\ \mathbf{X}_2(t) &= \mathbf{E} + \int_{t_0}^t \mathbf{A}(\tau_1) d\tau_1 + \int_{t_0}^t \int_{t_0}^{\tau_1} \mathbf{A}(\tau_1) \mathbf{A}(\tau_2) d\tau_2 d\tau_1, \\ &\dots\dots\dots \\ \mathbf{X}_k(t) &= \mathbf{E} + \int_{t_0}^t \mathbf{A}(\tau_1) d\tau_1 + \dots + \int_{t_0}^t \int_{t_0}^{\tau_1} \dots \int_{t_0}^{\tau_{k-1}} \mathbf{A}(\tau_1) \dots \mathbf{A}(\tau_k) d\tau_k \dots d\tau_1. \end{aligned}$$

*Ch. É. Picard (1856–1941) francia matematikus.

A keresett rezolvensmátrixra tehát a következő – iterált integrálokból álló – végtelen sort nyerjük:

$$(3.6.22) \quad \mathbf{X}(t) = \mathbf{E} + \int_{t_0}^t \mathbf{A}(\tau_1) d\tau_1 + \int_{t_0}^t \int_{t_0}^{\tau_1} \mathbf{A}(\tau_1) \mathbf{A}(\tau_2) d\tau_2 d\tau_1 + \dots$$

A rezolvensmátrix fenti előállítását *matrizánsnak* nevezik.

Mint látható, a rezolvensmátrix meghatározása ebben az általános esetben igen bonyolult eljárást kíván.

Az eljárást a következő – viszonylag egyszerű – példán mutatjuk be.

24. Példa. Határozzuk meg az

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 - tx_2 \end{aligned}$$

differenciálegyenlet-rendszernek az $x_1(0) = 1$, $x_2(0) = 0$ kezdeti feltételeket kielégítő megoldását!

Megoldás. Esetünkben $\mathbf{A}(t) = \begin{bmatrix} 0 & 1 \\ -1 & -t \end{bmatrix}$, tehát $\int_0^t \mathbf{A}(\tau) d\tau = \begin{bmatrix} 0 & t \\ -t & -\frac{t^2}{2} \end{bmatrix}$.

A (3.6.22) sor következő tagjai:

$$\begin{aligned} & \int_0^t \begin{bmatrix} 0 & 1 \\ -1 & -\tau \end{bmatrix} \begin{bmatrix} 0 & \tau \\ -\tau & -\frac{1}{2}\tau^2 \end{bmatrix} d\tau = \begin{bmatrix} -\frac{1}{2}t^2 & -\frac{1}{3!}t^3 \\ \frac{1}{3}t^3 & -\frac{1}{2}t^2 + \frac{1}{2 \cdot 4}t^4 \end{bmatrix}, \\ & \int_0^t \begin{bmatrix} 0 & 1 \\ -1 & -\tau \end{bmatrix} \begin{bmatrix} -\frac{1}{2}\tau^2 & -\frac{1}{3!}\tau^3 \\ \frac{1}{3}\tau^3 & -\frac{1}{2}\tau^2 + \frac{\tau^4}{2 \cdot 4} \end{bmatrix} d\tau = \\ & = \begin{bmatrix} \frac{1}{3 \cdot 4}t^4 & -\frac{1}{3!}t^3 + \frac{1}{2 \cdot 4 \cdot 5}t^5 \\ \frac{1}{2 \cdot 3}t^3 - \frac{1}{3 \cdot 5}t^5 & \frac{1}{4!}t^4 + \frac{1}{2 \cdot 4}t^4 - \frac{1}{2 \cdot 4 \cdot 6}t^6 \end{bmatrix}, \\ & \int_0^t \begin{bmatrix} 0 & 1 \\ -1 & -\tau \end{bmatrix} \begin{bmatrix} \frac{1}{3 \cdot 4}\tau^4 & -\frac{1}{2 \cdot 3}\tau^3 + \frac{1}{2 \cdot 4 \cdot 5}\tau^5 \\ \frac{1}{2 \cdot 3}\tau^3 - \frac{1}{3 \cdot 5}\tau^5 & \frac{1}{2 \cdot 3}\tau^4 - \frac{1}{2 \cdot 4 \cdot 6}\tau^6 \end{bmatrix} d\tau = \end{aligned}$$

$$= \begin{bmatrix} \frac{1}{2 \cdot 3 \cdot 4} t^4 - \frac{1}{3 \cdot 5 \cdot 6} t^6 & \frac{1}{2 \cdot 3 \cdot 5} t^5 - \frac{1}{2 \cdot 4 \cdot 6 \cdot 7} t^7 \\ -\frac{1}{3 \cdot 4 \cdot 5} t^5 - \frac{1}{2 \cdot 3 \cdot 5} t^5 + \frac{1}{3 \cdot 5 \cdot 7} t^7 & \frac{1}{4!} t^4 - \frac{23}{6!} t^6 + \frac{t^8}{2 \cdot 4 \cdot 6 \cdot 8} \end{bmatrix}$$

stb.

Tehát a rezolvensmátrix:

$$\mathbf{X}(t) = \begin{bmatrix} 1 - \frac{t^2}{2} + \frac{t^4}{2 \cdot 4} - \frac{t^6}{3 \cdot 5 \cdot 6} + \dots & t - \frac{t^3}{3} + \frac{7}{120} t^5 - \frac{t^7}{2 \cdot 4 \cdot 6 \cdot 7} + \dots \\ -t + \frac{t^3}{2} - \frac{7}{3 \cdot 4 \cdot 5} t^5 + \dots & 1 - t^2 + \frac{t^4}{3} - \frac{19}{360} t^6 + \dots \end{bmatrix}.$$

Ha ezt jobbról megszorozzuk az $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ vektorral, akkor a megoldásra azt kapjuk, hogy

$$\begin{aligned} x_1 &= 1 - \frac{t^2}{2} + \frac{t^4}{2 \cdot 4} - \frac{t^6}{3 \cdot 5 \cdot 6} + \dots, \\ x_2 &= -t + \frac{t^3}{2} - \frac{7}{3 \cdot 4 \cdot 5} t^5 + \dots \end{aligned}$$

Megjegyzés. Könnyen meggyőződhetünk arról, hogy a feladat megoldása $x_1 = e^{-\frac{t^2}{2}}$, $x_2 = -te^{-\frac{t^2}{2}}$; az iterációval kapott megoldás ezen függvények Taylor-sorát szolgáltatják. Mint látható, a fenti közelítés x_1 esetében három, x_2 esetében pedig csak két tagig egyezik a Taylor-sorokkal.

* * *

3.6.4 A rezolvensmátrix előállítása a felcserélhetőségi reláció teljesülése esetén

Lényegesen egyszerűbb a (3.6.3) differenciálegyenlet-rendszer megoldása abban az esetben, amikor az együtthatómátrixra teljesül a felcserélhetőségi reláció:

$$(3.6.23) \quad \mathbf{A}(t) \cdot \int_{t_0}^t \mathbf{A}(\tau) d\tau = \int_{t_0}^t \mathbf{A}(\tau) d\tau \cdot \mathbf{A}(t),$$

azaz ha az *együtthatómátrix felcserélhető az integráljával*. Ekkor ugyanis a rezolvensmátrixra kapott (3.6.22) alatti matrizáns egyes tagjai egyszerűbb alakra hozhatók.

Vezessük be a $\mathbf{B}(t) = \int_{t_0}^t \mathbf{A}(\tau) d\tau$ jelölést, ahonnan $\mathbf{A}(t) = \frac{d\mathbf{B}}{dt}$. Ezzel a (3.6.22) sor harmadik tagja a következő alakba írható:

$$(3.6.24) \quad \int_{t_0}^t \int_{t_0}^{\tau_1} \mathbf{A}(\tau_1) \mathbf{A}(\tau_2) d\tau_2 d\tau_1 = \int_{t_0}^t \frac{d\mathbf{B}}{d\tau_1} \mathbf{B}(\tau_1) d\tau_1.$$

Tekintettel arra, hogy a (3.6.23) reláció a $\mathbf{B}(t)$ mátrixra vonatkozóan azt jelenti, hogy az a *deriváltjával felcserélhető*, az adódik, hogy

$$(3.6.25) \quad \frac{d}{dt} \mathbf{B}^2(t) = \frac{d\mathbf{B}}{dt} \mathbf{B} + \mathbf{B} \frac{d\mathbf{B}}{dt} = 2 \frac{d\mathbf{B}}{dt} \mathbf{B}.$$

Ez az összefüggés általában természetesen nem érvényes, mivel általában $\mathbf{B}(t)$ és deriváltja nem felcserélhető.

Helyettesítsük be a (3.6.25) összefüggést (3.6.24)-be:

$$\int_{t_0}^t \frac{d\mathbf{B}}{d\tau_1} \mathbf{B}(\tau_1) d\tau_1 = \frac{1}{2} \int_{t_0}^t \frac{d}{d\tau_1} \mathbf{B}^2(\tau_1) d\tau_1 = \frac{1}{2} \mathbf{B}^2(t).$$

Hasonlóképpen adódnak a (3.6.22) sor következő tagjaira az alábbi kifejezések:

$$\frac{1}{3!} \mathbf{B}^3(t), \dots, \frac{1}{n!} \mathbf{B}^n(t), \dots,$$

tehát a (3.6.22) matrizáns a következő mátrixhatványsor alakjában írható:

$$\mathbf{X}(t) = \mathbf{E} + \mathbf{B}(t) + \frac{1}{2!} \mathbf{B}^2(t) + \dots + \frac{1}{n!} \mathbf{B}^n(t) + \dots$$

A mátrixfüggvény 3.3 szakaszban megadott definíciója értelmében ez a mátrixhatványsor a T intervallumban mindenütt konvergens, és az

$$\mathbf{X}(t) = e^{\int_{t_0}^t \mathbf{A}(\tau) d\tau}$$

exponenciális mátrixfüggvényt állítja elő. Belátható továbbá, hogy

$$(3.6.26) \quad \mathbf{X}(t) \mathbf{X}^{-1}(\tau) = e^{\int_{t_0}^t \mathbf{A}(u) du} e^{-\int_{t_0}^{\tau} \mathbf{A}(u) du} = e^{\int_{\tau}^t \mathbf{A}(u) du},$$

mivel a kitevők *kommutatívak*. Ugyanis

$$\int_{t_0}^t \mathbf{A}(u) du \cdot \int_{t_0}^{\tau} \mathbf{A}(v) dv = \int_{t_0}^t \int_{t_0}^{\tau} \mathbf{A}(u) \mathbf{A}(v) dv du =$$

$$= \int_{t_0}^{\tau} \int_{t_0}^t \mathbf{A}(u) \mathbf{A}(v) du dv = \int_{t_0}^{\tau} \left\{ \int_{t_0}^t \mathbf{A}(u) du \right\} \mathbf{A}(v) dv;$$

kihasználva itt a (3.6.23) felcserélhetőségi feltételt, az integranduszban álló tényezők felcserélésével a fenti összefüggésből

$$\int_{t_0}^{\tau} \mathbf{A}(v) \left\{ \int_{t_0}^t \mathbf{A}(u) du \right\} dv = \int_{t_0}^{\tau} \mathbf{A}(v) dv \cdot \int_{t_0}^t \mathbf{A}(u) du$$

adódik. Ebből következik, hogy a (3.6.26) összefüggésben a kitevők a szorzásra valóban kommutatívák, ezért az exponenciális függvények szorzatát képezhetjük úgy, hogy a kitevőket összeadjuk.

A (3.6.3) mátrix-differenciálegyenlet (3.6.18) megoldását tehát a következő alakban kapjuk:

$$(3.6.27) \quad \mathbf{x}(t; t_0, \mathbf{x}_0) = e^{\int_{t_0}^t \mathbf{A}(\tau) d\tau} \mathbf{x}_0 + \int_{t_0}^t e^{\int_{t_0}^{\tau} \mathbf{A}(u) du} \mathbf{f}(\tau) d\tau.$$

Láthatjuk innen, hogy a (3.6.3) egyenlet explicit alakú megoldásához exponenciális mátrixfüggvény előállítására van szükség.

A (3.6.23) felcserélhetőségi reláció teljesül minden olyan együtthatómátrixra, amely egy *állandó* mátrix függvényegyütthatós polinomja. A legegyszerűbb – és egyben legkönnyebben felismerhető – ilyen mátrixok a *ciklikus* mátrixok, amelyek az elemi ciklikus mátrix polinomjaként írhatók fel:

$$\mathbf{C}(c_0(t), c_1(t), \dots, c_{n-1}(t)) = \sum_{\nu=0}^{n-1} c_{\nu}(t) \mathbf{\Omega}^{\nu}.$$

Ennek $\varphi(\mathbf{C})$ függvénye a következő kanonikus alakra hozható (lásd a (3.2.44) és (3.2.52) képleteket):

$$\varphi(\mathbf{C}) = \frac{1}{n} \sum_{k=0}^{n-1} \varphi \left(\sum_{\nu=0}^{n-1} c_{\nu}(t) \omega_k^{\nu} \right) \begin{bmatrix} 1 \\ \omega_k \\ \omega_k^2 \\ \vdots \\ \omega_k^{n-1} \end{bmatrix} \begin{bmatrix} 1 & \bar{\omega}_k & \bar{\omega}_k^2 & \dots & \bar{\omega}_k^{n-1} \end{bmatrix},$$

ahol $\omega_k = e^{\frac{2k\pi i}{n}}$ ($k = 0, 1, \dots, n-1$). Ennek felhasználásával $\mathbf{A}(t) = \mathbf{C}(c_0(t), c_1(t), \dots, c_{n-1}(t))$ esetén a (3.6.3) egyenlet megoldásának tetszőleges komponense a következő alakban írható fel:

$$(3.6.28) \quad x_l(t) = \frac{1}{n} \sum_{k=0}^{n-1} \omega_k^{l-1} \left\{ e^{\int_0^t \sum_{\nu=0}^{n-1} c_\nu(\tau) \omega_k^\nu d\tau} \sum_{\nu=0}^{n-1} \bar{\omega}_k^\nu x_{\nu+1,0} + \right. \\ \left. + \int_{t_0}^t e^{\int_\tau^t \sum_{\nu=0}^{n-1} c_\nu(u) \omega_k^\nu du} \sum_{\nu=0}^{n-1} \bar{\omega}_k^\nu f_{\nu+1}(\tau) d\tau \right\} \quad (l = 1, 2, \dots, n).$$

3.6.5 Állandó együtthatómátrixú differenciálegyenlet-rendszerek megoldása

Abban a legegyszerűbb esetben, amikor az \mathbf{A} együtthatómátrix elemei állandók, tehát a (3.6.23) felcserélhetőségi reláció triviálisan teljesül, a rezolvensmátrix a következő: $\mathbf{X}(t) = e^{\mathbf{A}(t-t_0)}$. Az állandó együtthatójú differenciálegyenlet-rendszerek megoldására ekkor azt kapjuk, hogy

$$(3.6.29) \quad \mathbf{x}(t; t_0, \mathbf{x}_0) = e^{\mathbf{A}(t-t_0)} \mathbf{x}_0 + \int_{t_0}^t e^{\mathbf{A}(t-\tau)} \mathbf{f}(\tau) d\tau.$$

Az itt előforduló exponenciális mátrixfüggvények kanonikus felbontását aszerint kell elvégezni, hogy az \mathbf{A} mátrix minimálegyenletének gyökei egyszeresek-e vagy sem (lásd a 3.2 és a 3.4 szakaszt).

(a) *Ha az \mathbf{A} minimálegyenletének csupa egyszeres gyöke van, akkor a Lagrange-féle mátrixpolinomok alkalmazásával a (3.6.29) megoldás alakja a következő:*

$$(3.6.30) \quad \mathbf{x}(t; t_0, \mathbf{x}_0) = \sum_{k=1}^s e^{\lambda_k(t-t_0)} L_k(\mathbf{A}) \mathbf{x}_0 + \sum_{k=1}^s \int_{t_0}^t e^{\lambda_k(t-\tau)} L_k(\mathbf{A}) \mathbf{f}(\tau) d\tau.$$

Emlékeztetünk arra, hogy az $L_k(\mathbf{A})$ mátrixpolinomok kifejezhetők a karakterisztikus mátrix adjungáltja segítségével (lásd (3.2.53)):

$$(3.6.31) \quad L_k(\mathbf{A}) = \frac{1}{\Delta'(\lambda_k)} \left(\frac{\text{adj}(\lambda \mathbf{E} - \mathbf{A})}{\theta(\lambda)} \right)_{\lambda=\lambda_k},$$

és hogy az $L_k(\mathbf{A})$ mátrixpolinomok minimális diadikus felbontása a λ_k sajátértékhez tartozó $\mathbf{u}_{\nu k}$ jobb oldali és $\mathbf{v}_{\nu k}^\top$ bal oldali sajátvektorokat szolgáltatja:

$$(3.6.32) \quad L_k(\mathbf{A}) = \sum_{\nu=1}^{\alpha_k} \mathbf{u}_{\nu k} \mathbf{v}_{\nu k}^\top \quad (k = 1, 2, \dots, s),$$

ahol α_k a λ_k sajátérték multiplicitása. Egyszerűbb írásmód kedvéért lássuk el most a sajátértékeket – és ennek megfelelően a sajátvektorokat is – 1-től n -ig futó l indexszel, megjegyezve, hogy a λ_l sajátértékek között itt egyenlők is lehetnek. A (3.6.30) megoldás tehát felírható a következő alakban:

$$(3.6.33) \quad \mathbf{x}(t; t_0, \mathbf{x}_0) = \sum_{l=1}^n \left\{ e^{\lambda_l(t-t_0)} \mathbf{u}_l \mathbf{v}_l^T \mathbf{x}_0 + \int_{t_0}^t e^{\lambda_l(t-\tau)} \mathbf{u}_l \mathbf{v}_l^T \mathbf{f}(\tau) d\tau \right\}.$$

Az ismeretlen függvények ebből adódóan:

$$(3.6.34) \quad x_j(t; t_0, \mathbf{x}_0) = \sum_{l=1}^n u_{jl} \left\{ e^{\lambda_l(t-t_0)} (\mathbf{v}_l^T \mathbf{x}_0) + \int_{t_0}^t e^{\lambda_l(t-\tau)} (\mathbf{v}_l^T \mathbf{f}(\tau)) d\tau \right\} \\ (j = 1, 2, \dots, n).$$

(b) Ha az \mathbf{A} minimálegyenletének többszörös gyöke is van, akkor az Hermite-féle mátrixpolinomokat alkalmazzuk. A (3.4.7) képlet szerint ezek együtthatói az itt szereplő függvény deriváltjainak a λ_k helyen felvett értékei. Hangsúlyoznunk kell, hogy a függvény λ szerinti deriváltjait kell képezni! Mivel a (3.6.3) differenciálegyenlet megoldásában fellépő függvények $e^{\mathbf{A}(t-t_0)}$, ill. $e^{\mathbf{A}(t-\tau)}$, ezért most az $e^{\lambda(t-t_0)}$, ill. $e^{\lambda(t-\tau)}$ függvényeket kell λ szerint deriválni. (A mátrixfüggvények előállítás szempontjából a differenciálegyenlet független változója – vagyis t – csupán paraméter!) Így ebben az általános esetben a megoldásban az $e^{\lambda_k(t-t_0)}$ alakú függvények mellett $(t-t_0)e^{\lambda_k(t-t_0)}$, $(t-t_0)^2 e^{\lambda_k(t-t_0)}$ stb. alakú függvények is fellépnek. A (3.6.29) megoldás alakja tehát a következő:

$$(3.6.35) \quad \mathbf{x}(t; t_0, \mathbf{x}_0) = \sum_{k=1}^s e^{\lambda_k(t-t_0)} \sum_{\nu=0}^{\gamma_k-1} (t-t_0)^\nu H_{k\nu}(\mathbf{A}) \mathbf{x}_0 + \\ + \sum_{k=1}^s \int_{t_0}^t e^{\lambda_k(t-\tau)} \sum_{\nu=0}^{\gamma_k-1} (t-\tau)^\nu H_{k\nu}(\mathbf{A}) \mathbf{f}(\tau) d\tau,$$

ahol γ_k a λ_k gyök multiplicitása a minimálegyenletben.

Látjuk tehát, hogy az állandó együtthatójú lineáris differenciálegyenlet-rendszerek megoldásait egyszerű struktúrájú együtthatómátrixok esetén kizárólag exponenciális függvények alkotják, nem egyszerű struktúrájú együtthatómátrixok esetében azonban fellépnek polinommal szorzott exponenciális függvények is.

A fentiek megvilágítására bemutatunk néhány példát.

25. Példa. Határozzuk meg az

$$\dot{x}_1 = x_1 \operatorname{sh} t + x_2 \operatorname{ch} t$$

$$\dot{x}_2 = x_1 \operatorname{ch} t + x_2 \operatorname{sh} t$$

differentiálegyenlet-rendszernek az $\mathbf{x}(0) = \mathbf{x}_0 = \begin{bmatrix} x_{10} \\ x_{20} \end{bmatrix}$ kezdeti feltételeket kielégítő megoldását!

Megoldás. E differenciálegyenlet-rendszer változó együtthatómátrixa:

$$\mathbf{A}(t) = \begin{bmatrix} \operatorname{sh} t & \operatorname{ch} t \\ \operatorname{ch} t & \operatorname{sh} t \end{bmatrix} = \mathbf{C}(\operatorname{sh} t, \operatorname{ch} t);$$

az együtthatómátrix ciklikus, és így alkalmazható a (3.6.27), ill. a (3.6.28) megoldási formula.

A megoldás menete a következő. Mivel

$$\int_0^t \mathbf{A}(\tau) d\tau = \begin{bmatrix} \operatorname{ch} t - 1 & \operatorname{sh} t \\ \operatorname{sh} t & \operatorname{ch} t - 1 \end{bmatrix},$$

és $n = 2$ miatt $\omega_0 = 1$, $\omega_1 = -1$, továbbá esetünkben $c_0 = \operatorname{ch} t - 1$ és $c_1 = \operatorname{sh} t$, tehát

$$\begin{aligned} e^{\int_0^t \mathbf{A}(\tau) d\tau} &= e^{c_0 + c_1 \omega_0} \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} + e^{c_0 + c_1 \omega_1} \frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} = \\ &= \frac{1}{2} e^{\operatorname{ch} t - 1 + \operatorname{sh} t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} + \frac{1}{2} e^{\operatorname{ch} t - 1 - \operatorname{sh} t} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} = \\ &= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} e^{e^t - 1} & 0 \\ 0 & \frac{1}{2} e^{e^{-t} - 1} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \end{aligned}$$

Ezzel a keresett megoldás:

$$x_1(t) = \frac{1}{2} e^{e^t - 1} (x_{10} + x_{20}) + \frac{1}{2} e^{e^{-t} - 1} (x_{10} - x_{20})$$

$$x_2(t) = \frac{1}{2} e^{e^t - 1} (x_{10} + x_{20}) - \frac{1}{2} e^{e^{-t} - 1} (x_{10} - x_{20}).$$

* * *

26. Példa. Határozzuk meg az

$$\dot{x}_1 = -3x_1 + 2x_2 - 5x_3$$

$$\dot{x}_2 = -6x_1 - 5x_2 - 30x_3$$

$$\dot{x}_3 = x_1 + 3x_3$$

differenciálegyenlet-rendszernek az $\mathbf{x}(0) = \begin{bmatrix} x_{10} \\ x_{20} \\ x_{30} \end{bmatrix}$ kezdeti feltételeket kielégítő megoldását!

Megoldás. Tekintettel arra, hogy állandó együtthatójú rendszerről van szó, a (3.6.29) megoldási formulát alkalmazhatjuk. Ehhez szükségünk van az $e^{\mathbf{A}t}$ mátrixfüggvény kanonikus felbontására, ahol

$$\mathbf{A} = \begin{bmatrix} -3 & 2 & -5 \\ -6 & -5 & -30 \\ 1 & 0 & 3 \end{bmatrix}.$$

A karakterisztikus egyenletre egyszerű számolással adódik

$$|\lambda \mathbf{E} - \mathbf{A}| \equiv \lambda^3 + 5\lambda^2 + 8\lambda + 4 = 0,$$

innen

$$\begin{aligned} \lambda_{1,2} &= -2 \quad (\text{kétszeres gyök}), \\ \lambda_3 &= -1. \end{aligned}$$

Most a minimálpolinomot kell meghatározni. Képezzük ezért a karakterisztikus mátrix másodrendű minorainak (vagyis az adjungált mátrix elemeinek) legnagyobb közös osztóját. Az adjungált mátrix:

$$\text{adj}(\lambda \mathbf{E} - \mathbf{A}) = \begin{bmatrix} (\lambda - 3)(\lambda + 5) & 2(\lambda - 3) & -5(\lambda + 17) \\ -6(\lambda + 2) & (\lambda + 2)(\lambda - 2) & -30(\lambda + 2) \\ 5 + \lambda & 2 & \lambda^2 + 8\lambda + 27 \end{bmatrix},$$

tehát az elemek legnagyobb közös osztója 1; ez azt jelenti, hogy a karakterisztikus polinom és a $\Delta(\lambda)$ minimálpolinom megegyezik: $\Delta(\lambda) \equiv (\lambda + 2)^2(\lambda + 1)$.

Megjegyzés. Mivel ebben a feladatban a karakterisztikus polinomnak egyetlen többszörös gyöke $\lambda_{1,2} = -2$ (kétszeres gyök), ezért itt egyszerűbben is meggyőződhetünk arról, hogy a karakterisztikus polinom és a minimálpolinom megegyezik. Ehhez elég azt figyelembe venni, hogy $\lambda_{1,2} = -2$ helyettesítésével a

$$-2\mathbf{E} - \mathbf{A} = \begin{bmatrix} 1 & -2 & 5 \\ 6 & 3 & 30 \\ -1 & 0 & -5 \end{bmatrix}$$

mátrix rangja 2, tehát a kétszeres sajátértékhez egyetlen sajátvektor, vagyis egyetlen másodrendű Jordan-blokk tartozik.

Mivel a minimálpolinomnak van többszörös gyöke, a kanonikus felbontáshoz az Hermite-féle mátrixpolinomokat használjuk fel. Ezek meghatározására kétféle módszert mutatunk be.

(a) Meghatározzuk a (3.4.3) definíció alapján a $H_{k\nu}(z)$ alappolinomokat, azután a z skalár változó helyére az \mathbf{A} mátrixot helyettesítjük. Esetünkben $\lambda_1 = -2$, $\lambda_2 = -1$. A $H_{10}(z)$ alappolinomot a következő alakban keressük:

$$H_{10}(z) = (Az + B)(z + 1).$$

Az A és B együtthatókat abból a feltételből határozzuk meg, hogy $H_{10}(-2) = 1$ és $H'_{10}(-2) = 0$:

$$2A - B = 1$$

$$-3A + B = 0;$$

ebből $A = -1$ és $B = -3$. Tehát $H_{10}(z) = -(z + 3)(z + 1)$. A második alappolinomot a $H_{11}(z) = C(z + 2)(z + 1)$ alakban keressük. A $H'_{11}(-2) = 1$ feltételből $C = -1$ adódik, tehát $H_{11}(z) = -(z + 2)(z + 1)$.

Végül a harmadik alappolinomot $H_{20}(z) = D(z + 2)^2$ alakban keressük, a D együtthatóra pedig a $H_{20}(-1) = 1$ feltételből $D = 1$ adódik. Tehát $H_{20}(z) = (z + 2)^2$. A fentiek alapján a keresett mátrixpolinomokat könnyen felírhatjuk:

$$H_{10}(\mathbf{A}) = -(\mathbf{A} + 3\mathbf{E})(\mathbf{A} + \mathbf{E}) = \begin{bmatrix} 17 & 8 & 80 \\ 6 & 4 & 30 \\ -4 & -2 & -19 \end{bmatrix},$$

$$H_{11}(\mathbf{A}) = -(\mathbf{A} + 2\mathbf{E})(\mathbf{A} + \mathbf{E}) = \begin{bmatrix} 15 & 10 & 75 \\ 0 & 0 & 0 \\ -3 & -2 & -15 \end{bmatrix},$$

$$H_{20}(\mathbf{A}) = -(\mathbf{A} + 2\mathbf{E})^2 = \begin{bmatrix} -16 & -8 & -80 \\ -6 & -3 & -30 \\ 4 & 2 & 20 \end{bmatrix}.$$

(b) A $H_{k\nu}(\mathbf{A})$ mátrixpolinomokat meghatározhatjuk közvetlenül is, a (3.4.18) egyenletek alapján. Tekintettel arra, hogy esetünkben három ilyen alappolinom szerepel, a következő egyenleteket írhatjuk fel:

$$H_{10}(\mathbf{A}) + H_{20}(\mathbf{A}) = \mathbf{E},$$

$$-2H_{10}(\mathbf{A}) + H_{11}(\mathbf{A}) - H_{20}(\mathbf{A}) = \mathbf{A},$$

$$2H_{10}(\mathbf{A}) - 2H_{11}(\mathbf{A}) + \frac{1}{2}H_{20}(\mathbf{A}) = \frac{1}{2}\mathbf{A}^2.$$

Az így nyert egyenletrendszer megoldására természetesen ugyanazt az eredményt kapjuk, mint az (a) esetben.

Az $e^{\mathbf{A}t}$ mátrixfüggvény ezek alapján az alábbi alakban írható fel:

$$e^{\mathbf{A}t} = e^{-2t}H_{10}(\mathbf{A}) + te^{-2t}H_{11}(\mathbf{A}) + e^{-t}H_{20}(\mathbf{A});$$

ezt jobbról szorozva a kezdeti értékekből alkotott \mathbf{x}_0 vektorral, az adott differenciálegyenlet-rendszer keresett megoldását kapjuk:

$$\begin{aligned}x_1(t) &= e^{-2t} \{ (17 + 15t)x_{10} + (8 + 10t)x_{20} + (80 + 75t)x_{30} \} - \\ &\quad - 8e^{-t}(2x_{10} + x_{20} + 10x_{30}), \\x_2(t) &= 2e^{-2t}(3x_{10} + 2x_{20} + 15x_{30}) - 3e^{-t}(2x_{10} + x_{20} + 10x_{30}), \\x_3(t) &= e^{-2t} \{ (-4 - 3t)x_{10} + (-2 - 2t)x_{20} + (-19 - 15t)x_{30} \} + \\ &\quad + 2e^{-t}(2x_{10} + x_{20} + 10x_{30}).\end{aligned}$$

E példa kapcsán ismételten felhívjuk a figyelmet arra, hogy az $e^{\mathbf{A}t}$ mátrixfüggvény meghatározásához alkalmazott (3.4.7) képletben szereplő $f'(\lambda)$ derivált λ szerinti deriválást jelent, tehát esetünkben az $e^{\lambda t}$ függvénynek λ szerinti deriváltja $te^{\lambda t}$. (Ha a minimálegyenlet gyökének a multiplicitása nagyobb, mint 2, akkor ehhez hasonlóan fellépnének a $t^2e^{\lambda t}$, $t^3e^{\lambda t}$ stb. kifejezések is.)

* * *

27. Példa. Határozzuk meg az

$$\begin{aligned}\dot{x}_1 &= 3x_1 - 3x_2 + 2x_3 + \sin 2t \\ \dot{x}_2 &= -x_1 + 5x_2 - 2x_3 + t \\ \dot{x}_3 &= -x_1 + 3x_2 + e^{2t}\end{aligned}$$

differenciálegyenlet-rendszernek az

$$\mathbf{x}(0) = \begin{bmatrix} x_{10} \\ x_{20} \\ x_{30} \end{bmatrix} = \mathbf{0}$$

homogén kezdeti feltételeket kielégítő megoldását!

Megoldás. A feladat megoldására a (3.6.29) formulát alkalmazhatjuk, esetünkben $\mathbf{x}_0 = \mathbf{0}$. Szükségünk van az $e^{\mathbf{A}(t-\tau)}$ mátrixfüggvény kanonikus felbontására, ahol

$$\mathbf{A} = \begin{bmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{bmatrix}.$$

A karakterisztikus egyenlet: $|\lambda \mathbf{E} - \mathbf{A}| \equiv (\lambda - 4)(\lambda - 2)^2 = 0$. Látható, hogy $\lambda = 2$ kétszeres gyöke a karakterisztikus egyenletnek, meg kell tehát határozni a minimálegyenletet is. Képezzük e célból a karakterisztikus mátrix adjungáltját:

$$\text{adj}(\lambda \mathbf{E} - \mathbf{A}) = \begin{bmatrix} (\lambda - 2)(\lambda - 3) & -3(\lambda - 2) & 2(\lambda - 2) \\ -(\lambda - 2) & (\lambda - 1)(\lambda - 2) & -2(\lambda - 2) \\ -(\lambda - 2) & 3(\lambda - 2) & (\lambda - 2)(\lambda - 6) \end{bmatrix}.$$

Innen kiolvasható, hogy az adjungált mátrix elemeinek – azaz a karakterisztikus mátrix másodrendű minorainak – legnagyobb közös osztója $\lambda - 2$; a minimálegyenlet tehát $\Delta(\lambda) \equiv (\lambda - 2)(\lambda - 4) = 0$.

Megjegyzés. Az előző példához hasonlóan itt is egyszerűbben megállapítható, hogy a minimálpolinom gyökei mind egyszeresek, mivel a karakterisztikus polinom egyetlen többszörös gyöke $\lambda = 2$ (kétszeres gyök). A $\lambda \mathbf{E} - \mathbf{A}$ karakterisztikus mátrix $\lambda = 2$ helyettesítésével

$$2\mathbf{E} - \mathbf{A} = \begin{bmatrix} -1 & 3 & -2 \\ +1 & -3 & 2 \\ +1 & -3 & 2 \end{bmatrix},$$

ennek rangja nyilvánvalóan 1, azaz a $\lambda = 2$ sajátértékhez két, lineárisan független sajátvektor tartozik. Az adott mátrixnak van tehát teljes sajátvektorrendszere, azaz diagonalizálható, és a minimálpolinomjának csupa egyszeres gyöke van.

Mivel a minimálpolinom gyökei egyszeresek, a kanonikus felbontáshoz a gyökhelyekhez tartozó Lagrange-féle alappolinomokat használjuk fel. Ezek, mivel

$$L_1(2) = 1, \quad L_2(2) = 0, \quad L_1(4) = 0, \quad L_2(4) = 1,$$

a következő lineáris függvények:

$$L_1(z) = -\frac{1}{2}(z - 4), \quad L_2(z) = \frac{1}{2}(z - 2).$$

A mátrixpolinomok pedig (a diadikus felbontást is elvégezve):

$$\begin{aligned} L_1(\mathbf{A}) &= \frac{1}{2}(4\mathbf{E} - \mathbf{A}) = \frac{1}{2} \begin{bmatrix} 1 & 3 & -2 \\ 1 & -1 & 2 \\ 1 & -3 & 4 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 3 & -2 \\ 0 & -2 & 2 \end{bmatrix}, \\ L_2(\mathbf{A}) &= \frac{1}{2}(\mathbf{A} - 2\mathbf{E}) = \frac{1}{2} \begin{bmatrix} 1 & -3 & 2 \\ -1 & 3 & -2 \\ -1 & 3 & -2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & -3 & 2 \end{bmatrix}. \end{aligned}$$

Ezzel $e^{\mathbf{A}(t-\tau)}$ kanonikus felbontása:

$$e^{\mathbf{A}(t-\tau)} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & -1 \\ 1 & 3 & -1 \end{bmatrix} \left\langle \frac{1}{2} e^{2(t-\tau)} \quad \frac{1}{2} e^{2(t-\tau)} \quad \frac{1}{2} e^{4(t-\tau)} \right\rangle \begin{bmatrix} 1 & 3 & -2 \\ 0 & -2 & 2 \\ 1 & -3 & 2 \end{bmatrix}.$$

A keresett megoldást tehát az alábbi integrál szolgáltatja:

$$\mathbf{x}(t) = \int_0^t e^{\mathbf{A}(t-\tau)} \mathbf{f}(\tau) d\tau = \frac{1}{2} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & -1 \\ 1 & 3 & -1 \end{bmatrix} \int_0^t \begin{bmatrix} e^{2(t-\tau)} \{\sin 2\tau + 3\tau - 2e^{2\tau}\} \\ e^{2(t-\tau)} \{-2\tau + 2e^{2\tau}\} \\ e^{4(t-\tau)} \{\sin 2\tau - 3\tau + 2e^{2\tau}\} \end{bmatrix} d\tau.$$

Tekintetbe véve, hogy

$$\int_0^t e^{\lambda_k(t-\tau)} \sin 2\tau \, d\tau = \frac{1}{4 + \lambda_k^2} \{2e^{\lambda_k t} - 2 \cos 2t - \lambda_k \sin 2t\},$$

$$\int_0^t \tau e^{\lambda_k(t-\tau)} \, d\tau = \frac{1}{\lambda_k} \left(t + \frac{e^{\lambda_k t} - 1}{\lambda_k} \right),$$

$$\int_0^t e^{2\tau} e^{\lambda_k(t-\tau)} \, d\tau = \begin{cases} \frac{1}{2 - \lambda_k} (e^{2t} - e^{\lambda_k t}), & \text{ha } \lambda_k \neq 2 \\ t e^{\lambda_k t}, & \text{ha } \lambda_k = 2 \end{cases},$$

behelyettesítés és összevonás után a megoldást a következő alakban kapjuk:

$$\begin{aligned} x_1(t) &= -t e^{2t} + \frac{73}{160} e^{4t} - \frac{9}{32} + \frac{3}{8} t - \frac{7}{40} \cos 2t - \frac{9}{40} \sin 2t, \\ x_2(t) &= \left(t + \frac{1}{2} \right) e^{2t} - \frac{73}{160} e^{4t} + \frac{1}{32} + \frac{1}{8} t - \frac{3}{40} \cos 2t - \frac{1}{40} \sin 2t, \\ x_3(t) &= \left(2t + \frac{1}{4} \right) e^{2t} - \frac{73}{160} e^{4t} + \frac{9}{32} - \frac{3}{8} t - \frac{3}{40} \cos 2t - \frac{1}{40} \sin 2t. \end{aligned}$$

* * *

28. Példa. Határozzuk meg az

$$\begin{aligned} \dot{x}_1 &= 4x_1 + 2x_2 + x_3 + e^{3t} \\ \dot{x}_2 &= x_1 + 5x_2 + 4x_3 + 3e^{3t} \\ \dot{x}_3 &= -x_1 - 2x_2 - 2e^{3t} \end{aligned}$$

differenciálegyenlet-rendszernek az $\mathbf{x}(0) = \begin{bmatrix} x_{10} \\ x_{20} \\ x_{30} \end{bmatrix}$ kezdeti feltételeket kielégítő megoldását!

Megoldás. Az egyenlet együtthatómátrixa $\mathbf{A} = \begin{bmatrix} 4 & 2 & 1 \\ 1 & 5 & 4 \\ -1 & -2 & 0 \end{bmatrix}$, és az $\mathbf{f}(t)$

oszlopvektor $\mathbf{f}(t) = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} e^{3t}$. A karakterisztikus egyenlet:

$$|\lambda \mathbf{E} - \mathbf{A}| = (\lambda - 3)^3 = 0,$$

tehát $\lambda = 3$ háromszoros sajátérték. A karakterisztikus mátrix adjungáltja

$$\text{adj}(\lambda \mathbf{E} - \mathbf{A}) = \begin{bmatrix} \lambda^2 - 5\lambda + 8 & 2(\lambda - 1) & \lambda + 3 \\ \lambda - 4 & \lambda^2 - 4\lambda + 1 & 4\lambda - 15 \\ -(\lambda - 3) & -2(\lambda - 3) & (\lambda - 6)(\lambda - 3) \end{bmatrix},$$

elemeinek legnagyobb közös osztója 1, tehát a minimálegyenlet megegyezik a karakterisztikus egyenlettel: $\Delta(\lambda) \equiv (\lambda - 3)^3$.

Megjegyzés. A 26. és 27. példához hasonlóan a minimálpolinom itt is egyszerűbben meghatározható. Elegendő azt figyelembe venni, hogy ha a karakterisztikus mátrixba a $\lambda = 3$ háromszoros gyököt helyettesítjük, akkor a

$$3\mathbf{E} - \mathbf{A} = \begin{bmatrix} -1 & -2 & -1 \\ -1 & -2 & -4 \\ 1 & 2 & 3 \end{bmatrix}$$

mátrix rangja 2, tehát a $\lambda = 3$ sajátértékhez egyetlen sajátvektor, és így egyetlen harmadrendű Jordan-blokk tartozik, vagyis a minimálpolinom megegyezik a karakterisztikus polinommal.

A minimálpolinom meghatározásának ez a módja – ugyanúgy, mint a 26. és 27. példában – igen egyszerű, azonban csak kis rendszámú mátrixok (és főként kis multiplicitású sajátértékek) esetén célszerű alkalmazni.

Mivel a minimálegyenletnek többszörös gyöke van, az Hermite-féle alapolinomokat alkalmazzuk. Esetünkben, tekintettel arra, hogy a minimálpolinomnak egyetlen (háromszoros) gyöke van, a (3.4.11) egyenletből $s = 1$ miatt $H_{10}(\mathbf{A}) = \mathbf{E}$ adódik. A mátrixfüggvény felbontásának (3.4.57) alakját felhasználva, a következő összefüggésre jutunk:

$$e^{\mathbf{A}t} = e^{3t}\mathbf{E} + te^{3t}(\mathbf{A} - 3\mathbf{E}) + \frac{t^2}{2}e^{3t}(\mathbf{A} - 3\mathbf{E})^2.$$

Mivel

$$(\mathbf{A} - 3\mathbf{E})^2 = \begin{bmatrix} 2 & 4 & 6 \\ -1 & -2 & -3 \\ 0 & 0 & 0 \end{bmatrix},$$

azt kapjuk, hogy

$$e^{\mathbf{A}t} = e^{3t} \begin{bmatrix} 1 + t + t^2 & 2t + 2t^2 & t + 3t^2 \\ t - \frac{t^2}{2} & 1 + 2t - t^2 & 4t - \frac{3}{2}t^2 \\ -t & -2t & 1 - 3t \end{bmatrix}.$$

Hátra van még az alábbi integrál kiszámítása:

$$\int_0^t e^{\mathbf{A}(t-\tau)} \mathbf{f}(\tau) d\tau.$$

Behelyettesítve eddig kapott összefüggéseinket, az alábbi eredményre jutunk:

$$\begin{aligned} \int_0^t e^{3(t-\tau)} \begin{bmatrix} 1+(t-\tau)+(t-\tau)^2 & 2(t-\tau)+2(t-\tau)^2 & (t-\tau)+3(t-\tau)^2 \\ (t-\tau)-\frac{1}{2}(t-\tau)^2 & 1+2(t-\tau)-(t-\tau)^2 & 4(t-\tau)-\frac{3}{2}(t-\tau)^2 \\ -(t-\tau) & -2(t-\tau) & 1-3(t-\tau) \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} e^{3\tau} d\tau = \\ = e^{3t} \int_0^t \begin{bmatrix} 1+5(t-\tau)+(t-\tau)^2 \\ 3-(t-\tau)-\frac{1}{2}(t-\tau)^2 \\ -2-(t-\tau) \end{bmatrix} d\tau = e^{3t} \begin{bmatrix} t+\frac{5}{2}t^2+\frac{1}{3}t^3 \\ 3t-\frac{1}{2}t^2-\frac{1}{6}t^3 \\ -2t-\frac{1}{2}t^2 \end{bmatrix}. \end{aligned}$$

A keresett megoldás tehát

$$\begin{aligned} x_1(t) &= e^{3t} \left\{ (1+t+t^2)x_{10} + 2t(1+t)x_{20} + t(1+3t)x_{30} + t + \frac{5}{2}t^2 + \frac{1}{3}t^3 \right\}, \\ x_2(t) &= e^{3t} \left\{ t\left(1-\frac{t}{2}\right)x_{10} + (1+2t-t^2)x_{20} + t\left(4-\frac{3}{2}t\right)x_{30} + 3t - \frac{1}{2}t^2 - \frac{1}{6}t^3 \right\}, \\ x_3(t) &= e^{3t} \left\{ -tx_{10} - 2tx_{20} + (1-3t)x_{30} - 2t - \frac{1}{2}t^2 \right\}. \end{aligned}$$

* * *

3.6.6 Elsőrendű közönséges differenciálegyenlet-rendszer periodikus megoldása

A következő feladatban az

$$(3.6.36) \quad \dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{f}(t)$$

differenciálegyenlet-rendszernek a szokásostól eltérő feltételeket kielégítő megoldását keressük, nevezetesen a rendszer *periodikus* megoldását.

Mint ismeretes, akkor mondjuk az $\mathbf{x}(t)$ függvényre, hogy p periódusú periodikus függvény, ha teljesül rá a következő feltétel: $\mathbf{x}(t) = \mathbf{x}(t+p)$. Nyilvánvaló, hogy ha $\mathbf{x}(t)$ periodikus függvény p periódussal, akkor $\dot{\mathbf{x}}(t)$, valamint \mathbf{Ax} ugyanilyen tulajdonságúak. Ebből következik, hogy a periodikus megoldás létezésének szükséges feltétele

$$(3.6.37) \quad \mathbf{f}(t+p) = \mathbf{f}(t).$$

29. Példa. Tekintsük a (3.6.36) differenciálegyenlet-rendszert, és legyen $\mathbf{f}(t)$ egy p periódusú periodikus vektorfüggvény. Határozzuk meg e differenciálegyenlet-rendszer p periódusú periodikus megoldását!

Megoldás. A periodikus megoldást a (3.6.29) alakban írhatjuk fel:

$$\mathbf{x}(t) = e^{\mathbf{A}(t-t_0)}\mathbf{x}_0 + \int_{t_0}^t e^{\mathbf{A}(t-\tau)}\mathbf{f}(\tau) d\tau.$$

A megoldás a $t+p$ helyen:

$$(3.6.38) \quad \mathbf{x}(t+p) = e^{\mathbf{A}(t+p-t_0)}\mathbf{x}_0 + \int_{t_0}^{t+p} e^{\mathbf{A}(t+p-\tau)}\mathbf{f}(\tau) d\tau.$$

Szorozzuk meg $\mathbf{x}(t)$ fenti kifejezését balról $e^{\mathbf{A}p}$ -vel, és vonjuk ki a (3.6.38) egyenletből. Ekkor, mivel $\mathbf{x}(t+p) = \mathbf{x}(t)$, a következő összefüggésre jutunk:

$$(\mathbf{E} - e^{\mathbf{A}p})\mathbf{x}(t) = \int_t^{t+p} e^{\mathbf{A}(t+p-\tau)}\mathbf{f}(\tau) d\tau,$$

és innen – feltéve, hogy az $\mathbf{E} - e^{\mathbf{A}p}$ együtthatómátrix nonszinguláris –

$$(3.6.39) \quad \mathbf{x}(t) = (e^{-\mathbf{A}p} - \mathbf{E})^{-1} \int_t^{t+p} e^{\mathbf{A}(t-\tau)}\mathbf{f}(\tau) d\tau.$$

A kapott eredményből kiolvasható, hogy (3.6.39) az adott differenciálegyenlet-rendszernek az

$$\mathbf{x}(0) = (e^{-\mathbf{A}p} - \mathbf{E})^{-1} \int_0^p e^{-\mathbf{A}\tau}\mathbf{f}(\tau) d\tau$$

kezdeti feltételeket kielégítő megoldása.

* * *

30. Példa. Határozzuk meg az $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{f}_1 \sin \omega t + \mathbf{f}_2 \cos \omega t$ differenciálegyenlet-rendszer $p = \frac{2\pi}{\omega}$ periódusú megoldását!

Megoldás. A megoldáshoz felhasználhatjuk az előző példa (3.6.39) képletét:

$$\mathbf{x}(t) = (e^{-\mathbf{A}p} - \mathbf{E})^{-1} \int_t^{t+p} e^{\mathbf{A}(t-\tau)}\{\mathbf{f}_1 \sin \omega \tau + \mathbf{f}_2 \cos \omega \tau\} d\tau.$$

Elvégezve az integrálást:

$$\mathbf{x}(t) = (\mathbf{A}^2 + \omega^2 \mathbf{E})^{-1} \{\omega \mathbf{E}(\mathbf{f}_2 \sin \omega t - \mathbf{f}_1 \cos \omega t) - \mathbf{A}(\mathbf{f}_2 \cos \omega t + \mathbf{f}_1 \sin \omega t)\}.$$

Innen kiolvasható, hogy ha az \mathbf{A} mátrixnak van olyan tiszta képzetes sajátértéke, amelynek abszolút értéke megegyezik ω értékével, akkor a rendszernek nincs periodikus megoldása, mivel akkor az $(\mathbf{A}^2 + \omega^2 \mathbf{E})$ mátrix szinguláris, tehát nem invertálható. Ezt az esetet *rezonanciának*, ω értékét a rendszer *rezonanciafrekvenciájának* nevezzük.

* * *

31. Példa. Határozzuk meg az alábbi differenciálegyenlet-rendszer $p = \frac{2\pi}{13}$ periódusú megoldását:

$$\begin{aligned}\dot{x}_1 &= -3x_2 + 4x_3 + \sin 13t \\ \dot{x}_2 &= 3x_1 - 12x_3 - 2\sin 13t \\ \dot{x}_3 &= -4x_1 + 12x_2 + 3\sin 13t.\end{aligned}$$

Megoldás. A megoldásra a 30. példa eredményét használjuk fel. Esetünkben

$$\mathbf{A}^2 + \omega^2 \mathbf{E} = \begin{bmatrix} 0 & -3 & 4 \\ 3 & 0 & -12 \\ -4 & 12 & 0 \end{bmatrix}^2 + 169\mathbf{E} = \begin{bmatrix} 144 & 48 & 36 \\ 48 & 16 & 12 \\ 36 & 12 & 9 \end{bmatrix}.$$

Mivel $|\mathbf{A}^2 + \omega^2 \mathbf{E}| = 0$, tehát $\mathbf{A}^2 + \omega^2 \mathbf{E}$ nem invertálható, vagyis a keresett periodikus megoldás nem létezik. Egyébként az \mathbf{A} mátrix karakterisztikus egyenlete $|\lambda \mathbf{E} - \mathbf{A}| = \lambda^3 + 169\lambda = 0$, ahonnan $\lambda_1 = 0$, $\lambda_{2,3} = \pm 13i$, tehát ω értéke megegyezik a tiszta képzetes $\lambda_{2,3}$ sajátérték abszolút értékével. (Az $\omega = 13$ érték a rendszer rezonanciafrekvenciája.)

* * *

3.6.7 Rezgő rendszerek stabilitásvizsgálata

A következőkben egy, az alkalmazások szempontjából igen jelentős esettel foglalkozunk, amikor az együtthatómátrix minimálegyenletének többszörös tiszta képzetes gyöke van.

A polinomok elméletéből ismeretes, hogy egy valós együtthatójú polinom komplex gyöke mindig a konjugált párjával együtt fordul elő. Abban az esetben, ha ez a gyök tiszta képzetes, akkor a konjugált komplex gyökpárt jelölje $\pm i\omega$. Ez azt jelenti, hogy az együtthatómátrix exponenciális függvénye tartalmazza az $e^{i\omega t}$ és $e^{-i\omega t}$ függvényeket. Ezek lineáris kombinációja a valós $\cos \omega t$ és $\sin \omega t$ függvényekhez vezet, ami azt jelenti, hogy a rendszer megoldását – legalábbis részben – csillapítatlan rezgések jellemzik. Abban az esetben azonban, amikor a $\pm i\omega$ gyökpár többszörös gyöke a minimálpolinomnak, akkor az általános elmélet szerint szükségképpen fellépnek olyan megoldások, amelyekben a trigonometrikus függvények polinommal vett szorzatai is megjelennek. Például kétszeres gyök esetén megjelennek $t \cos \omega t$ és $t \sin \omega t$ alakú

kifejezések, ami azt jelenti, hogy a csillapítatlan rezgéseken kívül növekvő amplitúdójú rezgések is létrejönnek, vagyis a rezgő rendszer „elszáll”, pontosabban szólva elveszíti a stabilitását. Ez a jelenség világít rá legjobban arra, hogy hogyan lehet tisztán mátrixelméleti vizsgálattal a differenciálegyenlet-rendszer megoldásának a jellegére következtetni. A mondottak illusztrálására tekintsük az alábbi példát.

32. Példa. Határozzuk meg az

$$\begin{aligned}\dot{x}_1 &= \frac{1}{2}x_1 - \frac{1}{2}x_2 + 2x_3 \\ \dot{x}_2 &= \frac{1}{2}x_1 - \frac{1}{2}x_2 + 2x_4 \\ \dot{x}_3 &= -2x_1 + \frac{1}{2}x_3 - \frac{1}{2}x_4 \\ \dot{x}_4 &= -2x_2 + \frac{1}{2}x_3 - \frac{1}{2}x_4\end{aligned}$$

differenciálegyenlet-rendszer

$$\mathbf{x}(0) = \begin{bmatrix} x_{10} \\ x_{20} \\ x_{30} \\ x_{40} \end{bmatrix}$$

kezdeti feltételeket kielégítő megoldását!

Megoldás. Vegyük észre, hogy az együtthatómátrixot célszerű négy blokkra partícionált másodrendű hipermátrixként felírni:

$$\mathbf{A} = \begin{bmatrix} \mathbf{u}\mathbf{v}^T & 2\mathbf{E} \\ -2\mathbf{E} & \mathbf{u}\mathbf{v}^T \end{bmatrix}, \quad \text{ahol } \mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}^T = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

Az is látható azonnal, hogy a blokkok kommutatívak, tehát a karakterisztikus polinomot a 3.3.1 tétel alapján, mint a *hiperdetermináns* determinánsát számíthatjuk:

$$\lambda\mathbf{E} - \mathbf{A} = \begin{bmatrix} \lambda\mathbf{E} - \mathbf{u}\mathbf{v}^T & -2\mathbf{E} \\ 2\mathbf{E} & \lambda\mathbf{E} - \mathbf{u}\mathbf{v}^T \end{bmatrix},$$

tehát $|\lambda\mathbf{E} - \mathbf{A}| = |(\lambda\mathbf{E} - \mathbf{u}\mathbf{v}^T)(\lambda\mathbf{E} - \mathbf{u}\mathbf{v}^T) + 4\mathbf{E}|$. Mivel $\mathbf{v}^T\mathbf{u} = 0$, a beszorzás után $|\lambda\mathbf{E} - \mathbf{A}| = |(\lambda^2 + 4)\mathbf{E} - 2\lambda\mathbf{u}\mathbf{v}^T|$ adódik, ami nem más, mint egyetlen diáddal módosított mátrix determinánsa. Az 1.5.3 tételt alkalmazzuk: az (1.5.16) és (1.5.17) összefüggések jobb oldalának egyenlőségéből ($|\mathbf{D}| = 1$ helyettesítéssel) azt kapjuk, hogy

$$|\lambda\mathbf{E} - \mathbf{A}| = |(\lambda^2 + 4)\mathbf{E}| \cdot \left(1 - \frac{2\lambda}{\lambda^2 + 4}\mathbf{v}^T\mathbf{u}\right) = (\lambda^2 + 4)^2.$$

Ebből következik, hogy a sajátértékek $\lambda_1 = 2i$ és $\lambda_2 = -2i$ (mindkettő kétszeres gyök), tehát éppen azzal az esettel állunk szemben, amikor kétszeres tiszta képzetes konjugált komplex gyöke van a karakterisztikus polinomnak. Ez egyben a minimálpolinom is, erről legegyszerűbben úgy győződhetünk meg, hogy képezzük a karakterisztikus mátrix jobb felső sarokeleméhez tartozó minort:

$$(\lambda \mathbf{E} - \mathbf{A})_{14} = \begin{vmatrix} -\frac{1}{2} & \lambda + \frac{1}{2} & 0 \\ 2 & 0 & \lambda - \frac{1}{2} \\ 0 & 2 & -\frac{1}{2} \end{vmatrix} = 2\lambda.$$

Tehát sem $\lambda - 2i$, sem $\lambda + 2i$ nem lehet az adjungált mátrix elemeinek közös osztója. Az $e^{\mathbf{A}t}$ mátrixfüggvény meghatározására ezért az Hermite-féle interpoláló alappolinomok segítségével a (3.4.3) képletet használjuk fel.

Az Hermite-féle interpoláló alappolinomokat az alábbi alakban kapjuk meg:

$$\begin{aligned} H_{10}(z) &= \left(\frac{1}{32i}z - \frac{1}{8} \right) (z + 2i)^2, \\ H_{20}(z) &= \left(-\frac{1}{32i}z - \frac{1}{8} \right) (z - 2i)^2, \\ H_{11}(z) &= -\frac{1}{16}(z^2 + 4)(z + 2i), \\ H_{21}(z) &= -\frac{1}{16}(z^2 + 4)(z - 2i). \end{aligned}$$

Behelyettesítve z helyére az \mathbf{A} mátrixot és elvégezve a műveleteket

$$\begin{aligned} H_{10}(\mathbf{A}) &= \frac{1}{2} \begin{bmatrix} \mathbf{E} & -i\mathbf{E} \\ i\mathbf{E} & \mathbf{E} \end{bmatrix}, & H_{20}(\mathbf{A}) &= \frac{1}{2} \begin{bmatrix} \mathbf{E} & i\mathbf{E} \\ -i\mathbf{E} & \mathbf{E} \end{bmatrix}, \\ H_{11}(\mathbf{A}) &= \frac{1}{2} \begin{bmatrix} \mathbf{u}\mathbf{v}^\top & -i\mathbf{u}\mathbf{v}^\top \\ i\mathbf{u}\mathbf{v}^\top & \mathbf{u}\mathbf{v}^\top \end{bmatrix}, & H_{21}(\mathbf{A}) &= \frac{1}{2} \begin{bmatrix} \mathbf{u}\mathbf{v}^\top & i\mathbf{u}\mathbf{v}^\top \\ -i\mathbf{u}\mathbf{v}^\top & \mathbf{u}\mathbf{v}^\top \end{bmatrix} \end{aligned}$$

adódik. Az exponenciális mátrixfüggvény így a következő alakban írható fel:

$$\begin{aligned} e^{\mathbf{A}t} &= \frac{e^{2it}}{2} \left\{ \begin{bmatrix} \mathbf{E} & -i\mathbf{E} \\ i\mathbf{E} & \mathbf{E} \end{bmatrix} + t \begin{bmatrix} \mathbf{u}\mathbf{v}^\top & -i\mathbf{u}\mathbf{v}^\top \\ i\mathbf{u}\mathbf{v}^\top & \mathbf{u}\mathbf{v}^\top \end{bmatrix} \right\} + \\ &+ \frac{e^{-2it}}{2} \left\{ \begin{bmatrix} \mathbf{E} & i\mathbf{E} \\ -i\mathbf{E} & \mathbf{E} \end{bmatrix} + t \begin{bmatrix} \mathbf{u}\mathbf{v}^\top & i\mathbf{u}\mathbf{v}^\top \\ -i\mathbf{u}\mathbf{v}^\top & \mathbf{u}\mathbf{v}^\top \end{bmatrix} \right\}. \end{aligned}$$

Innen egyszerű összevonással és a trigonometrikus alak segítségével megkapjuk a trigonometrikus függvényeket tartalmazó végeredményt. Az is látható,

hogy mindenütt megjelenik a t szorzó:

$$(3.6.40) \quad e^{\mathbf{A}t} = \begin{bmatrix} (\mathbf{E} + t\mathbf{uv}^T) \cos 2t & (\mathbf{E} + t\mathbf{uv}^T) \sin 2t \\ -(\mathbf{E} + t\mathbf{uv}^T) \sin 2t & (\mathbf{E} + t\mathbf{uv}^T) \cos 2t \end{bmatrix}.$$

Ezt egyébként direkt szorzatként is felírhatjuk:

$$e^{\mathbf{A}t} = (\mathbf{E} + t\mathbf{uv}^T) \otimes \begin{bmatrix} \cos 2t & \sin 2t \\ -\sin 2t & \cos 2t \end{bmatrix},$$

vagy részletesebben:

$$e^{\mathbf{A}t} = \begin{bmatrix} 1 + \frac{t}{2} & -\frac{t}{2} \\ \frac{t}{2} & 1 - \frac{t}{2} \end{bmatrix} \otimes \begin{bmatrix} \cos 2t & \sin 2t \\ -\sin 2t & \cos 2t \end{bmatrix}.$$

A feladat megoldását (3.6.29) szerint az

$$\mathbf{x}(t) = e^{\mathbf{A}t} \mathbf{x}_0$$

szorzat adja.

Megjegyezzük, hogy ezt az eredményt a mátrixfüggvény kanonikus előállításával is megkaphatjuk. Tekintsük a mátrixfüggvény (3.4.58) alatti előállítását. A $H_{10}(\mathbf{A})$ és $H_{20}(\mathbf{A})$ diadikus felbontásával nyert \mathbf{U} mátrixszal végzett hasonlósági transzformációval először kvázidiagonalizáljuk a mátrixfüggvényt, azután a diagonálblokkokat Jordan-féle normálalakra transzformáljuk. A példákra alkalmazva:

$$\begin{aligned} H_{10}(\mathbf{A}) &= \frac{1}{2} \begin{bmatrix} \mathbf{E} & -i\mathbf{E} \\ i\mathbf{E} & \mathbf{E} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \mathbf{E} \\ i\mathbf{E} \end{bmatrix} \begin{bmatrix} \mathbf{E} & -i\mathbf{E} \end{bmatrix} = \mathbf{U}_1 \mathbf{V}_1^T, \\ H_{20}(\mathbf{A}) &= \frac{1}{2} \begin{bmatrix} \mathbf{E} & i\mathbf{E} \\ -i\mathbf{E} & \mathbf{E} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \mathbf{E} \\ -i\mathbf{E} \end{bmatrix} \begin{bmatrix} \mathbf{E} & i\mathbf{E} \end{bmatrix} = \mathbf{U}_2 \mathbf{V}_2^T. \end{aligned}$$

Tehát az

$$\mathbf{U} = \begin{bmatrix} \mathbf{E} & \mathbf{E} \\ i\mathbf{E} & -i\mathbf{E} \end{bmatrix}; \quad \mathbf{V}^T = \mathbf{U}^{-1} = \frac{1}{2} \begin{bmatrix} \mathbf{E} & -i\mathbf{E} \\ \mathbf{E} & i\mathbf{E} \end{bmatrix}$$

mátrixok segítségével elvégezve a hasonlósági transzformációt:

$$\begin{aligned} \mathbf{V}^T \mathbf{A} \mathbf{U} &= \\ &= \frac{1}{2} \begin{bmatrix} \mathbf{E} & -i\mathbf{E} \\ \mathbf{E} & i\mathbf{E} \end{bmatrix} \begin{bmatrix} \mathbf{uv}^T & 2\mathbf{E} \\ -2\mathbf{E} & \mathbf{uv}^T \end{bmatrix} \begin{bmatrix} \mathbf{E} & \mathbf{E} \\ i\mathbf{E} & -i\mathbf{E} \end{bmatrix} = \begin{bmatrix} 2i\mathbf{E} + \mathbf{uv}^T & \mathbf{0} \\ \mathbf{0} & -2i\mathbf{E} + \mathbf{uv}^T \end{bmatrix}. \end{aligned}$$

A diagonálblokkokban a sajátértékeket tartalmazó $2i\mathbf{E}$, illetve $-2i\mathbf{E}$ mellett megjelenik az \mathbf{uv}^T nilpotens blokk, következő lépésben tehát hasonlósági transzformációval ezt kell Jordan-blokkra transzformálni. Ha $\mathbf{N} = \mathbf{uv}^T$,

$\mathbf{N}^2 = \mathbf{u}\mathbf{v}^T\mathbf{u}\mathbf{v}^T = \mathbf{0}$, a fővektorokat pedig úgy kapjuk meg, hogy előbb választunk \mathbf{y} és \mathbf{z} vektort, amelyekkel $\mathbf{y}^T\mathbf{N}\mathbf{z} = 1$. Például $\mathbf{y} = \mathbf{e}_1$, $\mathbf{z} = 2\mathbf{e}_2$ esetén ez teljesül. Az \mathbf{x} vektort $\mathbf{x} = (\mathbf{E} + c\mathbf{N})\mathbf{z}$ alakban keressük és a c paramétert úgy határozzuk meg, hogy $\mathbf{y}^T\mathbf{x} = 0$ legyen (lásd a 3.4.6 tételt):

$$\mathbf{y}^T\mathbf{x} = \mathbf{y}^T\mathbf{z} + c\mathbf{y}^T\mathbf{N}\mathbf{z} = 0,$$

és mivel $\mathbf{y}^T\mathbf{N}\mathbf{z} = 1$, $c = -\mathbf{y}^T\mathbf{z} = -2$. Ezzel $\mathbf{x} = (\mathbf{E} - 2\mathbf{N})\mathbf{z} = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$. Ha felső Jordan-blokkra kívánjuk transzformálni az \mathbf{N} nilpotens mátrixot, akkor a fővektorok a következőképpen adódnak:

$$\begin{bmatrix} \mathbf{y}^T \\ \mathbf{y}^T\mathbf{N} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} = \mathbf{T}^{-1}, \text{ ill. } \begin{bmatrix} \mathbf{N}\mathbf{x} & \mathbf{x} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -2 \end{bmatrix} = \mathbf{T}.$$

Ezzel elvégezve a transzformációt:

$$\mathbf{T}^{-1}\mathbf{N}\mathbf{T} = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \mathbf{J},$$

valóban megkaptuk a Jordan-blokkot. Ha most a (3.4.61) összefüggés szerint a Jordan-féle normálalakra transzformált diagonáloblokkok exponenciális függvényét kívánjuk felírni, akkor nyilvánvalóan

$$e^{(2i\mathbf{E}+\mathbf{J})t} = e^{2it}e^{\mathbf{J}t} = e^{2it}(\mathbf{E} + t\mathbf{J}) = e^{2it} \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix},$$

illetve

$$e^{(-2i\mathbf{E}+\mathbf{J})t} = e^{-2it}e^{\mathbf{J}t} = e^{-2it}(\mathbf{E} + t\mathbf{J}) = e^{-2it} \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$$

adódik. Tehát az adott mátrix exponenciális függvényét az alábbi transzformációk segítségével lehet előállítani:

$$\begin{aligned} e^{\mathbf{A}t} &= \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{T} & \\ & \mathbf{T} \end{bmatrix} \begin{bmatrix} e^{2it}(\mathbf{E} + t\mathbf{J}) & \\ & e^{-2it}(\mathbf{E} + t\mathbf{J}) \end{bmatrix} \begin{bmatrix} \mathbf{T}^{-1} & \\ & \mathbf{T}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix} = \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{E} & \mathbf{E} \\ i\mathbf{E} & -i\mathbf{E} \end{bmatrix} \begin{bmatrix} e^{2it}(\mathbf{E} + t\mathbf{T}\mathbf{J}\mathbf{T}^{-1}) & \\ & e^{-2it}(\mathbf{E} + t\mathbf{T}\mathbf{J}\mathbf{T}^{-1}) \end{bmatrix} \begin{bmatrix} \mathbf{E} & -i\mathbf{E} \\ \mathbf{E} & i\mathbf{E} \end{bmatrix} \frac{1}{\sqrt{2}}. \end{aligned}$$

Elvégezve a beszorzásokat, figyelembe véve, hogy $\mathbf{T}\mathbf{J}\mathbf{T}^{-1} = \mathbf{u}\mathbf{v}^T$ és a trigonometrikus alak segítségével áttérve a trigonometrikus függvényekre, végül ugyanazt a (3.6.40) alatti eredményt kapjuk, mint az előző módszer esetén.

* * *

3.6.8 Explicit alakban megadott másodrendű rendszerek

Azt a legegyszerűbb esetet tekintjük, amikor keressük az

$$(3.6.41) \quad \ddot{\mathbf{x}} + \mathbf{A}\mathbf{x} = \mathbf{f}(t)$$

differentiálegyenlet-rendszernek az

$$(3.6.42) \quad \begin{aligned} \mathbf{x}(t_0) &= \mathbf{x}_0 \\ \dot{\mathbf{x}}(t_0) &= \dot{\mathbf{x}}_0 \end{aligned}$$

kezdeti feltételeket kielégítő megoldását, amit a következőképpen jelölünk: $\mathbf{x}(t; t_0, \mathbf{x}_0, \dot{\mathbf{x}}_0)$. Az inhomogén lineáris differentiálegyenlet-rendszer inhomogén kezdeti feltételeket kielégítő megoldását most is az egyenlethez tartozó homogén rendszer inhomogén kezdeti feltételeket kielégítő megoldásának és az inhomogén rendszer homogén kezdeti feltételeket kielégítő megoldásának összegeként állítjuk elő.

Könnyen igazolható, hogy egyetlen differentiálegyenlet esetében az

$$(3.6.43) \quad \ddot{x} + ax = f(t)$$

egyenlet

$$(3.6.44) \quad x(t_0) = x_0, \quad \dot{x}(t_0) = \dot{x}_0$$

kezdeti feltételeket kielégítő megoldása $a > 0$ esetén az alábbi alakban írható fel:

$$(3.6.45) \quad \begin{aligned} x(t; t_0, x_0, \dot{x}_0) &= \\ &= \cos \sqrt{a}(t-t_0) \cdot x_0 + \frac{\sin \sqrt{a}(t-t_0)}{\sqrt{a}} \dot{x}_0 + \int_{t_0}^t \frac{\sin \sqrt{a}(t-\tau)}{\sqrt{a}} f(\tau) d\tau. \end{aligned}$$

Itt az első két tag az $\ddot{x} + ax = 0$ homogén egyenlet (3.6.44) inhomogén feltételeket kielégítő megoldása, az integrált tartalmazó tag pedig a (3.6.43) inhomogén egyenlet $x(t_0) = \dot{x}(t_0) = 0$ homogén feltételeket kielégítő megoldása.

Visszatérve az egyetlen egyenlet esetéről a több egyenletből álló differentiálegyenlet-rendszer tárgyalására, ha teljesen formálisan akarunk eljárni, akkor megkísérelhetjük az általunk vizsgált (3.6.41) rendszernek a (3.6.42) feltételeket kielégítő megoldását is a (3.6.45) kifejezésnek megfelelő alakban előállítani. Ekkor természetesen az a skalár helyett \mathbf{A} mátrixot kell írni, x, x_0, \dot{x}_0 és $f(t)$ helyett pedig vektorok állnak. Mindenekelőtt meg kell

vizsgálni, hogy értelmezhetők-e ekkor egyáltalán a (3.6.45) kifejezésnek megfelelő

$$\cos \sqrt{\mathbf{A}}(t - t_0) \quad \text{és} \quad \frac{\sin \sqrt{\mathbf{A}}(t - t_0)}{\sqrt{\mathbf{A}}}$$

mátrixfüggvények. A 3.3 szakaszban láttuk, hogy a mátrixfüggvényt hatványsorral értelmezzük. Ebből a célból felírjuk a szóban forgó függvények Taylor-sorát:

$$(3.6.46) \quad \begin{aligned} \cos \sqrt{a}(t - t_0) &= 1 - a \frac{(t - t_0)^2}{2!} + a^2 \frac{(t - t_0)^4}{4!} - \dots, \\ \frac{\sin \sqrt{a}(t - t_0)}{\sqrt{a}} &= (t - t_0) - a \frac{(t - t_0)^3}{3!} + a^2 \frac{(t - t_0)^5}{5!} - \dots \end{aligned}$$

Ezekben a függvényekben a csak pozitív egész kitevős hatványon fordul elő és a sorok a komplex számsík bármely véges tartományában konvergensnek. Azt is látjuk, hogy még az $a > 0$ feltételtől is eltekinthetünk, hiszen erre csupán azért volt szükségünk, hogy a (3.6.43) egyenlet valós megoldását trigonometrikus függvények segítségével írassuk le. Negatív a esetén is helyesek maradnak a formuláink, csupán valós argumentumú trigonometrikus függvények helyett képzetes argumentumúak lépnek fel, vagy ami ezzel ekvivalens, hiperbolikus függvényekkel fejezhetők ki a (3.6.46) sorok, azaz a (3.6.43) egyenlet megoldása. A (3.6.46) sorok segítségével tehát tetszőleges \mathbf{A} mátrixra definiálhatók a

$$(3.6.47) \quad \cos \sqrt{\mathbf{A}}(t - t_0), \quad \text{ill.} \quad \frac{\sin \sqrt{\mathbf{A}}(t - t_0)}{\sqrt{\mathbf{A}}}$$

függvények, és éppen a (3.6.46) előállításból következik, hogy mind a négyzetgyökvonás, mind pedig az \mathbf{A} mátrixszal való „osztás” csupán látszólagos, mivel mindkét függvény az \mathbf{A} mátrixnak – pozitív egész kitevős – konvergens hatványsora. Ezzel indokolható, hogy a továbbiakban megengedjük magunknak azt az – egyébként súlyos hibának tekintendő – pongyolaságot, hogy egy mátrixot formálisan a nevezőben szerepeltetünk.

Ezek szerint a (3.6.47) mátrixfüggvények a (3.6.46) sorok segítségével a következőképpen értelmezhetők:

$$(3.6.48) \quad \begin{aligned} \cos \sqrt{\mathbf{A}}(t - t_0) &= \mathbf{E} - \frac{(t - t_0)^2}{2!} \mathbf{A} + \frac{(t - t_0)^4}{4!} \mathbf{A}^2 - \dots, \\ \frac{\sin \sqrt{\mathbf{A}}(t - t_0)}{\sqrt{\mathbf{A}}} &= (t - t_0) \mathbf{E} - \frac{(t - t_0)^3}{3!} \mathbf{A} + \frac{(t - t_0)^5}{5!} \mathbf{A}^2 - \dots \end{aligned}$$

Hátra van még annak igazolása, hogy ezek segítségével a (3.6.45)-nek megfelelően képzett mátrixfüggvény, vagyis

$$(3.6.49) \quad \mathbf{x}(t; t_0, \mathbf{x}_0, \dot{\mathbf{x}}_0) = \\ = \cos \sqrt{\mathbf{A}}(t-t_0)\mathbf{x}_0 + \frac{\sin \sqrt{\mathbf{A}}(t-t_0)}{\sqrt{\mathbf{A}}}\dot{\mathbf{x}}_0 + \int_{t_0}^t \frac{\sin \sqrt{\mathbf{A}}(t-\tau)}{\sqrt{\mathbf{A}}}\mathbf{f}(\tau) d\tau,$$

valóban kielégíti a (3.6.41) differenciálegyenletet és a (3.6.42) kezdeti feltételeket. A (3.6.48) sorokat t szerint kétszer tagonként deriválva, meggyőződhetünk az alábbi azonosságok helyességéről:

$$\begin{aligned} \frac{d}{dt} \frac{\sin \sqrt{\mathbf{A}}(t-t_0)}{\sqrt{\mathbf{A}}} &= \cos \sqrt{\mathbf{A}}(t-t_0), \\ \frac{d^2}{dt^2} \frac{\sin \sqrt{\mathbf{A}}(t-t_0)}{\sqrt{\mathbf{A}}} &= -\sqrt{\mathbf{A}} \sin \sqrt{\mathbf{A}}(t-t_0), \\ \frac{d^2}{dt^2} \cos \sqrt{\mathbf{A}}(t-t_0) &= -\mathbf{A} \cos \sqrt{\mathbf{A}}(t-t_0). \end{aligned}$$

A paraméteres integrált tartalmazó tagot kétszer deriválva (lásd pl. [18]),

$$\begin{aligned} \frac{d}{dt} \int_{t_0}^t \frac{\sin \sqrt{\mathbf{A}}(t-\tau)}{\sqrt{\mathbf{A}}} \mathbf{f}(\tau) d\tau &= \int_{t_0}^t \cos \sqrt{\mathbf{A}}(t-\tau) \mathbf{f}(\tau) d\tau, \\ \frac{d}{dt} \int_{t_0}^t \cos \sqrt{\mathbf{A}}(t-\tau) \mathbf{f}(\tau) d\tau &= \mathbf{f}(t) - \int_{t_0}^t \sqrt{\mathbf{A}} \sin \sqrt{\mathbf{A}}(t-\tau) \mathbf{f}(\tau) d\tau. \end{aligned}$$

A (3.6.49) függvényt behelyettesítve a (3.6.41) differenciálegyenlet-rendszerbe, megállapíthatjuk, hogy azt valóban kielégíti a kezdeti feltételekkel együtt.

Attól függően, hogy \mathbf{A} egyszerű struktúrájú mátrix-e vagy sem, a (3.6.49)-ben előforduló mátrixfüggvények a Lagrange-féle, ill. az Hermite-féle mátrixpolinomok segítségével állíthatók elő.

(a) Ha \mathbf{A} egyszerű struktúrájú mátrix, akkor a keresett megoldást az alábbi alakban kapjuk:

$$(3.6.50) \quad \mathbf{x}(t; t_0, \mathbf{x}_0, \dot{\mathbf{x}}_0) = \\ = \sum_{k=1}^s \left\{ \cos \sqrt{\lambda_k}(t-t_0) L_k(\mathbf{A}) \mathbf{x}_0 + \frac{\sin \sqrt{\lambda_k}(t-t_0)}{\sqrt{\lambda_k}} L_k(\mathbf{A}) \dot{\mathbf{x}}_0 + \right. \\ \left. + \int_{t_0}^t \frac{\sin \sqrt{\lambda_k}(t-\tau)}{\sqrt{\lambda_k}} L_k(\mathbf{A}) \mathbf{f}(\tau) d\tau \right\}.$$

Látható ebből, hogy az \mathbf{A} mátrix pozitív sajátértékeihez trigonometrikus függvények tartoznak, a negatív sajátértékekhez pedig hiperbolikus függvények. Ha \mathbf{A} szinguláris, tehát van zérus sajátértéke – legyen pl. $\lambda_0 = 0$ –, akkor $\cos \sqrt{\lambda_0}(t - t_0) = 1$ és

$$\lim_{\lambda \rightarrow 0} \frac{\sin \sqrt{\lambda}(t - t_0)}{\sqrt{\lambda}} = t - t_0,$$

tehát a (3.6.50) megoldásban szereplő összeg megfelelő tagja

$$(3.6.51) \quad L_0(\mathbf{A}) \left\{ \mathbf{x}_0 + (t - t_0)\dot{\mathbf{x}}_0 + \int_{t_0}^t (t - \tau)\mathbf{f}(\tau)d\tau \right\}.$$

Abban az esetben, ha \mathbf{A} sajátértékei között komplex szám is előfordul, akkor nem feltétlenül kapunk valós megoldást.

(b) Ha \mathbf{A} nem egyszerű struktúrájú mátrix, akkor – amint azt az elsőrendű rendszereknél láttuk – fellépnek

$$\begin{aligned} & (t - t_0) \cos \sqrt{\lambda_k}(t - t_0), \quad (t - t_0)^2 \cos \sqrt{\lambda_k}(t - t_0), \dots \\ & (t - t_0) \frac{\sin \sqrt{\lambda_k}(t - t_0)}{\sqrt{\lambda_k}}, \quad (t - t_0)^2 \frac{\sin \sqrt{\lambda_k}(t - t_0)}{\sqrt{\lambda_k}}, \dots \end{aligned}$$

alakú tagok is. Ebben az esetben az explicit megoldás:

$$\begin{aligned} (3.6.52) \quad \mathbf{x}(t; t_0, \mathbf{x}_0, \dot{\mathbf{x}}_0) = & \sum_{k=1}^s \left\{ \left[\cos \sqrt{\lambda_k}(t - t_0) H_{k0}(\mathbf{A}) - (t - t_0) \frac{\sin \sqrt{\lambda_k}(t - t_0)}{2\sqrt{\lambda_k}} H_{k1}(\mathbf{A}) + \dots \right] \mathbf{x}_0 + \right. \\ & + \left[\frac{\sin \sqrt{\lambda_k}(t - t_0)}{\sqrt{\lambda_k}} H_{k0}(\mathbf{A}) + \frac{1}{2\lambda_k} \left((t - t_0) \cos \sqrt{\lambda_k}(t - t_0) - \right. \right. \\ & \left. \left. - \frac{\sin \sqrt{\lambda_k}(t - t_0)}{\sqrt{\lambda_k}} \right) H_{k1}(\mathbf{A}) + \dots \right] \dot{\mathbf{x}}_0 + \int_{t_0}^t \left[\frac{\sin \sqrt{\lambda_k}(t - \tau)}{\sqrt{\lambda_k}} H_{k0}(\mathbf{A}) + \right. \\ & \left. + \frac{1}{2\lambda_k} \left((t - \tau) \cos \sqrt{\lambda_k}(t - \tau) - \frac{\sin \sqrt{\lambda_k}(t - \tau)}{\sqrt{\lambda_k}} \right) H_{k1}(\mathbf{A}) + \dots \right] \mathbf{f}(\tau) d\tau \Big\}. \end{aligned}$$

Megjegyezzük, hogy a $H_{k\nu}(\mathbf{A})$ polinomok együtthatói itt a

$$\cos \sqrt{\lambda}(t - t_0), \quad \frac{\sin \sqrt{\lambda}(t - t_0)}{\sqrt{\lambda}}, \quad \text{ill.} \quad \frac{\sin \sqrt{\lambda}(t - \tau)}{\sqrt{\lambda}}$$

függvények λ szerinti ν -edik deriváltjai a $\lambda = \lambda_k$ helyeken, ezért válnak az együtthatók egyre bonyolultabbá.

33. Példa. Határozzuk meg az

$$\begin{aligned}\ddot{x}_1 + \frac{4}{3}x_1 + \frac{4}{3}x_2 + \frac{4}{3}x_3 &= 2t, \\ \ddot{x}_2 + \frac{4}{3}x_1 + \frac{11}{6}x_2 + \frac{5}{6}x_3 &= 3t, \\ \ddot{x}_3 + \frac{4}{3}x_1 + \frac{5}{6}x_2 + \frac{11}{6}x_3 &= -t\end{aligned}$$

differenciálegyenlet-rendszer

$$\mathbf{x}(0) = \begin{bmatrix} x_{10} \\ x_{20} \\ x_{30} \end{bmatrix} \quad \text{és} \quad \dot{\mathbf{x}}(0) = \begin{bmatrix} \dot{x}_{10} \\ \dot{x}_{20} \\ \dot{x}_{30} \end{bmatrix}$$

kezdeti feltételeket kielégítő megoldását!

Megoldás. A megoldásra a (3.6.50) képletet használjuk fel. A karakterisztikus egyenlet $|\lambda \mathbf{E} - \mathbf{A}| \equiv \lambda(\lambda - 1)(\lambda - 4) = 0$, innen a sajátértékek $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 4$. Mivel csupa egyszeres sajátérték van, a keresett mátrixfüggvények kanonikus felbontásához a Lagrange-féle alappolinomokat használjuk fel. Ezek:

$$L_1(z) = \frac{1}{4}(z-1)(z-4), \quad L_2(z) = -\frac{1}{3}z(z-4), \quad L_3(z) = \frac{1}{12}z(z-1).$$

Innen

$$\begin{aligned}L_1(\mathbf{A}) &= \frac{1}{4}(\mathbf{A} - \mathbf{E})(\mathbf{A} - 4\mathbf{E}) = \begin{bmatrix} \frac{4}{6} & -\frac{2}{6} & -\frac{2}{6} \\ -\frac{2}{6} & \frac{1}{6} & \frac{1}{6} \\ -\frac{2}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix} = \\ &= \begin{bmatrix} \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{6}} - \frac{1}{\sqrt{6}} - \frac{1}{\sqrt{6}} \end{bmatrix}, \\ L_2(\mathbf{A}) &= \frac{1}{3}\mathbf{A}(4\mathbf{E} - \mathbf{A}) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix},\end{aligned}$$

$$L_3(\mathbf{A}) = \frac{1}{12} \mathbf{A}(\mathbf{A} - \mathbf{E}) = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}.$$

Ezeket behelyettesítve a (3.6.50) képletbe, a következőt kapjuk:

$$\begin{aligned} \mathbf{x}(t) &= \mathbf{T} \langle 1, \cos t, \cos 2t \rangle \mathbf{T}^T \mathbf{x}_0 + \mathbf{T} \left\langle t, \sin t, \frac{1}{2} \sin 2t \right\rangle \mathbf{T}^T \dot{\mathbf{x}}_0 + \\ &+ \int_0^t \mathbf{T} \left\langle (t - \tau), \sin(t - \tau), \frac{1}{2} \sin 2(t - \tau) \right\rangle \mathbf{T}^T \mathbf{f}(\tau) d\tau, \end{aligned}$$

$$\text{ahol} \quad \mathbf{T} = \begin{bmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

a sajátvektorokból alkotott modálmátrix, amely – mivel \mathbf{A} szimmetrikus – ortogonális, azaz $\mathbf{T}^{-1} = \mathbf{T}^T$. A kijelölt integrálást a következőképpen végezzük el:

$$\begin{aligned} \mathbf{T} \int_0^t \left\langle t - \tau, \sin(t - \tau), \frac{1}{2} \sin 2(t - \tau) \right\rangle &\begin{bmatrix} \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} 2\tau \\ 3\tau \\ -\tau \end{bmatrix} d\tau = \\ &= \mathbf{T} \int_0^t \begin{bmatrix} \frac{2}{\sqrt{6}}(t - \tau)\tau \\ \frac{4}{\sqrt{2}}\tau \sin(t - \tau) \\ \frac{2}{\sqrt{3}}\tau \sin 2(t - \tau) \end{bmatrix} d\tau = \mathbf{T} \begin{bmatrix} \frac{1}{3\sqrt{6}}t^3 \\ \frac{4}{\sqrt{2}}(t - \sin t) \\ \frac{1}{\sqrt{3}}\left(t - \frac{1}{2}\sin 2t\right) \end{bmatrix}. \end{aligned}$$

Tehát a keresett megoldás:

$$\begin{aligned}
x_1(t) &= \frac{1}{3}(2x_{10} - x_{20} - x_{30}) + \frac{1}{3}(x_{10} + x_{20} + x_{30}) \cos 2t + \\
&\quad + \frac{1}{3}t(2\dot{x}_{10} - \dot{x}_{20} - \dot{x}_{30}) + \frac{1}{6}(\dot{x}_{10} + \dot{x}_{20} + \dot{x}_{30}) \sin 2t + \\
&\quad + \frac{1}{9}t^3 + \frac{1}{3}t - \frac{1}{6} \sin 2t, \\
x_2(t) &= -\frac{1}{6}(2x_{10} - x_{20} - x_{30}) + \frac{1}{2}(x_{20} - x_{30}) \cos t + \\
&\quad + \frac{1}{3}(x_{10} + x_{20} + x_{30}) \cos 2t - \frac{1}{6}t(2\dot{x}_{10} - \dot{x}_{20} - \dot{x}_{30}) + \\
&\quad + \frac{1}{2}(\dot{x}_{20} - \dot{x}_{30}) \sin t + \frac{1}{6}(\dot{x}_{10} + \dot{x}_{20} + \dot{x}_{30}) \sin 2t - \\
&\quad - \frac{1}{18}t^3 + \frac{7}{3}t - 2 \sin t - \frac{1}{6} \sin 2t, \\
x_3(t) &= -\frac{1}{6}(2x_{10} - x_{20} - x_{30}) - \frac{1}{2}(x_{20} - x_{30}) \cos t + \\
&\quad + \frac{1}{3}(x_{10} + x_{20} + x_{30}) \cos 2t - \frac{1}{6}t(2\dot{x}_{10} - \dot{x}_{20} - \dot{x}_{30}) - \\
&\quad - \frac{1}{2}(\dot{x}_{20} - \dot{x}_{30}) \sin t + \frac{1}{6}(\dot{x}_{10} + \dot{x}_{20} + \dot{x}_{30}) \sin 2t - \\
&\quad - \frac{1}{18}t^3 - \frac{5}{3}t + 2 \sin t - \frac{1}{6} \sin 2t.
\end{aligned}$$

* * *

34. Példa. Számítsuk ki az

$$\begin{aligned}
\ddot{x}_1 &\quad - 8x_2 + 6x_3 = 3 \cos \omega t \\
\ddot{x}_2 - 8x_1 + 12x_2 - 12x_3 &= 2 \cos \omega t \\
\ddot{x}_3 + 6x_1 - 12x_2 + 5x_3 &= -\cos \omega t
\end{aligned}$$

differenciálegyenlet-rendszer rezonanciafrekvenciáit, és határozzuk meg a rendszernek ezekhez tartozó és $\mathbf{x}(0) = \mathbf{0}$, $\dot{\mathbf{x}}(0) = \mathbf{0}$ homogén kezdeti feltételeket kielégítő megoldásait!*

Megoldás. A differenciálegyenlet-rendszer együtthatómátrixa

$$\mathbf{A} = \begin{bmatrix} 0 & -8 & 6 \\ -8 & 12 & -12 \\ 6 & -12 & 5 \end{bmatrix},$$

* Másodrendű rendszer esetén rezonanciáról van szó, ha a rendszer valamelyik sajátrezgése megegyezik a gerjesztő rezgéssel, azaz ha a rendszer együtthatómátrixának van olyan pozitív λ_k sajátértéke, amelynek a négyzetgyöke (sajátfrekvencia) megegyezik ω értékével (gerjesztő frekvencia).

karakterisztikus egyenlete:

$$|\lambda \mathbf{E} - \mathbf{A}| \equiv (\lambda + 4)^2(\lambda - 25) = 0,$$

és így sajátértékei $\lambda_{1,2} = -4$ (kétszeres sajátérték) és $\lambda_3 = 25$. A rendszernek tehát csak egy rezonanciafrekvenciája van: $\omega = 5$. Az ehhez tartozó megoldás a (3.6.50) képlet segítségével számítható. Megjegyezzük, hogy bár a karakterisztikus egyenletnek van többszörös gyöke, mégis – tekintettel arra, hogy az \mathbf{A} mátrix szimmetrikus – a minimálegyenletnek biztosan csupa egyszeres gyöke van. Ezért a (3.6.50) képletben előforduló mátrixfüggvények a Lagrange-féle mátrixpolinomokkal határozhatók meg:

$$\begin{aligned} L_1(\mathbf{A}) &= \frac{1}{29}(25\mathbf{E} - \mathbf{A}) = \frac{1}{29} \begin{bmatrix} 25 & 8 & -6 \\ 8 & 13 & 12 \\ -6 & 12 & 20 \end{bmatrix} = \\ &= \begin{bmatrix} \frac{5}{\sqrt{29}} & 0 \\ \frac{8}{5\sqrt{29}} & \frac{3}{5} \\ -\frac{6}{5\sqrt{29}} & \frac{4}{5} \end{bmatrix} \begin{bmatrix} \frac{5}{\sqrt{29}} & \frac{8}{5\sqrt{29}} & \frac{-6}{5\sqrt{29}} \\ 0 & \frac{3}{5} & \frac{4}{5} \end{bmatrix}, \\ L_2(\mathbf{A}) &= \frac{1}{29}(\mathbf{A} + 4\mathbf{E}) = \frac{1}{29} \begin{bmatrix} 4 & -8 & 6 \\ -8 & 16 & -12 \\ 6 & -12 & 9 \end{bmatrix} = \\ &= \begin{bmatrix} \frac{2}{\sqrt{29}} \\ \frac{4}{\sqrt{29}} \\ \frac{3}{\sqrt{29}} \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{29}} & -\frac{4}{\sqrt{29}} & \frac{3}{\sqrt{29}} \end{bmatrix}. \end{aligned}$$

Így a keresett megoldás

$$\mathbf{x}(t) = \mathbf{T} \int_0^t \left\langle \frac{\sin \sqrt{\lambda_k}(t - \tau)}{\sqrt{\lambda_k}} \right\rangle \mathbf{T}^T \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix} \cos 5\tau \, d\tau,$$

ahol

$$\mathbf{T} = \begin{bmatrix} \frac{5}{\sqrt{29}} & 0 & \frac{2}{\sqrt{29}} \\ \frac{8}{5\sqrt{29}} & \frac{3}{5} & -\frac{4}{\sqrt{29}} \\ -\frac{6}{5\sqrt{29}} & \frac{4}{5} & \frac{3}{\sqrt{29}} \end{bmatrix}.$$

Tekintetbe véve, hogy

$$\int_0^t \frac{\sin \sqrt{\lambda_k}(t-\tau)}{\sqrt{\lambda_k}} \cos \omega \tau d\tau = \begin{cases} \frac{1}{\lambda_k - \omega^2} (\cos \omega t - \cos \sqrt{\lambda_k} t), & \text{ha } \omega \neq \sqrt{\lambda_k}, \\ \frac{t}{2\sqrt{\lambda_k}} \sin \sqrt{\lambda_k} t, & \text{ha } \omega = \sqrt{\lambda_k}, \end{cases}$$

azt kapjuk, hogy

$$\mathbf{x}(t) = \mathbf{T} \begin{bmatrix} \frac{97}{5\sqrt{29}} \cdot \frac{1}{-4-25} (\cos 5t - \operatorname{ch} 2t) \\ \frac{2}{5} \cdot \frac{1}{-4-25} (\cos 5t - \operatorname{ch} 2t) \\ -\frac{5}{\sqrt{29}} \frac{t}{10} \sin 5t \end{bmatrix},$$

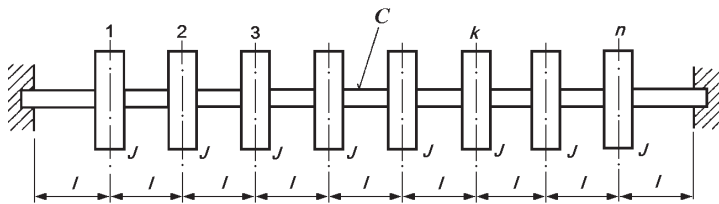
$$x_1(t) = -\frac{97}{29^2} (\cos 5t - \operatorname{ch} 2t) - \frac{t}{29} \sin 5t,$$

$$x_2(t) = -\frac{38}{29^2} (\cos 5t - \operatorname{ch} 2t) + \frac{2t}{29} \sin 5t,$$

$$x_3(t) = \frac{14}{29^2} (\cos 5t - \operatorname{ch} 2t) - \frac{3t}{58} \sin 5t.$$

* * *

35. Példa. *Egydimenziós feladat első peremfeltételekkel.* Két végén befogott, súlytalan tengelyre egyenlő távolságokban azonos inercianyomatékú tárcsákat ékelünk, amelyekre az időtől függő $M_i(t)$ csavarónyomaték hat. A tárcsák száma legyen n , egymástól való távolságuk l , inercianyomatékuk I és a tengely csavarási rugalmassági modulusa C (lásd az ábrát). Határozzuk meg a rendszer csavarórezgéseit!



Megoldás. Jelölje φ_i az egyes tárcsák szögelfordulását; ezekre a következő mozgásegyenletek írhatók fel:

$$(3.6.53) \quad I\ddot{\varphi}_i = \frac{C}{l}(\varphi_{i-1} - \varphi_i) + \frac{C}{l}(\varphi_{i+1} - \varphi_i) + M_i(t).$$

A befogás miatt a 0 és $n + 1$ indexszel jelölt tengelyvégek mozdulatlanok, tehát a (3.6.53) egyenleteket kiegészíthetjük a

$$(3.6.54) \quad \varphi_0 = \varphi_{n+1} = 0$$

feltételekkel, ezeket *első peremfeltételeknek* nevezzük. A (3.6.53) differenciálegyenlet-rendszert $i = 1, 2, \dots, n$ értékeire felírva, és figyelembe véve a (3.6.54) peremfeltételeket, az alábbi differenciálegyenlet-rendszert nyerjük:

$$(3.6.55) \quad I\dot{\varphi} + \frac{C}{l}\mathbf{A}\varphi = \mathbf{M}(t),$$

ahol

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & & \\ -1 & 2 & & \\ & & \ddots & -1 \\ & & -1 & 2 \end{bmatrix}, \quad \varphi(t) = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_n \end{bmatrix}, \quad \mathbf{M}(t) = \begin{bmatrix} M_1(t) \\ M_2(t) \\ \vdots \\ M_n(t) \end{bmatrix}.$$

Ennek a differenciálegyenlet-rendszernek keressük a $\varphi(t_0) = \varphi_0$ és $\dot{\varphi}(t_0) = \dot{\varphi}_0$ kezdeti feltételeket kielégítő megoldását. Bevezetve az $\frac{1}{T^2} = \frac{C}{lI}$ jelölést, és felhasználva a másodrendű, állandó együtthatójú lineáris differenciálegyenlet-rendszerek megoldására vonatkozó (3.6.49) képletet, a keresett megoldást a következő alakban kapjuk:

$$\begin{aligned} \varphi(t) = & \left\{ \cos \left[\frac{1}{T} \sqrt{\mathbf{A}}(t - t_0) \right] \right\} \varphi_0 + \frac{\sin \left[\frac{1}{T} \sqrt{\mathbf{A}}(t - t_0) \right]}{\frac{1}{T} \sqrt{\mathbf{A}}} \dot{\varphi}_0 + \\ & + \int_{t_0}^t \frac{\sin \left[\frac{1}{T} \sqrt{\mathbf{A}}(t - \tau) \right]}{\frac{1}{T} \sqrt{\mathbf{A}}} \frac{1}{I} \mathbf{M}(\tau) d\tau. \end{aligned}$$

Mivel az \mathbf{A} mátrix sajátértékei és sajátvektorainak az elemei (3.2.70) alapján

$$\lambda_k = 4 \sin^2 \frac{k\pi}{2(n+1)}, \quad u_{ik} = \sqrt{\frac{2}{n+1}} \sin \frac{ik\pi}{n+1} \quad (i, k = 1, 2, \dots, n),$$

az egyes tárcsák szögelfordulására az alábbi adódik:

$$(3.6.56) \quad \varphi(t) = \sum_{k=1}^n \mathbf{u}_k \mathbf{u}_k^T \left\{ \cos \frac{\sqrt{\lambda_k}(t - t_0)}{T} \varphi_0 + \frac{\sin \frac{\sqrt{\lambda_k}(t - t_0)}{T}}{\frac{\sqrt{\lambda_k}}{T}} \dot{\varphi}_0 + \right. \\ \left. + \frac{1}{I} \int_{t_0}^t \frac{\sin \frac{\sqrt{\lambda_k}(t - \tau)}{T}}{\frac{\sqrt{\lambda_k}}{T}} \mathbf{M}(\tau) d\tau \right\},$$

azaz

$$\begin{aligned} \varphi_i(t) = & \frac{2}{n+1} \sum_{k=1}^n \sin \frac{ik\pi}{n+1} \left\{ \cos \left(\frac{t-t_0}{T} 2 \sin \frac{k\pi}{2(n+1)} \right) \sum_{j=1}^n \varphi_{0j} \sin \frac{jk\pi}{n+1} + \right. \\ & + \frac{\sin \left(\frac{t-t_0}{T} 2 \sin \frac{k\pi}{2(n+1)} \right)}{\frac{2}{T} \sin \frac{k\pi}{2(n+1)}} \sum_{j=1}^n \dot{\varphi}_{0j} \sin \frac{jk\pi}{n+1} + \\ & \left. + \frac{1}{I} \int_{t_0}^t \frac{\sin \left(\frac{t-\tau}{T} 2 \sin \frac{k\pi}{2(n+1)} \right)}{\frac{2}{T} \sin \frac{k\pi}{2(n+1)}} \sum_{j=1}^n M_j(\tau) \sin \frac{jk\pi}{n+1} d\tau \right\}. \end{aligned}$$

* * *

36. Példa. Egydimenziós feladat második peremfeltételekkel. Oldjuk meg az 35. példát két végén csapágyazott tengely esetére.

Megoldás. Mivel ekkor a tengelyvég együtt mozog a hozzá legközelebb levő tárcsával, ezért az ún. második peremérték-feladatnak megfelelő $\varphi_0 = \varphi_1$, $\varphi_n = \varphi_{n+1}$ feltételeket kell figyelembe venni. Ezzel a rendszer első és utolsó egyenlete a következőképpen módosul:

$$\begin{aligned} I\ddot{\varphi}_1 &= \frac{C}{l}(\varphi_2 - \varphi_1), \\ I\ddot{\varphi}_n &= \frac{C}{l}(\varphi_{n-1} - \varphi_n), \end{aligned}$$

vagyis a (3.6.55) egyenletben szereplő \mathbf{A} mátrix helyett most az

$$\tilde{\mathbf{A}} = \begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & & \\ & & & 2 & -1 \\ & & & -1 & 1 \end{bmatrix}$$

mátrixot kell írni. Mivel a 3.2 szakasz 9. példája alapján ennek sajátértékei és sajátvektorainak elemei

$$\begin{aligned} \lambda_k &= 4 \sin^2 \frac{k\pi}{2n} \quad (k = 0, 1, 2, \dots, n-1), \\ u_{i0} &= \frac{1}{\sqrt{n}}, \quad u_{ik} = \sqrt{\frac{2}{n}} \cos \frac{(2i-1)k\pi}{2n} \quad (k = 1, 2, \dots, n-1), \end{aligned}$$

a (3.6.56) megoldás felhasználásával az egyes tárcsák szögelfordulására most az alábbi adódik:

$$\begin{aligned}\varphi_i(t) = & \frac{2}{n} \sum_{k=1}^{n-1} \cos \frac{(2i-1)k\pi}{2n} \left\{ \cos \left(\frac{t-t_0}{T} 2 \sin \frac{k\pi}{2n} \right) \sum_{j=1}^n \varphi_{0j} \cos \frac{(2j-1)k\pi}{2n} + \right. \\ & + \frac{\sin \left(\frac{t-t_0}{T} 2 \sin \frac{k\pi}{2n} \right)}{\frac{2}{T} \sin \frac{k\pi}{2n}} \sum_{j=1}^n \dot{\varphi}_{0j} \cos \frac{(2j-1)k\pi}{2n} + \\ & + \frac{1}{I} \int_{t_0}^t \frac{\sin \left(\frac{t-\tau}{T} 2 \sin \frac{k\pi}{2n} \right)}{\frac{2}{T} \sin \frac{k\pi}{2n}} \sum_{j=1}^n \cos \frac{(2j-1)k\pi}{2n} M_j(\tau) d\tau \left. \right\} + \\ & + \frac{1}{n} \sum_{j=1}^n \varphi_{0j} + \frac{t-t_0}{n} \sum_{j=1}^n \dot{\varphi}_{0j} + \frac{1}{I} \int_{t_0}^t \frac{t-\tau}{n} \sum_{j=1}^n M_j(\tau) d\tau.\end{aligned}$$

Az utolsó két tag a tengelynek egyenletes szögsebességű forgását fejezi ki, amelyre szuperponálódnak a tárcsák torziós rezgései.

* * *

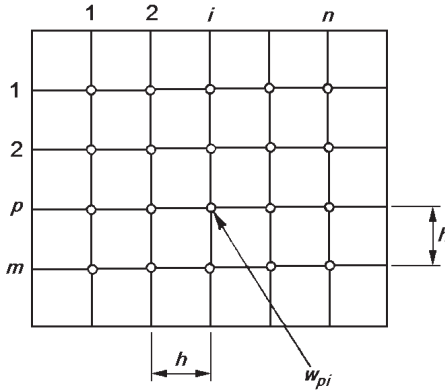
37. Példa. *Kétdimenziós feladat első peremfeltételekkel.* Tekintsünk egy téglalap alakú korpuszkuális membránt, melyet úgy nyerünk, hogy az $x = ph$ (ahol $p = 1, 2, \dots, m$), $y = ih$ (ahol $i = 1, 2, \dots, n$) egyenesek mentén tömegtelen, rugalmas fonalakat helyezünk el, és ezek végeit a $p = 0$, $p = m + 1$, illetve $i = 0$, $i = n + 1$ vonalak mentén úgy rögzítjük, hogy a bennük ébredő S feszítőerő „nagy” legyen, a fonalak metszéspontjaiba – az ip jelű rácspontokba – pedig M tömegű tömegpontokat helyezünk. Vizsgáljuk meg az így kapott rugalmas rendszer szabad rezgéseit.

Megoldás (Kronecker-polinom spektrálfelbontásával). Mivel az egyensúlyi helyzet körüli kis elmozdulások vizsgálatára szorítkozunk, a mozgásegyenletek felírásakor azzal a feltevessel élünk, hogy a mozgás során a fonalakban működő feszítőerő nagysága mindenütt S , iránya pedig a nyugalmi helyzetben elfoglalt irányával olyan kis szöget zár be, amelynek szinusz helyett tangense vehető. A pi indexű pont kitérését jelölje w_{pi} . A mozgásegyenletek a következő alakban írhatók (lásd az ábrát):

$$M\ddot{w}_{pi} = \frac{S}{h} [(w_{p,i+1} - w_{pi}) + (w_{p,i-1} - w_{pi}) + (w_{p+1,i} - w_{pi}) + (w_{p-1,i} - w_{pi})].$$

Átrendezve:

$$(3.6.57) \quad \ddot{w}_{pi} + \frac{S}{Mh} [4w_{pi} - w_{p,i+1} - w_{p,i-1} - w_{p+1,i} - w_{p-1,i}] = 0.$$



Figyelembe véve a

$$w_{p0} = w_{p,n+1} = w_{0i} = w_{m+1,i} = 0$$

peremfeltételeket, ez olyan differenciálegyenlet-rendszerként írható fel, amelynek együtthatómátrixa a (3.3.26) Kronecker-polinom alakú hiper-mátrix. Vezessük be az s^{-2} dimenziójú S/Mh mennyiségre az $\frac{1}{T^2} = \frac{S}{Mh}$ jelölést; ezzel a (3.6.57) egyenlet az alábbi alakban írható fel:

$$(3.6.58) \quad \ddot{\mathbf{w}} + \frac{1}{T^2} \mathbf{A} \mathbf{w} = 0,$$

ahol

$$(3.6.59) \quad \mathbf{A} = (2\mathbf{E}_m - \mathbf{K}_m) \otimes \mathbf{E}_n + \mathbf{E}_m \otimes (2\mathbf{E}_n - \mathbf{K}_n),$$

és

$$(3.6.60) \quad \mathbf{K}_m = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & 1 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix} \begin{matrix} (1) \\ (2) \\ (3) \\ \vdots \\ (m-1) \\ (m) \end{matrix}.$$

A (3.6.58) egyenletnek a $\mathbf{w}(0) = \mathbf{w}_0$, $\dot{\mathbf{w}}(0) = \dot{\mathbf{w}}_0$ kezdeti feltételeket kielégítő megoldása a (3.6.49) képlet alapján felírható mint

$$(3.6.61) \quad \mathbf{w} = \cos\left(\sqrt{\mathbf{A}} \frac{t}{T}\right) \mathbf{w}_0 + \frac{\sin \sqrt{\mathbf{A}} \frac{t}{T}}{\frac{1}{T} \sqrt{\mathbf{A}}} \dot{\mathbf{w}}_0.$$

Felhasználva a (3.6.60) alatti egyszerű kontinuáns mátrix ismert spektrálfelbontását, ennek segítségével a 3.3.4 tétel alapján felírhatók a (3.6.59) direkt

polinomnak a (3.6.61) összefüggésben szereplő függvényei. Jelölje $\lambda_k^{(m)}$, ill. $\lambda_l^{(n)}$ a \mathbf{K}_m , illetve \mathbf{K}_n mátrix sajátértékeit, és $\mathbf{u}_k^{(m)}$, ill. $\mathbf{u}_l^{(n)}$ a sajátvektorait. Ezek segítségével a megoldás a következő alakban írható fel:

$$\mathbf{w} = \sum_{l=1}^n \sum_{k=1}^m \left\{ \cos \left(\sqrt{\lambda_k^{(m)} + \lambda_l^{(n)}} \frac{t}{T} \right) (\mathbf{u}_k^{(m)} \otimes \mathbf{u}_l^{(n)}) (\mathbf{u}_k^{(m)} \otimes \mathbf{u}_l^{(n)})^T \mathbf{w}_0 + \right. \\ \left. + \frac{\sin \left(\sqrt{\lambda_k^{(m)} + \lambda_l^{(n)}} \frac{t}{T} \right)}{\frac{1}{T} \sqrt{\lambda_k^{(m)} + \lambda_l^{(n)}}} (\mathbf{u}_k^{(m)} \otimes \mathbf{u}_l^{(n)}) (\mathbf{u}_k^{(m)} \otimes \mathbf{u}_l^{(n)})^T \dot{\mathbf{w}}_0 \right\}.$$

Behelyettesítve a sajátértékek és sajátvektorok elemeinek ismert kifejezéseit (vö. (3.3.28)), a rácsmembrán tetszőleges pi indexű pontjának mozgását meghatározó következő kifejezést nyerjük:

$$w_{pi} = \frac{4}{(m+1)(n+1)} \sum_{l=1}^n \sum_{k=1}^m \sin \frac{pk\pi}{m+1} \sin \frac{il\pi}{n+1} \times \\ \times \left\{ \cos \omega_{kl} \frac{t}{T} \sum_{j=1}^n \sum_{q=1}^m w_{jq}(0) \sin \frac{qk\pi}{m+1} \sin \frac{j l \pi}{n+1} + \right. \\ \left. + \frac{\sin \omega_{kl} \frac{t}{T}}{\frac{1}{T} \omega_{kl}} \sum_{j=1}^n \sum_{q=1}^m \dot{w}_{jq}(0) \sin \frac{qk\pi}{m+1} \sin \frac{j l \pi}{n+1} \right\},$$

ahol

$$\omega_{kl} = \sqrt{4 \sin^2 \frac{k\pi}{2(m+1)} + 4 \sin^2 \frac{l\pi}{2(n+1)}}.$$

* * *

38. Példa. Határozzuk meg a

$$(3.6.62) \quad \frac{d\mathbf{V}}{dt} = \mathbf{A}^T \mathbf{V} + \mathbf{V} \mathbf{A}$$

mátrix-differenciálegyenletnek a

$$(3.6.63) \quad \mathbf{V}(0) = \mathbf{V}_0$$

kezdeti feltételeket kielégítő megoldását! (Lásd [2].)

Megoldás. Mindenekelőtt alakítsuk át a (3.6.62) mátrix-differenciálegyenletet differenciálegyenlet-rendszerre oly módon, hogy *oszloponként* írjuk fel a differenciálegyenletet. Ha a \mathbf{V} és az \mathbf{A} mátrixot oszlopaira particionálva $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_n]$, ill. $\mathbf{A} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_n]$ alakban írjuk fel, akkor a következő egyenleteket kapjuk:

$$\begin{aligned}\dot{\mathbf{v}}_1 &= \mathbf{A}^\top \mathbf{v}_1 + \mathbf{V} \mathbf{a}_1 = \mathbf{A}^\top \mathbf{v}_1 + \sum_{k=1}^n a_{k1} \mathbf{v}_k, \\ \dot{\mathbf{v}}_2 &= \mathbf{A}^\top \mathbf{v}_2 + \mathbf{V} \mathbf{a}_2 = \mathbf{A}^\top \mathbf{v}_2 + \sum_{k=1}^n a_{k2} \mathbf{v}_k, \\ &\vdots \\ \dot{\mathbf{v}}_n &= \mathbf{A}^\top \mathbf{v}_n + \mathbf{V} \mathbf{a}_n = \mathbf{A}^\top \mathbf{v}_n + \sum_{k=1}^n a_{kn} \mathbf{v}_k.\end{aligned}$$

Bevezetve a \mathbf{v}_i oszlopokból alkotott hipervektort, a fenti differenciálegyenlet-rendszer a következő alakba írható:

$$\begin{bmatrix} \dot{\mathbf{v}}_1 \\ \dot{\mathbf{v}}_2 \\ \vdots \\ \dot{\mathbf{v}}_n \end{bmatrix} = \left\{ \begin{bmatrix} \mathbf{A}^\top & & \\ & \mathbf{A}^\top & \\ & & \ddots \\ & & & \mathbf{A}^\top \end{bmatrix} + \begin{bmatrix} a_{11}\mathbf{E} & a_{21}\mathbf{E} & \dots & a_{n1}\mathbf{E} \\ a_{12}\mathbf{E} & a_{22}\mathbf{E} & \dots & a_{n2}\mathbf{E} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n}\mathbf{E} & a_{2n}\mathbf{E} & \dots & a_{nn}\mathbf{E} \end{bmatrix} \right\} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_n \end{bmatrix},$$

azaz a direkt szorzatok bevezetésével

$$(3.6.64) \quad \dot{\mathbf{v}} = (\mathbf{A}^\top \otimes \mathbf{E} + \mathbf{E} \otimes \mathbf{A}^\top) \mathbf{v}$$

ahol \mathbf{v} a \mathbf{v}_i vektorokból alkotott hipervektort jelöli. A (3.6.64) egyenlet $\mathbf{v}(0) = \mathbf{v}_0$ kezdeti feltételeket kielégítő megoldása

$$(3.6.65) \quad \mathbf{v} = e^{(\mathbf{A}^\top \otimes \mathbf{E} + \mathbf{E} \otimes \mathbf{A}^\top)t} \mathbf{v}_0.$$

A 3.3.6 tétel felhasználásával ez $\mathbf{v} = (e^{\mathbf{A}^\top t} \otimes e^{\mathbf{A}^\top t}) \mathbf{v}_0$ alakban írható, majd ismét visszatérve a \mathbf{V} mátrixra, a (3.6.62) egyenlet (3.6.63) kezdeti feltételt kielégítő megoldása

$$\mathbf{V} = e^{\mathbf{A}^\top t} \mathbf{V}_0 e^{\mathbf{A} t}.$$

* * *

4. fejezet

NEMNEGATÍV ELEMŰ MÁTRIXOK

Arra a kérdésre, hogy a valószínűesszámitás valóban a matematika egy ága-e, a válasz persze attól függ, mit értünk matematikán....

Én ... e tekintetben Descartes pártján állok, aki azt mondotta, hogy a matematikához kell számítani minden olyan vizsgálatot, amely a rend és a mérték kutatására irányul, függetlenül attól, mi a tárgya, minek a rendjét és mértékét keresi.

(Rényi Alfréd: Ars Mathematica, Levelek a valószínűségről)

Ebben a fejezetben olyan mátrixok tulajdonságait vizsgáljuk, amelyeknek elemei nemnegatív valós számok. Számos alkalmazási területen ugyanis ilyen típusú mátrixok fordulnak elő, ezért indokolt ezeket külön tárgyalni. Az eredmények nagy része elég egyszerűen megfogalmazható tételekben foglalható össze – egyes tételek bizonyítása azonban mély megfontolást és igen finom megkülönböztetéseket igényel. Már a fejezet címéből is kitűnik, hogy általában nem *pozitív*, hanem *nemnegatív* elemű mátrixokkal foglalkozunk. Elsősorban a nemnegatív elemű mátrixok sajátértékeinek és sajátvektorainak a tulajdonságait vizsgáljuk – ezeket együttesen a mátrix spektrális tulajdonságainak nevezzük. Bebonyóítjuk az irreducibilis, nemnegatív elemű mátrixok spektrális tulajdonságaira vonatkozó Frobenius-féle tételeket. Ezek bizonyítását az Olvasó első olvasáskor átlapozhatja, mert elég nehezen követhetők. E könyvből mégsem akartuk a bizonyításokat kihagyni, mert egyrészt az egész további elmélet alapját képezik, másrészt betekintést nyújtanak az igényesebb Olvasó számára a finomabb megfontolásokat igénylő mátrixelméleti bizonyításokban használatos módszerekbe (lásd pl. [6], [44]).

A nemnegatív elemű mátrixok tárgyalása során meg kell különböztetni a *reducibilis* és az *irreducibilis* mátrixok osztályát – de megjegyezzük, hogy maga a reducibilitás, ill. irreducibilitás fogalma tetszőleges elemű mátrixokra is érvényes.

4.0.1 definíció. *Reducibilisnek* nevezzük egy (kvadrátikus) mátrixot, ha egy alkalmasan választott permutáló mátrixszal végzett ortogonális transzformációval – azaz sorainak és oszlopainak ugyanolyan permutációjával – olyan szimmetrikusan partícionált, négy blokkból álló mátrixba transzformálható, amelynek jobb felső vagy bal alsó blokkja zérusmátrix; tehát az adott A

mátrix reducibilis, ha létezik olyan \mathbf{P} permutáló mátrix, amellyel

$$\mathbf{PAP}^T = \left[\begin{array}{c|c} \mathbf{B} & \mathbf{0} \\ \hline \mathbf{G} & \mathbf{C} \end{array} \right], \quad \text{vagy} \quad \mathbf{PAP}^T = \left[\begin{array}{c|c} \mathbf{B} & \mathbf{F} \\ \hline \mathbf{0} & \mathbf{C} \end{array} \right],$$

ahol \mathbf{B} és \mathbf{C} a transzformált mátrix particionálásával keletkezett hipermátrix egy-egy kvadratikus blokkja. A mátrix **irreducibilis**, ha ilyen \mathbf{P} permutáló mátrix nem létezik.

A nemnegatív elemű mátrixokra vonatkozó legfontosabb eredmények a legnagyobb abszolút értékű sajátértékük tulajdonságaira, becslésére és a hozzá tartozó sajátvektorra vonatkoznak. Külön pontban foglalkozunk a sztochasztikus mátrixokkal; ezeket a homogén Markov-láncok* elméletében alkalmazzuk. A többlépéses átmenetvalószínűségek meghatározása mellett a határeloszlás létezésének feltételeire és kiszámítására vonatkozóan kapunk eredményeket. Ezzel kapcsolatban érintjük az ergodicitás fogalmát is. Ezután megmutatjuk, hogyan alkalmazhatók az eredmények bolyongási feladatok vizsgálatára, végül meghatározzuk egy nevezetes sztochasztikus mátrix spektrálfelbontását generátorfüggvény segítségével.

4.1 IRREDUCIBILIS MÁTRIXOK

Az irreducibilis mátrixok tárgyalásában központi helyet foglal el a Frobeniustól** származó alaptétel (lásd [6], [11], [19]). Mivel ennek bizonyítása elég hosszadalmas, megértésének a megkönnyítése céljából két részre bontva tárgyaljuk és a továbbiakban Frobenius I. illetve Frobenius II. tételeként hivatkozunk majd rájuk. Bevezetőben néhány segédttétellel készítjük elő a bizonyításukat, majd bemutatjuk a speciális esetként adódó, pozitív elemű mátrixokra vonatkozó Perron-féle tételt. Végül megadjuk a nemnegatív elemű mátrixok legnagyobb abszolút értékű sajátértékeinek becslésére vonatkozó – igen egyszerű, de fontos – összefüggést. Az alábbiakban mindig valós elemű mátrixokról lesz szó.

4.1.1 Út a Frobenius-tételekhez

Vezessük be a következő jelöléseket:

$$\begin{aligned} \mathbf{A} > \mathbf{0}, & \text{ ha } a_{\mu\nu} > 0 & \text{ minden } \mu\nu \text{ indexpárra,} \\ \mathbf{A} \geq \mathbf{0}, & \text{ ha } a_{\mu\nu} \geq 0 & \text{ minden } \mu\nu \text{ indexpárra,} \end{aligned}$$

* A. A. Markov (1865–1922) orosz matematikus.

** G. Frobenius (1849–1917) német matematikus.

hasonlóan

$\mathbf{A} > \mathbf{B}$, ha $a_{\mu\nu} > b_{\mu\nu}$ minden $\mu\nu$ indexpárra,

$\mathbf{A} \geq \mathbf{B}$, ha $a_{\mu\nu} \geq b_{\mu\nu}$ minden $\mu\nu$ indexpárra.

Tehát $\mathbf{A} \geq \mathbf{0}$ azt jelenti, hogy \mathbf{A} nemnegatív elemű mátrix, vagy röviden nemnegatív mátrix. Ennek megfelelően használjuk a nemnegatív vektor, továbbá a pozitív mátrix, ill. pozitív vektor elnevezéseket.

Megjegyzés. Itt felhívjuk a figyelmet arra, hogy ez a jelölés a skaláris mennyiségekre megszokott jelölés általánosítása. Ezért mátrixok esetében az $\mathbf{A} \geq 0$ és $\mathbf{A} \neq 0$ relációk együttes teljesülése azt jelenti, hogy \mathbf{A} minden eleme nemnegatív és \mathbf{A} nem a zérusmátrix – tehát van pozitív eleme is.

Először bebizonyítjuk a következő tételt.

4.1.1 tétel. Ha $\mathbf{A} \geq 0$ irreducibilis, n -edrendű mátrix, akkor

$$(4.1.1) \quad (\mathbf{E} + \mathbf{A})^{n-1} > \mathbf{0}$$

teljesül.

Bizonyítás. A tétel bizonyításához elegendő azt belátni, hogy tetszőleges nemnegatív elemű \mathbf{y} vektorra (amely azonban nem a zérusvektor), azaz tetszőleges $\mathbf{y} \geq \mathbf{0}$, $\mathbf{y} \neq \mathbf{0}$ esetén

$$(4.1.2) \quad (\mathbf{E} + \mathbf{A})^{n-1} \mathbf{y} > \mathbf{0}.$$

Ez az állítás viszont közvetlenül adódik abból, hogy a $\mathbf{z} = (\mathbf{E} + \mathbf{A})\mathbf{y}$ vektornak mindig több, zérustól különböző eleme van, mint az \mathbf{y} vektornak. Ezt indirekt úton bizonyítjuk be: feltesszük, hogy a \mathbf{z} és az \mathbf{y} vektornak ugyanannyi eleme zérustól különböző. Mivel

$$(4.1.3) \quad \mathbf{z} = \mathbf{y} + \mathbf{A}\mathbf{y}, \quad \mathbf{A}\mathbf{y} \geq \mathbf{0},$$

emiatt azokra a μ indexekre, amelyekre $y_\mu > 0$, egyúttal $z_\mu > 0$, ezért feltevésünk szerint a \mathbf{z} és az \mathbf{y} vektornak ugyanazon indexű elemei zérustól különbözőek. Ez a feltevés azonban ellentétben áll az \mathbf{A} mátrix irreducibilitásával. Ugyanis a koordináták átszámozásával és az \mathbf{A} mátrix sorainak és oszlopainak megfelelő átrendezésével (azaz egy permutáló mátrixszal végzett ortogonális transzformációval), valamint megfelelő particionálással a (4.1.3) összefüggés

$$(4.1.4) \quad \begin{bmatrix} \mathbf{v} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{u} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \tilde{\mathbf{A}}_{11} & \tilde{\mathbf{A}}_{12} \\ \tilde{\mathbf{A}}_{21} & \tilde{\mathbf{A}}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{0} \end{bmatrix} \quad (\mathbf{u} > \mathbf{0}, \mathbf{v} > \mathbf{0})$$

alakban írható, ahol $\tilde{\mathbf{A}}$ jelöli az átrendezett \mathbf{A} mátrixot. Innen kiolvasható, hogy $\tilde{\mathbf{A}}_{21}\mathbf{u} = \mathbf{0}$, ami $\mathbf{u} > \mathbf{0}$ miatt csak akkor állhatna fenn, ha $\tilde{\mathbf{A}}_{21} = \mathbf{0}$ lenne.

Mivel ez valóban azt jelentené, hogy \mathbf{A} reducibilis, ezzel a tételt bebizonyítottuk. ■

Ebből az eredményből következik az alábbi tétel.

4.1.2 tétel. Jelölje $a_{\mu\nu}^{(q)}$ az \mathbf{A}^q mátrix elemeit. Ha \mathbf{A} nemnegatív elemű irreducibilis mátrix, akkor minden $\mu\nu$ indexpárhoz található olyan q pozitív szám, amelyre

$$(4.1.5) \quad a_{\mu\nu}^{(q)} > 0,$$

és ha m jelöli az \mathbf{A} mátrix $\Delta(\lambda)$ minimálpolinomjának fokszámát, akkor q mindig megválasztható úgy, hogy

$$(4.1.6) \quad \begin{aligned} q &\leq m-1, \text{ ha } \mu \neq \nu, \\ q &\leq m, \quad \text{ ha } \mu = \nu \end{aligned}$$

teljesüljön.

Bizonyítás. Képezzük a $(\lambda + 1)^n$ polinomnak a $\Delta(\lambda)$ minimálpolinomra vett osztási maradékát: $(\lambda + 1)^n \equiv \Delta(\lambda)Q(\lambda) + R(\lambda)$. Helyettesítsük be ebbe az egyváltozós racionális egész skalár azonosságba λ helyére az \mathbf{A} mátrixot:

$$(4.1.7) \quad (\mathbf{A} + \mathbf{E})^n = \Delta(\mathbf{A})Q(\mathbf{A}) + R(\mathbf{A}).$$

Mivel $\Delta(\mathbf{A}) = \mathbf{0}$, a 4.1.1 tétel értelmében pedig $(\mathbf{A} + \mathbf{E})^n > \mathbf{0}$, ezért ebből

$$(4.1.8) \quad R(\mathbf{A}) > \mathbf{0},$$

ahol $R(\lambda)$ egy legfeljebb $(m-1)$ -edfokú polinom. A kapott (4.1.8) egyenlőtlenségből következik, hogy minden $\mu\nu$ indexpárra a nemnegatív

$$\delta_{\mu\nu}, \quad a_{\mu\nu}, \quad a_{\mu\nu}^{(2)}, \quad \dots, \quad a_{\mu\nu}^{(m-1)}$$

számok közül legalább az egyik pozitív. Mivel $\mu \neq \nu$ esetén $\delta_{\mu\nu} = 0$, innen azonnal adódik a (4.1.6) egyenlőtlenségek közül az első, a $\mu = \nu$ esetre vonatkozót pedig úgy nyerjük, hogy a (4.1.8) egyenlőtlenséget az

$$(4.1.9) \quad \mathbf{A}R(\mathbf{A}) > \mathbf{0}$$

egyenlőtlenséggel helyettesítjük. (Pozitív elemű mátrix és irreducibilis, nemnegatív elemű mátrix szorzata ugyanis nyilvánvalóan pozitív elemű mátrix.) Ezzel a tételt bebizonyítottuk. ■

Megjegyzés. A most bebizonyított tétel alapján a (4.1.1) egyenlőtlenségben szereplő $n-1$ kitevő $(m-1)$ -gyel helyettesíthető.

Vezessük be egy mátrix, ill. vektor elemeinek abszolút értékeiből alkotott mátrixra, ill. vektorra a következő jelölést:

$$(4.1.10) \quad \mathbf{A}^+ = [|a_{\mu\nu}|], \quad \text{ill.} \quad \mathbf{x}^+ = [|x_\mu|].$$

A háromszög-egyenlőtlenségből következik, hogy ha \mathbf{A} nemnegatív elemű mátrix és fennáll $\mathbf{y} = \mathbf{Ax}$, azaz $y_\mu = \sum_{\nu=1}^n a_{\mu\nu}x_\nu$, akkor

$$|y_\mu| = \left| \sum_{\nu=1}^n a_{\mu\nu}x_\nu \right| \leq \sum_{\nu=1}^n a_{\mu\nu}|x_\nu|,$$

vagyis

$$(4.1.11) \quad \mathbf{y}^+ \leq \mathbf{Ax}^+.$$

Legyen \mathbf{x} rögzített, nemnegatív elemű, zérustól különböző vektor. Ekkor az

$$(4.1.12) \quad \mathbf{y} = \mathbf{Ax}$$

transzformációval adódó \mathbf{y} is nyilván nemnegatív elemű vektor. Rendeljük hozzá az \mathbf{x} vektorhoz az

$$(4.1.13) \quad r(\mathbf{x}) = \min_{\mu} \frac{y_\mu}{x_\mu}$$

összefüggéssel meghatározott $r(\mathbf{x})$ számot*, kirekesztve azokat a μ_κ értékeket, amelyekre

$$(4.1.14) \quad x_{\mu_\kappa} = 0.$$

A (4.1.12) összefüggésből következik, hogy

$$(4.1.15) \quad r(\mathbf{x}) \geq 0,$$

és egyúttal $r(\mathbf{x})$ az a legnagyobb ϱ valós szám, amelyre még fennáll

$$(4.1.16) \quad \varrho \mathbf{x} \leq \mathbf{Ax}.$$

Megmutatjuk majd, hogy az $r(\mathbf{x})$ függvény az egységgömbön** valamely nemnegatív \mathbf{u} vektorra felveszi a maximumát, vagyis hogy az

$$(4.1.17) \quad |\mathbf{x}| = \sqrt{\sum_{\mu=1}^n x_\mu^2} = 1, \quad \mathbf{x} \geq \mathbf{0}$$

*Mivel a (4.1.13) összefüggéssel definiált mennyiség az \mathbf{x} vektor és az \mathbf{A} mátrix függvénye, ezért indokoltabb volna az $r_{\mathbf{A}}(\mathbf{x})$ jelölés alkalmazása, egyszerűség kedvéért azonban az \mathbf{A} indexet elhagyjuk.

**A geometriai térből vett szemlélet alapján az n -dimenziós euklideszi tér egységnyi abszolút értékű \mathbf{x} vektorainak a halmazát egységgömbnek nevezzük.

mellékfeltételek mellett az

$$(4.1.18) \quad r = r(\mathbf{u}) = \max_{\substack{\mathbf{x} \geq 0 \\ |\mathbf{x}|=1}} r(\mathbf{x}) = \max_{\substack{\mathbf{x} \geq 0 \\ |\mathbf{x}|=1}} \min_{\mu} \frac{y_{\mu}}{x_{\mu}}$$

minimax-feladatnak van megoldása.

Erre a korlátozásra csupán azért van szükség, hogy a vizsgált tartományt korlátozottsá és zárttá tegyük, hiszen egyébként az $r(\mathbf{x})$ függvény (4.1.13) definiáló egyenletéből következik, hogy a függvény értéke nem változik, ha az \mathbf{x} vektort tetszőleges pozitív számmal megszorozzuk. Ha az egységgömbön mint korlátozott zárt tartományon az $r(\mathbf{x})$ függvény folytonos volna, akkor a maximum létezése biztosítva volna. Noha a tartomány minden olyan pontjában, ahol $\mathbf{x} > \mathbf{0}$, az $r(\mathbf{x})$ függvény folytonos, mégis azokban a pontokban, ahol a (4.1.14) összefüggés szerint \mathbf{x} valamelyik eleme 0, az $r(\mathbf{x})$ függvénynek szakadása lehet. Ezért bevezetjük azon \mathbf{z} vektorok halmazát, amelyeket az egységgömb pontjaihoz tartozó \mathbf{x} vektorokhoz a

$$(4.1.19) \quad \mathbf{z} = (\mathbf{E} + \mathbf{A})^{n-1} \mathbf{x}$$

összefüggéssel rendelünk hozzá. A \mathbf{z} vektoroknak ez a halmaza ugyancsak korlátozott és zárt, de a 4.1.1 tétel értelmében csupa pozitív vektorból áll.

Ha most az

$$(4.1.20) \quad r(\mathbf{x})\mathbf{x} \leq \mathbf{A}\mathbf{x}$$

egyenlőtlenség mindkét oldalát a pozitív elemű $(\mathbf{E} + \mathbf{A})^{n-1} > \mathbf{0}$ mátrixszal megszorozzuk, akkor (4.1.19) felhasználásával adódik:

$$(4.1.21) \quad r(\mathbf{x})\mathbf{z} \leq \mathbf{A}\mathbf{z}.$$

Az $r(\mathbf{z})$ függvény definícióját felhasználva (mely szerint $r(\mathbf{z})$ az a legnagyobb ϱ valós szám, amelyre a $\varrho\mathbf{z} \leq \mathbf{A}\mathbf{z}$ egyenlőtlenség fennáll), ebből következik

$$(4.1.22) \quad r(\mathbf{x}) \leq r(\mathbf{z}).$$

Az $r(\mathbf{x})$ függvény maximumának keresésekor tehát az egységgömb pontjaihoz tartozó \mathbf{x} vektorok halmaza a (4.1.19) összefüggés segítségével hozzájuk rendelt pozitív elemű \mathbf{z} vektorok halmazával helyettesíthető. Ezen a korlátozott és zárt halmazon az $r(\mathbf{x})$ függvény folytonos, tehát egy bizonyos $\mathbf{u} > \mathbf{0}$ vektorra felveszi a maximumát. Minden olyan $\mathbf{u} \geq \mathbf{0}$ vektort, amelyre

$$(4.1.23) \quad r(\mathbf{u}) = r,$$

extremális vektornak nevezzük.

4.1.2 A Frobenius-tételek

Ezen előkészítés után bebizonyítjuk az irreducibilis, nemnegatív elemű mátrixokra vonatkozó első alaptételt.

4.1.3 tétel (Frobenius I. tétele). *Bármely nemnegatív elemű irreducibilis \mathbf{A} mátrixnak mindig van egy maximális abszolút értékű, egyszeres, pozitív r sajátértéke, amelyhez tartozó sajátvektor minden eleme pozitív.*

Bizonyítás. Először megmutatjuk, hogy a (4.1.18) összefüggéssel definiált r szám az \mathbf{A} mátrix egy pozitív sajátértéke, és hogy minden extrémális vektor pozitív elemű és az \mathbf{A} mátrix r sajátértékéhez tartozó sajátvektor, azaz

$$(4.1.24) \quad r > 0, \quad \mathbf{u} > \mathbf{0}, \quad \mathbf{A}\mathbf{u} = r\mathbf{u}.$$

A csupa 1 elemből álló \mathbf{e} összegező vektorhoz tartozó $r(\mathbf{e})$ függvényérték pozitív, mert $r(\mathbf{e}) = \min_{\mu} \sum_{\nu=1}^n a_{\mu\nu}$, és egy irreducibilis mátrixnak nem lehet egyetlen, csupa 0 elemet tartalmazó sora sem. Tehát $r(\mathbf{e}) > 0$, viszont $r \geq r(\mathbf{e})$ miatt ebből

$$(4.1.25) \quad r > 0$$

következik.

Most indirekt módszerrel bebizonyítjuk, hogy r az \mathbf{A} mátrix sajátértéke és hogy a hozzá tartozó sajátvektor \mathbf{u} . Feltesszük tehát, hogy

$$(4.1.26) \quad \mathbf{A}\mathbf{u} - r\mathbf{u} \neq \mathbf{0}.$$

Mivel a (4.1.20) egyenlőtlenség miatt

$$(4.1.27) \quad \mathbf{A}\mathbf{u} - r\mathbf{u} \geq \mathbf{0},$$

tehát nemnegatív vektor, ezt az $(\mathbf{E} + \mathbf{A})^{n-1}$ mátrixszal megszorozva, a 4.1.1 tétel alapján pozitív vektort kapunk: $(\mathbf{E} + \mathbf{A})^{n-1}(\mathbf{A}\mathbf{u} - r\mathbf{u}) > \mathbf{0}$. Bevezetve a

$$(4.1.28) \quad \mathbf{v} = (\mathbf{E} + \mathbf{A})^{n-1}\mathbf{u}$$

vektort, az

$$(4.1.29) \quad \mathbf{A}\mathbf{v} - r\mathbf{v} > \mathbf{0}$$

egyenlőtlenség adódik. Ekkor azonban található olyan elég kicsiny pozitív ε , amellyel még fennáll $\mathbf{A}\mathbf{v} - (r + \varepsilon)\mathbf{v} > \mathbf{0}$. Mivel pedig $r(\mathbf{v})$ az a legnagyobb

szám, amelyre a (4.1.16) egyenlőtlenség teljesül, ezért ekkor $r(\mathbf{v}) \geq r + \varepsilon > r$ lenne, ami ellentmond az r szám (4.1.18) definíciójának. Következésképpen

$$(4.1.30) \quad \mathbf{A}\mathbf{u} = r\mathbf{u},$$

tehát a pozitív r szám valóban az \mathbf{A} mátrix sajátértéke, és a hozzá tartozó sajátvektor \mathbf{u} .

Most azt mutatjuk meg, hogy \mathbf{u} pozitív vektor. A (4.1.30) összefüggésből következik ugyanis, hogy

$$(4.1.31) \quad (\mathbf{E} + \mathbf{A})^{n-1}\mathbf{u} = (1 + r)^{n-1}\mathbf{u},$$

a 4.1.1 tétel szerint pedig $(\mathbf{E} + \mathbf{A})^{n-1}\mathbf{u} > \mathbf{0}$. Ezért a (4.1.28) alatt definiált \mathbf{v} vektor pozitív; (4.1.31) behelyettesítésével tehát

$$(4.1.32) \quad \mathbf{v} = (\mathbf{E} + \mathbf{A})^{n-1}\mathbf{u} = (1 + r)^{n-1}\mathbf{u} > \mathbf{0}$$

adódik, és innen közvetlenül megkapjuk az $\mathbf{u} > \mathbf{0}$ összefüggést, amit bizonyítani kellett.*

Be kell még látni, hogy az \mathbf{A} mátrix többi sajátértékének abszolút értéke nem nagyobb r -nél. Legyen α egy tetszőleges sajátérték és \mathbf{w} a hozzá tartozó sajátvektor, azaz

$$(4.1.33) \quad \mathbf{A}\mathbf{w} = \alpha\mathbf{w}.$$

Vegyük mindkét oldal elemeinek az abszolút értékét; ekkor (4.1.11) felhasználásával adódik

$$(4.1.34) \quad |\alpha|\mathbf{w}^+ \leq \mathbf{A}\mathbf{w}^+,$$

ahonnan – ismét $r(\mathbf{w}^+)$ és r definícióját felhasználva – az

$$(4.1.35) \quad |\alpha| \leq r(\mathbf{w}^+) \leq r$$

egyenlőtlenségre jutunk, tehát valóban nincs r -nél nagyobb abszolút értékű sajátérték.

Most azt látjuk be, hogy az r sajátértékhez csak egyetlen lineárisan független sajátvektor tartozik. Legyen ugyanis \mathbf{y} az r sajátértékhez tartozó tetszőleges sajátvektor:

$$(4.1.36) \quad \mathbf{A}\mathbf{y} = r\mathbf{y}.$$

*Ha az \mathbf{A} mátrix helyett az \mathbf{A}^T mátrixon alkalmazzuk a fenti bizonyítást, akkor ebből látható, hogy az \mathbf{A} mátrix r -hez tartozó bal oldali sajátvektora is pozitív.

Ismét felhasználva a (4.1.11) egyenlőtlenséget,

$$(4.1.37) \quad r\mathbf{y}^+ \leq \mathbf{A}\mathbf{y}^+.$$

A (4.1.35) összefüggésből következik, hogy $r \leq r(\mathbf{y}^+) \leq r$, így (4.1.23) alapján \mathbf{y}^+ extrémális vektor, és $\mathbf{y}^+ > \mathbf{0}$. Ezért az \mathbf{y} sajátvektor egyetlen eleme sem lehet 0, tehát

$$(4.1.38) \quad y_\mu \neq 0 \quad (\mu = 1, 2, \dots, n).$$

Ebből viszont következik, hogy az r sajátértékhez csak egyetlen lineárisan független sajátvektor tartozik. Ha ugyanis \mathbf{u}_1 és \mathbf{u}_2 lineárisan független sajátvektorok lennének, akkor biztosan található lenne olyan c_1 és c_2 szám, hogy az $\mathbf{y} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2$ sajátvektornak legalább egy eleme zérus volna, ami ellentmond a (4.1.38) feltételnek.

Végül bebizonyítjuk, hogy r nem lehet a karakterisztikus polinomnak többszörös gyöke; vagyis ha $D(\lambda)$ jelöli a karakterisztikus polinomot, akkor

$$(4.1.39) \quad D'(r) \neq 0.$$

Ehhez felhasználjuk a karakterisztikus mátrixra vonatkozó alábbi azonosságokat:

$$(4.1.40) \quad (\lambda\mathbf{E} - \mathbf{A})\text{adj}(\lambda\mathbf{E} - \mathbf{A}) \equiv D(\lambda)\mathbf{E}, \quad \text{adj}(\lambda\mathbf{E} - \mathbf{A})(\lambda\mathbf{E} - \mathbf{A}) \equiv D(\lambda)\mathbf{E}.$$

Mivel

$$(4.1.41) \quad \text{adj}(\lambda\mathbf{E} - \mathbf{A}) = [D_{\nu\mu}(\lambda)],$$

ahol $D_{\mu\nu}(\lambda)$ jelöli a karakterisztikus mátrix $\mu\nu$ indexű eleméhez tartozó $(n-1)$ -edrendű előjeles aldeterminánst, a (4.1.40) azonosságban λ helyére az r sajátértéket behelyettesítve

$$(4.1.42) \quad (r\mathbf{E} - \mathbf{A})[D_{\nu\mu}(r)] = \mathbf{0}, \quad [D_{\nu\mu}(r)](r\mathbf{E} - \mathbf{A}) = \mathbf{0}$$

adódik. Mivel az r sajátértékhez egyetlen lineárisan független $\mathbf{u} > \mathbf{0}$ sajátvektor tartozik, a $[D_{\nu\mu}(r)]$ mátrix rangja 1, tehát ez a mátrix az r -hez tartozó jobb és bal oldali sajátvektorból alkotott diádként írható fel. A (4.1.42) egyenlőségek értelmében a $[D_{\nu\mu}(r)]$ mátrix elemei mind azonos előjelű, zérustól különböző számok. Mivel pedig a karakterisztikus polinom deriváltjára érvényes a $D'(\lambda) = \sum_{\mu=1}^n D_{\mu\mu}(\lambda)$, összefüggés, ezért

$$(4.1.43) \quad D'(r) = \sum_{\mu=1}^n D_{\mu\mu}(r) \neq 0,$$

amit bizonyítani kellett. Megjegyezzük még, hogy mivel r a karakterisztikus polinom legnagyobb abszolút értékű, pozitív gyöke, ezért $D(\lambda)$ a $\lambda = r$ helyen monoton nő, tehát

$$(4.1.44) \quad D'(r) > 0.$$

Így a (4.1.43) összefüggésből következik, hogy

$$(4.1.45) \quad D_{\mu\mu}(r) > 0,$$

továbbá valamennyi indexpárra $D_{\nu\mu}(r) > 0$. Tehát

$$(4.1.46) \quad \text{adj}(r\mathbf{E} - \mathbf{A}) > \mathbf{0}. \quad \blacksquare$$

A most bizonyított 4.1.3 tétel nem tér ki arra, hogy az adott irreducibilis mátrixnak csak egyetlen vagy több maximális abszolút értékű sajátértéke van-e. Abban az esetben, ha több maximális abszolút értékű sajátértéke van, akkor a mátrix szerkezetére és spektrális tulajdonságaira vonatkozik a II. alaptétel.

Mielőtt erre rátérünk, bebizonyítunk egy segédtevélt, amelyet a II. alaptétel bizonyítása során használunk majd fel.

4.1.4 tétel. Tekintsük az n -edrendű \mathbf{A} és \mathbf{B} mátrixot és jelölje α_k , ill. β_k a sajátértékeiket. Tegyük fel, hogy $\mathbf{A} \geq \mathbf{0}$ irreducibilis, nemnegatív elemű mátrix az $r > 0$ maximális abszolút értékű, egyszeres sajátértékkel, és \mathbf{B} olyan (komplex elemű) mátrix, amelyre

$$(4.1.47) \quad \mathbf{B}^+ \leq \mathbf{A}.$$

A \mathbf{B} mátrix β_k sajátértékeire ekkor teljesül a

$$(4.1.48) \quad |\beta_k| \leq r \quad (k = 1, 2, \dots, n)$$

egyenlőtlenség, az egyenlőség pedig a \mathbf{B} mátrix valamely $\beta_j = re^{i\varphi_j}$ sajátértékére akkor és csak akkor áll fenn, ha

$$(4.1.49) \quad \mathbf{B} = e^{i\varphi_j} \mathbf{DAD}^{-1},$$

ahol \mathbf{D} olyan diagonálmátrix, amelynek diagonálelemei egységnyi abszolút értékűek, azaz $\mathbf{D}^+ = \mathbf{E}$.

Bizonyítás. Jelölje \mathbf{y}_k a \mathbf{B} mátrixnak a β_k sajátértékhez tartozó sajátvektorát, azaz

$$(4.1.50) \quad \mathbf{B}\mathbf{y}_k = \beta_k\mathbf{y}_k \quad (\beta_k \neq 0).$$

A (4.1.47) és a (4.1.50) összefüggésből következik, hogy

$$(4.1.51) \quad |\beta_k| \mathbf{y}_k^+ \leq \mathbf{B}^+ \mathbf{y}_k^+ \leq \mathbf{A} \mathbf{y}_k^+,$$

és innen

$$(4.1.52) \quad |\beta_k| \leq r_{\mathbf{A}}(\mathbf{y}_k^+) \leq r.$$

Ha a \mathbf{B} mátrix (4.1.49) alakú, akkor a (4.1.48)-ban fennáll az egyenlőség, tehát a feltétel elégséges.

Most megmutatjuk, hogy a feltétel szükséges is. Tegyük fel, hogy $|\beta_j| = r$; ekkor a (4.1.52) összefüggésből következik, hogy \mathbf{y}_j^+ az \mathbf{A} mátrix extrémális vektora, tehát $\mathbf{y}_j^+ > \mathbf{0}$ és \mathbf{y}_j^+ az \mathbf{A} mátrix r sajátértékéhez tartozó sajátvektor. Így a (4.1.51) összefüggésben az egyenlőségjel érvényes, vagyis

$$(4.1.53) \quad \mathbf{A} \mathbf{y}_j^+ = \mathbf{B}^+ \mathbf{y}_j^+ = r \mathbf{y}_j^+, \quad \mathbf{y}_j^+ > \mathbf{0}.$$

Mivel \mathbf{y}_j^+ pozitív elemű vektor, (4.1.47) alapján ebből következik

$$(4.1.54) \quad \mathbf{B}^+ = \mathbf{A}.$$

Tekintsük most az $\mathbf{y}_k^T = [y_{1k} \ y_{2k} \ \dots \ y_{nk}]$ sajátvektort, és írjuk fel ennek elemeit az alábbi alakban:

$$(4.1.55) \quad y_{\nu k} = |y_{\nu k}| e^{i\varphi_{\nu}} \quad (\nu = 1, 2, \dots, n).$$

Vezessük be a \mathbf{D} diagonálmátrixot:

$$(4.1.56) \quad \mathbf{D} = \langle e^{i\varphi_1} \ e^{i\varphi_2} \ \dots \ e^{i\varphi_n} \rangle.$$

Ezzel írható

$$(4.1.57) \quad \mathbf{y}_k = \mathbf{D} \mathbf{y}_k^+.$$

Helyettesítsük be ezt a (4.1.50) egyenletbe, és vegyük tekintetbe, hogy $k = j$ esetén $\beta_j = r e^{i\varphi_j}$, tehát $\mathbf{B} \mathbf{D} \mathbf{y}_j^+ = r e^{i\varphi_j} \mathbf{D} \mathbf{y}_j^+$, így az

$$(4.1.58) \quad \mathbf{F} = e^{-i\varphi_j} \mathbf{D}^{-1} \mathbf{B} \mathbf{D}$$

jelöléssel írható

$$(4.1.59) \quad \mathbf{F} \mathbf{y}_j^+ = r \mathbf{y}_j^+.$$

A (4.1.53) összefüggés figyelembevételével ekkor

$$(4.1.60) \quad \mathbf{F} \mathbf{y}_j^+ = \mathbf{B}^+ \mathbf{y}_j^+ = \mathbf{A} \mathbf{y}_j^+,$$

továbbá – mivel a (4.1.58) összefüggésből $\mathbf{F}^+ = \mathbf{B}^+$ következik – ezért a (4.1.54) egyenlőség felhasználásával az

$$(4.1.61) \quad \mathbf{F}^+ = \mathbf{A}$$

összefüggést kapjuk, és így, ezt a (4.1.60) egyenletbe behelyettesítve, $\mathbf{F}\mathbf{y}_j^+ = \mathbf{F}^+\mathbf{y}_j^+$. Tekintetbe véve, hogy \mathbf{y}_j^+ pozitív elemű vektor, a kapott összefüggés csak akkor állhat fenn, ha $\mathbf{F} = \mathbf{F}^+$. Végül ebbe behelyettesítve a (4.1.58) és (4.1.61) kifejezéseket, $e^{-i\varphi_j}\mathbf{D}^{-1}\mathbf{B}\mathbf{D} = \mathbf{A}$ adódik. Innen következik a $\mathbf{B} = e^{i\varphi_j}\mathbf{D}\mathbf{A}\mathbf{D}^{-1}$ összefüggés, és ezzel a 4.1.4 tételt bebizonyítottuk. ■

Ezek után a nemnegatív elemű irreducibilis mátrixokra vonatkozó II. alaptételt fogalmazzuk meg és bizonyítjuk be.

4.1.5 tétel (Frobenius II. tétele). *Ha a nemnegatív elemű irreducibilis \mathbf{A} mátrixnak összesen h maximális abszolút értékű sajátértéke van, akkor ezek a*

$$(4.1.62) \quad \lambda^h - r^h = 0$$

egyenlet gyökei, és az \mathbf{A} mátrix összes sajátértékeinek a rendszere – mint a komplex számsík pontthalmaza – a komplex sík $\frac{2\pi}{h}$ szögű elforgatásával szemben invariáns; továbbá az \mathbf{A} mátrix sorok és oszlopok ugyanolyan permutációjával az

$$(4.1.63) \quad \tilde{\mathbf{A}} = \begin{bmatrix} 0 & \tilde{\mathbf{A}}_{12} & 0 & \dots & 0 \\ 0 & 0 & \tilde{\mathbf{A}}_{23} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \tilde{\mathbf{A}}_{h-1,h} \\ \tilde{\mathbf{A}}_{h1} & 0 & 0 & \dots & 0 \end{bmatrix}$$

szimmetrikusan particionált h -adrendű hiperciklikus mátrixba vihető át.

Megjegyzés. Mint ismeretes, a $\lambda^h - r^h = 0$ egyenlet gyökei a h -adik egységgyökök r -szeresei, és ezért ezek a komplex számsíkon a 0 körüli r sugarú körbe rajzolt azon szabályos h -szög csúcsai, amelynek egyik csúcsa a pozitív valós tengelyen van, vagyis $\lambda_k = r \cdot e^{i\frac{2\pi k}{h}}$ ($k = 0, 1, \dots, h-1$).

Bizonyítás. Jelölje

$$(4.1.64) \quad \lambda_0 = re^{i\varphi_0}; \quad \lambda_1 = re^{i\varphi_1}; \dots; \lambda_{h-1} = re^{i\varphi_{h-1}} \\ (0 = \varphi_0 < \varphi_1 \leq \varphi_2 \leq \dots \leq \varphi_{h-1} < 2\pi)$$

az \mathbf{A} mátrix h számú maximális abszolút értékű sajátértékét, és alkalmazzuk a 4.1.4 tételt, ahol \mathbf{B} helyére is az \mathbf{A} mátrixot helyettesítjük. Ekkor β_j helyére λ_k ($k = 0, 1, \dots, h-1$) helyettesítendő, és így (4.1.49) alapján írható

$$(4.1.65) \quad \mathbf{A} = e^{i\varphi_k}\mathbf{D}_k\mathbf{A}\mathbf{D}_k^{-1},$$

ahol $\mathbf{D}_k^+ = \mathbf{E}$. Legyen \mathbf{u} az \mathbf{A} mátrix r sajátértékéhez tartozó pozitív elemű sajátvektor, azaz

$$(4.1.66) \quad \mathbf{A}\mathbf{u} = r\mathbf{u}, \quad \mathbf{u} > \mathbf{0}.$$

Ha most bevezetjük az

$$(4.1.67) \quad \mathbf{y}_k = \mathbf{D}_k\mathbf{u}, \quad \mathbf{y}_k^+ = \mathbf{u} > \mathbf{0}$$

összefüggéssel definiált \mathbf{y}_k vektort és ezzel jobbról megszorozzuk a (4.1.65) egyenletet, akkor (4.1.66) felhasználásával adódik $\mathbf{A}\mathbf{D}_k\mathbf{u} = e^{i\varphi_k}\mathbf{D}_k\mathbf{A}\mathbf{u} = e^{i\varphi_k}\mathbf{D}_kr\mathbf{u}$, azaz

$$(4.1.68) \quad \mathbf{A}\mathbf{y}_k = \lambda_k\mathbf{y}_k \quad (k = 0, 1, 2, \dots, h-1).$$

Ebből az összefüggésből kiolvasható, hogy a (4.1.67) alatti \mathbf{y}_k vektorok az \mathbf{A} mátrix (4.1.64) alatti λ_k sajátértékéhez tartozó sajátvektorok.

A (4.1.65) összefüggésből következik, hogy nemcsak a $\lambda_0 = r$, hanem a többi $\lambda_1, \lambda_2, \dots, \lambda_{h-1}$ sajátérték is egyszeres sajátérték. Ezért az \mathbf{y}_k sajátvektorok és így a \mathbf{D}_k mátrixok is egy állandó szorzótól eltekintve egyértelműen meghatározottak. Normáljuk tehát a \mathbf{D}_k mátrixokat oly módon, hogy mindegyikben az első diagonálem 1 legyen; így ezek a mátrixok egyértelműek, ahol $\mathbf{D}_0 = \mathbf{E}$ és $\mathbf{y}_0 = \mathbf{u} > \mathbf{0}$.

Ha most a (4.1.65) összefüggés jobb oldalán \mathbf{A} helyére – ugyanezen összefüggés alapján – az $\mathbf{A} = e^{i\varphi_j}\mathbf{D}_j\mathbf{A}\mathbf{D}_j^{-1}$ kifejezést helyettesítjük, akkor

$$(4.1.69) \quad \mathbf{A} = e^{i(\varphi_k + \varphi_j)}\mathbf{D}_k\mathbf{D}_j\mathbf{A}\mathbf{D}_j^{-1}\mathbf{D}_k^{-1}.$$

Ebből következik, hogy a $\mathbf{D}_k\mathbf{D}_j\mathbf{u}$ vektor az \mathbf{A} mátrix $re^{i(\varphi_k + \varphi_j)}$ sajátértékéhez tartozó sajátvektora, vagyis az \mathbf{A} mátrixnak van olyan $re^{i\varphi_l}$ sajátértéke, amelyre

$$(4.1.70) \quad re^{i\varphi_l} = re^{i(\varphi_k + \varphi_j)},$$

és a hozzá tartozó sajátvektor

$$(4.1.71) \quad \mathbf{D}_l\mathbf{u} = \mathbf{D}_k\mathbf{D}_j\mathbf{u}.$$

Ez annyit jelent, hogy az \mathbf{A} mátrix maximális abszolút értékű sajátértékének az a tulajdonságuk, hogy bármelyik kettő argumentumának az összege egy (harmadik) maximális abszolút értékű sajátérték argumentuma. Ebből következik, hogy ha egy mátrixnak éppen h számú különböző maximális abszolút értékű sajátértéke van, és ez a maximális abszolút érték r , akkor ezek a sajátértékek a komplex számsíkon a 0 körüli r sugarú körön egy h oldalú

szabályos sokszög csúcspontjaiban helyezkednek el. És mivel az egyik sajátérték pozitív valós szám, ez egyértelműen meghatározza az összes maximális abszolút értékű sajátértéket; ezek az r számnak és a h -adik egységgyököknek a szorzatai. Eszerint a sajátértékek argumentumai:

$$(4.1.72) \quad \varphi_k = \frac{2k\pi}{h} \quad (k = 0, 1, 2, \dots, h-1);$$

az

$$(4.1.73) \quad \varepsilon = e^{i\varphi_1} = e^{i\frac{2\pi}{h}}$$

jelöléssel

$$(4.1.74) \quad e^{i\varphi_k} = \varepsilon^k \quad (k = 0, 1, 2, \dots, h-1),$$

vagyis a λ_k maximális abszolút értékű sajátértékek az alábbi alakban írhatók:

$$(4.1.75) \quad \lambda_k = r\varepsilon^k \quad (k = 0, 1, 2, \dots, h-1).$$

A (4.1.71) összefüggésből következik, hogy a megfelelő sajátvektorok elemeinek argumentumát meghatározó \mathbf{D}_k diagonálmátrixokra fennáll

$$(4.1.76) \quad \mathbf{D}_l = \mathbf{D}_k \mathbf{D}_j,$$

ami a sajátvektorok elemeinek ugyanazt a tulajdonságát fejezi ki, amit a (4.1.70) összefüggés alapján a sajátértékekre megmutattunk. Eszerint bármely két sajátvektor megfelelő elemének az argumentumát összeadva, egy (harmadik) sajátvektor megfelelő elemének argumentumát kapjuk. A \mathbf{D}_k mátrix elemei tehát szintén a k -adik egységgyökök, így (4.1.73), (4.1.74) és (4.1.76) figyelembevételével

$$(4.1.77) \quad \mathbf{D}_k = \mathbf{D}_1^k.$$

A (4.1.65) összefüggésből $k = 1$ esetén

$$(4.1.78) \quad \mathbf{A} = \varepsilon \mathbf{D} \mathbf{A} \mathbf{D}^{-1},$$

ami azt jelenti, hogy az \mathbf{A} mátrix az

$$(4.1.79) \quad \varepsilon = e^{i\frac{2\pi}{h}}$$

számmal való szorzással önmagával hasonló mátrixba transzformálódik, vagyis a mátrix összes sajátértékeinek a rendszere a (4.1.79) alatti számmal szorozva – tehát a komplex síkon $\frac{2\pi}{h}$ szöggel elforgatva – önmagába megy át.

Mivel a \mathbf{D}_k mátrixok elemei, és így \mathbf{D}_1 elemei is h -adik egységgyökök, ezért

$$(4.1.80) \quad \mathbf{D}_1^h = \mathbf{E}.$$

Ha az \mathbf{A} mátrixon egy alkalmas permutáló mátrixszal ortogonális transzformációt végzünk, akkor a \mathbf{D}_1 diagonálmátrix elemei úgy rendezhetők át, hogy az azonos egységgyökök egymás mellé kerüljenek, és az argumentumok monoton nemcsökkenő sorozatot alkossanak. Jelölje \mathbf{D} az így átrendezett diagonálmátrixot, vagyis

$$(4.1.81) \quad \mathbf{D} = \langle \varepsilon^{j_0} \mathbf{E}_{\alpha_0} \quad \varepsilon^{j_1} \mathbf{E}_{\alpha_1} \dots \varepsilon^{j_{s-1}} \mathbf{E}_{\alpha_{s-1}} \rangle,$$

ahol $0 = j_0 < j_1 < j_2 < \dots < j_{s-1} < h$; j_ν egész számokat jelent és $s \leq h$. A (4.1.81) diagonálmátrixban álló α_ν rendű egység mátrixoknak megfelelően particionálva az átrendezett $\tilde{\mathbf{A}}$ mátrixot, azt s -edrendű hipermátrix alakjában írhatjuk fel:

$$(4.1.82) \quad \tilde{\mathbf{A}} = [\tilde{\mathbf{A}}_{pq}] \quad (p, q = 1, 2, \dots, s).$$

A (4.1.78) egyenletből ekkor

$$(4.1.83) \quad \varepsilon \tilde{\mathbf{A}}_{pq} = \frac{\varepsilon^{j_{q-1}}}{\varepsilon^{j_{p-1}}} \tilde{\mathbf{A}}_{pq} \quad (p, q = 1, 2, \dots, s).$$

Innen minden pq indexpárra

$$(4.1.84) \quad \text{vagy} \quad \frac{\varepsilon^{j_{q-1}}}{\varepsilon^{j_{p-1}}} = \varepsilon \quad \text{vagy} \quad \tilde{\mathbf{A}}_{pq} = \mathbf{0}.$$

Tekintsük először az első bloksort, vagyis legyen $p = 1$. Mivel az $\tilde{\mathbf{A}}_{12}, \tilde{\mathbf{A}}_{13}, \dots, \tilde{\mathbf{A}}_{1s}$ blokkok nem lehetnek mind zérussal egyenlők, ezért az

$$(4.1.85) \quad \frac{\varepsilon^{j_1}}{\varepsilon^{j_0}}, \frac{\varepsilon^{j_2}}{\varepsilon^{j_0}}, \dots, \frac{\varepsilon^{j_{s-1}}}{\varepsilon^{j_0}}$$

számok egyikének ε -nal kell egyenlőnek lennie. Mivel $j_0 = 0$, azaz $\varepsilon^{j_0} = 1$, ezért csak j_1 lehet 1, és így az $\frac{\varepsilon^{j_1}}{\varepsilon^{j_0}} = \varepsilon$ összefüggés alapján

$$(4.1.86) \quad \tilde{\mathbf{A}}_{12} \neq \mathbf{0}, \quad \tilde{\mathbf{A}}_{11} = \tilde{\mathbf{A}}_{13} = \tilde{\mathbf{A}}_{14} = \dots = \tilde{\mathbf{A}}_{1s} = \mathbf{0}.$$

Legyen most a (4.1.84) egyenletben $p = 2$. Ekkor a (4.1.85) hányadosok helyett az $\frac{\varepsilon^{j_2}}{\varepsilon^{j_1}}, \frac{\varepsilon^{j_3}}{\varepsilon^{j_1}}, \dots, \frac{\varepsilon^{j_{s-1}}}{\varepsilon^{j_1}}$ hányadosokat tekintjük, ahol $j_1 = 1$. Ezek közül ismét csak az első lehet egyenlő ε -nal, ha $j_2 = 2$, és innen

$$(4.1.87) \quad \tilde{\mathbf{A}}_{23} \neq \mathbf{0}, \quad \tilde{\mathbf{A}}_{21} = \tilde{\mathbf{A}}_{22} = \tilde{\mathbf{A}}_{24} = \dots = \tilde{\mathbf{A}}_{2s} = \mathbf{0}.$$

Ezt az eljárást folytatva, $j_1 = 1, j_2 = 2, \dots, j_{s-1} = s - 1$ adódik, tehát $\tilde{\mathbf{A}}$ particionált alakja:

$$\tilde{\mathbf{A}} = \begin{bmatrix} \mathbf{0} & \tilde{\mathbf{A}}_{12} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \tilde{\mathbf{A}}_{23} \dots & \mathbf{0} & \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \tilde{\mathbf{A}}_{s-1,s} \\ \tilde{\mathbf{A}}_{s1} & \tilde{\mathbf{A}}_{s2} & \tilde{\mathbf{A}}_{s3} \dots & \tilde{\mathbf{A}}_{ss} & \end{bmatrix}.$$

Az utolsó blokkSORra, vagyis $p = s$ esetén a (4.1.83) egyenletből következik $\varepsilon = \frac{\varepsilon^{q-1}}{\varepsilon^{s-1}}$, azaz $\varepsilon = \varepsilon^{q-s}$, ami csak olyan $q-s$ értékekre állhat fenn, amelyeket a h számmal osztva, maradékul 1-et kapunk*. Ebből következik $q = 1$ és $s = h$, és innen $\tilde{\mathbf{A}}_{s2} = \tilde{\mathbf{A}}_{s3} = \dots = \tilde{\mathbf{A}}_{ss} = \mathbf{0}$ ($s = h$).

A \mathbf{D} diagonálmátrix alakja tehát

$$(4.1.88) \quad \mathbf{D} = \langle \mathbf{E}_{\alpha_0} \quad \varepsilon \mathbf{E}_{\alpha_1} \quad \varepsilon^2 \mathbf{E}_{\alpha_2} \dots \varepsilon^{h-1} \mathbf{E}_{\alpha_{h-1}} \rangle,$$

ahol \mathbf{E}_{α_ν} az α_ν rendű egység mátrix, $\sum_{\nu=0}^{h-1} \alpha_\nu = n$, és az átrendezett \mathbf{A} mátrix particionált alakja valóban a (4.1.63) szerinti hiperciklikus alak. ■

4.1.3 A Frobenius-tételek következményei

A Frobenius-tételek segítségével bebizonyítható néhány további fontos tétel az irreducibilis, nemnegatív elemű mátrixokra.

4.1.6 tétel. *Legyen \mathbf{A} irreducibilis, nemnegatív elemű mátrix, melynek karakterisztikus polinomja $D_n(\lambda) \equiv |\lambda \mathbf{E} - \mathbf{A}|$, a karakterisztikus mátrix adjungáltja $\text{adj}(\lambda \mathbf{E} - \mathbf{A}) = [D_{\nu\mu}(\lambda)]$, a $D_{\nu\mu}(\lambda)$ polinomok legnagyobb közös osztója $\theta(\lambda)$, és a karakterisztikus mátrix redukált adjungáltja*

$$(4.1.89) \quad \mathbf{F}(\lambda) = \frac{\text{adj}(\lambda \mathbf{E} - \mathbf{A})}{\theta(\lambda)} = \frac{1}{\theta(\lambda)} [D_{\nu\mu}(\lambda)].$$

Ha r jelöli a $D_n(\lambda)$ karakterisztikus polinom legnagyobb abszolút értékű pozitív valós gyökét, akkor

$$(4.1.90) \quad \mathbf{F}(r) > \mathbf{0},$$

azaz a redukált adjungált mátrix a $\lambda = r$ helyen pozitív.

*Számelméleti terminológiával ezt úgy fejezhetjük ki, hogy $q - s$ kongruens 1 modulo h , vagy jelöléssel: $q - s \equiv 1 \pmod{h}$.

Bizonyítás. A Frobenius-tételek bizonyítása során kiderült, hogy a karakterisztikus mátrix adjungáltja a $\lambda = r$ helyen pozitív, azaz

$$(4.1.91) \quad D_{\nu\mu}(r) > 0$$

(lásd a (4.1.46) összefüggést). Innen következik, hogy a $D_{\nu\mu}(\lambda)$ polinomok legnagyobb közös osztója a $\lambda = r$ helyen nem lehet zérus: $\theta(r) \neq 0$.

Mivel továbbá a $D_n(\lambda)$ karakterisztikus polinom osztható a $\theta(\lambda)$ polinommal, ez utóbbinak a gyökei csak olyan sajátértékek lehetnek, amelyekre $|\lambda_k| < r$. S mivel $\theta(\lambda)$ legmagasabb fokú tagjának együtthatója 1, ezért

$$(4.1.92) \quad \theta(r) > 0.$$

A (4.1.91) és (4.1.92) összefüggést behelyettesítve a redukált adjungált mátrix (4.1.89) definiáló egyenletébe, (4.1.90) adódik. Ezzel a tételt bebizonyítottuk. ■

A (4.1.91) egyenlőtlenség lehetővé teszi, hogy az r legnagyobb abszolút értékű sajátértékre alsó és felső korlátot állapítsunk meg. Erre vonatkozik a következő tétel.

4.1.7 tétel. *Ha az irreducibilis, nemnegatív elemű \mathbf{A} mátrix sorelemeinek összegét s_k jelöli, azaz*

$$(4.1.93) \quad s_k = \sum_{\nu=1}^n a_{k\nu} \quad (k = 1, 2, \dots, n),$$

továbbá

$$(4.1.94) \quad s = \min_{1 \leq k \leq n} s_k, \quad S = \max_{1 \leq k \leq n} s_k,$$

akkor az r maximális sajátérték az

$$(4.1.95) \quad s \leq r \leq S$$

egyenlőtlenséggel becsülhető, ahol akár a bal, akár a jobb oldali egyenlőségjel csak akkor érvényes, ha $s = S$, azaz ha az s_1, s_2, \dots, s_n sorelemösszegek egyenlőek egymással.

Bizonyítás. Felhasználva az $\{\text{adj}(\lambda \mathbf{E} - \mathbf{A})\}(\lambda \mathbf{E} - \mathbf{A}) \equiv D_n(\lambda) \mathbf{E}$ azonosságot, helyettesítsük be λ helyére az r sajátértéket:

$$(4.1.96) \quad \{\text{adj}(r \mathbf{E} - \mathbf{A})\}(r \mathbf{E} - \mathbf{A}) = \mathbf{0},$$

és szorozzuk meg a kapott egyenletet jobbról a csupa 1 elemből álló összegező

$$(4.1.97) \quad \mathbf{e} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

vektorral. Mivel

$$(4.1.98) \quad (r\mathbf{E} - \mathbf{A})\mathbf{e} = \begin{bmatrix} r - s_1 \\ r - s_2 \\ \vdots \\ r - s_n \end{bmatrix},$$

a (4.1.96) egyenletből következik, hogy $\{\text{adj}(r\mathbf{E} - \mathbf{A})\}(r\mathbf{E} - \mathbf{A})\mathbf{e} = \mathbf{0}$, ill. (4.1.98) helyettesítésével

$$(4.1.99) \quad \sum_{k=1}^n D_{\mu k}(r)(r - s_k) = 0.$$

A (4.1.91) egyenlőtlenség alapján a $D_{\mu k}(r)$ mennyiségek pozitívak, és így a (4.1.99) egyenlet csak akkor állhat fenn, ha az $r - s_k$ különbségek nem jeltartók, vagyis ha fennáll a (4.1.95) egyenlőtlenség. Ebből a tétel további állítása már következik. ■

A Frobenius-tételek speciális esetként magukban foglalják a pozitív elemű mátrixok spektrális (azaz a sajátértékek és sajátvektorok) tulajdonságaira vonatkozó Perron-féle* tételt.

4.1.8 tétel (Perron tétele). *Minden pozitív elemű mátrixnak egyetlen legnagyobb abszolút értékű r sajátértéke van, amely pozitív és egyszeres sajátérték, és amelyhez tartozó sajátvektor elemei pozitívak.*

Bizonyítás. A tétel közvetlenül adódik a Frobenius-tételekből, ugyanis ha a mátrixnak lenne még egy másik r abszolút értékű sajátértéke, akkor azt sorok és oszlopok ugyanolyan permutációjával a (4.1.63) szerinti hiperciklikus alakra lehetne hozni, ami azonban lehetetlen, mert a mátrix valamennyi eleme pozitív. A tétel többi állítása a Frobenius-tételek alapján nyilvánvaló. ■

A következő tételek is közvetlenül adódnak a Frobenius-tételekből.

4.1.9 tétel. *Nemnegatív elemű, irreducibilis mátrixnak nem lehet két lineárisan független, nemnegatív elemű sajátvektora.*

Bizonyítás. A tételt indirekt úton bizonyítjuk. Legyen a nemnegatív elemű, irreducibilis \mathbf{A} mátrix maximális abszolút értékű sajátértéke a pozitív r szám, amelyhez tartozó pozitív elemű jobb oldali sajátvektor \mathbf{u} . Tegyük fel, hogy az \mathbf{A} mátrix α sajátértékéhez az \mathbf{u} vektortól lineárisan független, nemnegatív elemű \mathbf{v} sajátvektor tartozik, vagyis

$$(4.1.100) \quad \mathbf{A}\mathbf{v} = \alpha\mathbf{v}; \quad \mathbf{v} \neq \mathbf{0}, \quad \mathbf{v} \geq \mathbf{0}.$$

*O. Perron (1880–1975) német matematikus.

Mivel a Frobenius-tételek értelmében r egyszeres sajátérték, ezért

$$(4.1.101) \quad \alpha \neq r.$$

Jelölje \mathbf{w}^T az \mathbf{A} mátrix r sajátértékéhez tartozó bal oldali sajátvektort (mivel \mathbf{w} az \mathbf{A} mátrix transzponáltjának jobb oldali sajátvektora, ezért a Frobenius-tételek értelmében szintén pozitív elemű), azaz

$$(4.1.102) \quad \mathbf{w}^T \mathbf{A} = r \mathbf{w}^T; \quad \mathbf{w} > \mathbf{0}.$$

Szorozzuk meg a (4.1.100) egyenletet balról a \mathbf{w}^T vektorral, a (4.1.102) egyenletet jobbról a \mathbf{v} vektorral és képezzük a két egyenlet különbségét: $(\alpha - r) \mathbf{w}^T \mathbf{v} = 0$; ebből (4.1.101) miatt következik $\mathbf{w}^T \mathbf{v} = 0$, ami azonban $\mathbf{w} > \mathbf{0}$, $\mathbf{v} \geq \mathbf{0}$, de $\mathbf{v} \neq \mathbf{0}$ miatt nem lehetséges és ezzel a tételt bebizonyítottuk. ■

Frobenius I. tételének bizonyításában a legnagyobb sajátértéket a (4.1.18) minimax-feladat megoldásaként jellemeztük:

$$r = \max_{\mathbf{x} \geq \mathbf{0}} \min_{\mu} \frac{y_{\mu}}{x_{\mu}}, \quad \text{ahol } \mathbf{y} = \mathbf{A}\mathbf{x}.$$

Mivel $r(\mathbf{x}) = \min_{\mu} \frac{y_{\mu}}{x_{\mu}}$, $\mathbf{y} = \mathbf{A}\mathbf{x}$, vagyis $r(\mathbf{x})$ volt az a legnagyobb ϱ szám, amelyre még teljesült $\varrho \mathbf{x} \leq \mathbf{A}\mathbf{x}$, az r legnagyobb sajátértéket az alábbi maximumfeladat megoldása szolgáltatta: $r = \max_{\mathbf{x} \geq \mathbf{0}} r(\mathbf{x})$.

Hasonlóképpen megfogalmazható a következő feladat. Legyen $\tilde{r}(\mathbf{x})$ az a legkisebb σ szám, amelyre tetszőleges $\mathbf{x} \geq \mathbf{0}$ esetén teljesül

$$(4.1.103) \quad \sigma \mathbf{x} \geq \mathbf{A}\mathbf{x},$$

vagyis legyen

$$(4.1.104) \quad \tilde{r}(\mathbf{x}) = \max_{\mu} \frac{y_{\mu}}{x_{\mu}}, \quad \mathbf{y} = \mathbf{A}\mathbf{x}.$$

Ha valamelyik μ -re $x_{\mu} = 0$ és $y_{\mu} \neq 0$, akkor $\tilde{r}(\mathbf{x}) = +\infty$. Ugyanúgy, ahogy Frobenius I. tételének bizonyítása során megmutattuk, hogy az $r(\mathbf{x})$ függvény egy bizonyos pozitív elemű \mathbf{u} vektornál felveszi a maximumát (ez volt a mátrix r legnagyobb sajátértéke), itt is belátható, hogy a (4.1.104) alatti $\tilde{r}(\mathbf{x})$ függvény egy pozitív \mathbf{v} vektornál felveszi a minimumát, amelyet \tilde{r} -rel jelölünk. Most bebizonyítjuk a következő tételt.

4.1.10 tétel. Az

$$(4.1.105) \quad \tilde{r} = \min_{\mathbf{x} \geq \mathbf{0}} r(\mathbf{x}) = \min_{\mathbf{x} \geq \mathbf{0}} \max_{\mu} \frac{y_{\mu}}{x_{\mu}}, \quad \mathbf{y} = \mathbf{A}\mathbf{x}$$

minimax feladat megoldásaként definiált \tilde{r} szám egyenlő az \mathbf{A} mátrix r legnagyobb sajátértékével, és az a $\mathbf{v} \geq \mathbf{0}$, $\mathbf{v} \neq \mathbf{0}$ vektor, amelyre $\tilde{r}(\mathbf{x})$ felveszi a minimumát, az \mathbf{A} mátrix r sajátértékéhez tartozó sajátvektor.

Bizonyítás. Az \tilde{r} szám és a \mathbf{v} vektor definíciójából következik, hogy

$$(4.1.106) \quad \tilde{r}\mathbf{v} - \mathbf{A}\mathbf{v} \geq \mathbf{0} \quad (\mathbf{v} \geq \mathbf{0}, \mathbf{v} \neq \mathbf{0}).$$

Azt kell belátnunk, hogy ebben az összefüggésben az egyenlőségjel érvényes; ezt indirekt módszerrel bizonyítjuk. Tegyük fel tehát, hogy $\tilde{r}\mathbf{v} - \mathbf{A}\mathbf{v} \geq \mathbf{0}$, $\tilde{r}\mathbf{v} - \mathbf{A}\mathbf{v} \neq \mathbf{0}$. A 4.1.1 tételből következik

$$(4.1.107) \quad (\mathbf{E} + \mathbf{A})^{n-1}(\tilde{r}\mathbf{v} - \mathbf{A}\mathbf{v}) > \mathbf{0}.$$

Vezessük be a pozitív elemű

$$(4.1.108) \quad \mathbf{z} = (\mathbf{E} + \mathbf{A})^{n-1}\mathbf{v} > \mathbf{0}$$

vektort, és helyettesítsük be a (4.1.107) egyenlőtlenségbe. Ezzel $\tilde{r}\mathbf{z} > \mathbf{A}\mathbf{z}$ adódik, vagyis elég kicsiny pozitív ε -ra $(\tilde{r} - \varepsilon)\mathbf{z} > \mathbf{A}\mathbf{z}$ ($\mathbf{z} > \mathbf{0}$), ami ellentmond az \tilde{r} szám definíciójának. Így tehát érvényes

$$(4.1.109) \quad \mathbf{A}\mathbf{v} = \tilde{r}\mathbf{v}$$

Ekkor azonban $\mathbf{z} = (\mathbf{E} + \mathbf{A})^{n-1}\mathbf{v} = (1 + \tilde{r})^{n-1}\mathbf{v}$, és $\mathbf{z} > \mathbf{0}$ -ból következik, hogy $\mathbf{v} > \mathbf{0}$.

A 4.1.9 tétel felhasználásával innen adódik, hogy $\tilde{r} = r$ (mivel az \mathbf{A} mátrixnak nem lehet két lineárisan független pozitív sajátvektora). Ezzel a tételt bebizonyítottuk. ■

Az r számot jellemezhetjük tehát a következő két extrémális érték bármelyikeként:

$$(4.1.110) \quad r = \max_{\mathbf{x} \geq \mathbf{0}} \min_{\mu} \frac{y_{\mu}}{x_{\mu}} = \min_{\mathbf{x} \geq \mathbf{0}} \max_{\mu} \frac{y_{\mu}}{x_{\mu}};$$

beláttuk, hogy az az \mathbf{x} vektor, amelyre (4.1.110) teljesül, az r sajátértékhez tartozó egyetlen sajátvektor.

Az r sajátérték (4.1.110) definíciójából következik az alábbi fontos egyenlőtlenség:

$$(4.1.111) \quad \min_{1 \leq \mu \leq n} \frac{y_{\mu}}{x_{\mu}} \leq r \leq \max_{1 \leq \mu \leq n} \frac{y_{\mu}}{x_{\mu}},$$

ahol $\mathbf{A}\mathbf{x} = \mathbf{y}$, $\mathbf{x} \geq \mathbf{0}$, $\mathbf{x} \neq \mathbf{0}$.

Mivel a (4.1.13) összefüggésben $r(\mathbf{x})$ a maximumát, ill. a (4.1.104) összefüggésben $\tilde{r}(\mathbf{x})$ a minimumát csak pozitív elemű sajátvektornál veheti fel, érvényes a következő tétel.

4.1.11 tétel. *Ha \mathbf{A} irreducibilis, nemnegatív elemű mátrix, amelynek legnagyobb abszolút értékű sajátértéke r , akkor az $r(\mathbf{z}) \leq \mathbf{A}\mathbf{z}$, $\mathbf{z} \geq \mathbf{0}$, $\mathbf{z} \neq \mathbf{0}$, ill. az $r(\mathbf{z}) \geq \mathbf{A}\mathbf{z}$, $\mathbf{z} \geq \mathbf{0}$, $\mathbf{z} \neq \mathbf{0}$ egyenlőtlenségből $\mathbf{A}\mathbf{z} = r\mathbf{z}$, $\mathbf{z} > \mathbf{0}$ következik.*

4.2 REDUCIBILIS MÁTRIXOK

Ebben a szakaszban a legáltalánosabb nemnegatív elemű mátrixok néhány tulajdonságával foglalkozunk. Bevezetjük a (reducibilis) nemnegatív elemű mátrixok normálalakját, és ennek felhasználásával feltételeket adunk pozitív elemű sajátvektor létezésére. A kapott eredmények segítségével irreducibilis mátrixokra vonatkozóan is újabb tételekhez jutunk. Ezek között az a legfontosabb, amely bizonyos mátrixok inverzének a pozitivitására vonatkozik.

4.2.1 A reducibilis mátrixok alaptétele

Az alapvető tétel bizonyításához felhasználjuk a konvergens mátrixsorozatok fogalmát.

4.2.1 definíció. Ha az $\mathbf{A}^{(m)} = [a_{ij}^{(m)}]$ ($m = 1, 2, \dots$) mátrixsorozat elemeit a konvergens $a_{ij}^{(1)}, a_{ij}^{(2)}, \dots, a_{ij}^{(m)}, \dots$ ($i, j = 1, 2, \dots, n$) számsorozatok alkotják, akkor azt mondjuk, hogy az

$$\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \dots, \mathbf{A}^{(m)}, \dots$$

mátrixok konvergens mátrixsorozatot alkotnak, és ha

$$\lim_{m \rightarrow \infty} a_{ij}^{(m)} = a_{ij} \quad (i, j = 1, 2, \dots, n),$$

akkor az $\mathbf{A}^{(m)}$ mátrixsorozat határértékét a következőképpen definiáljuk:

$$\lim_{m \rightarrow \infty} \mathbf{A}^{(m)} = \mathbf{A}, \quad \text{ahol } \mathbf{A} = [a_{ij}].$$

E definíció segítségével tetszőleges, nemnegatív elemű mátrix, pozitív elemű mátrixok valamilyen sorozatának határértékeként értelmezhető, azaz tetszőleges $\mathbf{A} \geq \mathbf{0}$ előállítható az

$$(4.2.1) \quad \mathbf{A} = \lim_{m \rightarrow \infty} \mathbf{A}^{(m)}, \quad \mathbf{A}^{(m)} > \mathbf{0} \quad (m = 1, 2, \dots)$$

határértékként. Várható ezért, hogy az irreducibilis mátrixok néhány spektrális tulajdonsága bizonyos gyengébb formában érvényes reducibilis mátrixokra is.

4.2.1 tétel. Minden nemnegatív elemű \mathbf{A} mátrixnak van egy olyan nemnegatív (valós) r sajátértéke, amely nem kisebb, mint a mátrix bármely más sajátértékének abszolút értéke, és amelyhez tartozik egy nemnegatív elemű \mathbf{u} sajátvektor:

$$(4.2.2) \quad \mathbf{A}\mathbf{u} = r\mathbf{u}, \quad \mathbf{u} \geq \mathbf{0}, \quad \mathbf{u} \neq \mathbf{0},$$

a $(\lambda \mathbf{E} - \mathbf{A})$ karakterisztikus mátrix adjungáltjának $D_{ji}(\lambda)$ elemeire pedig $\lambda \geq r$ esetén

$$(4.2.3) \quad D_{ji}(\lambda) \geq 0, \quad \frac{d}{d\lambda} D_{ji}(\lambda) \geq 0$$

érvényes; vagyis mind a karakterisztikus mátrix adjungáltja, $\text{adj}(\lambda \mathbf{E} - \mathbf{A}) = [D_{ji}(\lambda)]$, mind ennek λ szerinti deriváltja nemnegatív elemű mátrix minden olyan λ -ra, amelynél nagyobb abszolút értékű sajátérték nem létezik.

Bizonyítás. Értelmezzük az \mathbf{A} mátrixot a (4.2.1) szerinti határértékként és jelölje a pozitív elemű $\mathbf{A}^{(m)}$ mátrix legnagyobb abszolút értékű sajátértékét $r^{(m)}$, a hozzá tartozó pozitív elemű normált sajátvektorát pedig $\mathbf{u}^{(m)}$; tehát

$$(4.2.4) \quad \mathbf{A}^{(m)} \mathbf{u}^{(m)} = r^{(m)} \mathbf{u}^{(m)}, \quad \mathbf{u}^{(m)} > \mathbf{0}, \quad \mathbf{u}^{(m)\top} \mathbf{u}^{(m)} = 1.$$

A 4.2.1 definícióból következik, hogy létezik a $\lim_{m \rightarrow \infty} r^{(m)} = r$ határérték, amely az \mathbf{A} mátrix egy sajátértéke. Mivel $r^{(m)} > 0$ és $r^{(m)} > |\lambda_i^{(m)}|$, ahol $\lambda_i^{(m)}$ az $\mathbf{A}^{(m)}$ mátrix $r^{(m)}$ -től különböző tetszőleges sajátértéke, a határátmenetet elvégezve és figyelembe véve, hogy pozitív számok sorozatának határértéke biztosan nemnegatív szám – tehát az $r^{(m)} - |\lambda_i^{(m)}| > 0$ különbségek sorozatának határértéke lehet zérus is – ezért $r \geq 0$, $r \geq |\lambda_i|$ következik, ahol λ_i az \mathbf{A} mátrix tetszőleges sajátértéke. Ugyanígy, a határátmenetet elvégezve, a (4.1.46) egyenlőtlenség helyett azt kapjuk, hogy

$$(4.2.5) \quad D_{ji}(r) \geq 0.$$

Az $\mathbf{u}^{(m)}$ normált sajátvektorok sorozatából kiválasztható olyan $\mathbf{u}^{(m_p)}$ ($p = 1, 2, \dots$) részsorozat, amely egy normált vektorhoz konvergál. Ha most a (4.2.4) sorozat m_p indexszel jellemzett részsorozatán keresztül képezzük a határértéket, akkor $\mathbf{A}\mathbf{u} = r\mathbf{u}$, $\mathbf{u} \geq \mathbf{0}$, $\mathbf{u} \neq \mathbf{0}$ adódik.

A (4.2.3) egyenlőtlenséget teljes indukcióval bizonyítjuk. Ha $n = 1$, akkor az egyenlőtlenség triviális módon teljesül, hiszen ekkor $\text{adj}(\lambda \mathbf{E} - \mathbf{A}) \equiv \mathbf{E}$ és $\frac{d}{d\lambda} \text{adj}(\lambda \mathbf{E} - \mathbf{A}) \equiv \mathbf{0}$. Tegyük most fel, hogy (4.2.3) érvényes $1, 2, \dots, (n-1)$ -re; be fogjuk látni, hogy ebből következik érvényessége n -re is.

Fejtsük ki a $D(\lambda) = |\lambda \mathbf{E} - \mathbf{A}|$ determinánst utolsó sora és oszlopa szerint. Jelölje $D_{(j \ n)}^{(i \ n)}(\lambda)$ a $\lambda \mathbf{E} - \mathbf{A}$ karakterisztikus mátrixnak azt az $(n-2)$ -edrendű minorát, amelyet i -edik és n -edik sorának, valamint j -edik és n -edik oszlopának elhagyásával kapunk. Ekkor a determináns

$$(4.2.6) \quad D(\lambda) = (\lambda - a_{nn})D_{nn}(\lambda) - \sum_{i,j=1}^{n-1} D_{(j \ n)}^{(i \ n)}(\lambda) a_{in} a_{nj}.$$

Jelölje r_n a $D_{nn}(\lambda)$ polinom legnagyobb abszolút értékű nemnegatív gyökét. Helyettesítsünk a (4.2.6) összefüggésbe $\lambda = r_n$ értéket, és vegyük figyelembe,

hogy az indukciós feltétel következtében $D\binom{i}{j} \binom{n}{n}(r_n) \geq 0$ ($i, j = 1, 2, \dots, n-1$). Így $D(r_n) \leq 0$. Másrészt, mivel a $D(\lambda)$ polinomban λ legmagasabb fokú hatványának együtthatója pozitív, ezért a polinom elég nagy pozitív λ értékek esetén biztosan pozitív. Ebből következik, hogy r_n vagy gyöke a $D(\lambda)$ polinomnak, vagy van a $D(\lambda)$ polinomnak olyan gyöke, amelyik nagyobb, mint r_n . Mindkét esetben $r_n \leq r$.

Jelölje r_k a $D_{kk}(\lambda)$ polinom legnagyobb valós gyökét. Mivel sorok és oszlopok ugyanolyan permutációjával bármelyik főminor a $D_{nn}(\lambda)$ sarokminor helyére hozható, ezért

$$(4.2.7) \quad r_k \leq r \quad (k = 1, 2, 3, \dots, n).$$

Az $(n-1)$ -edrendű $D_{ji}(\lambda)$ minorokat deriválva,

$$(4.2.8) \quad \frac{d}{d\lambda} D_{ji}(\lambda) = \sum_k D\binom{j}{i} \binom{k}{k}(\lambda) \quad k \neq i, \quad k \neq j.$$

A $D\binom{j}{i} \binom{k}{k}(\lambda)$ polinomok úgy is felfoghatók, mint a $(\lambda \mathbf{E} - \mathbf{A})$ karakterisztikus mátrix k -adik sorának és oszlopának törlésével kapott minormátrix adjungáltjának elemei. Az indukciós feltevés szerint ezekre $D\binom{j}{i} \binom{k}{k}(\lambda) \geq 0$, ha $\lambda \geq r_k$ ($k = 1, 2, \dots, n$), és ezért $\lambda \geq r$ esetén a (4.2.7) és (4.2.8) összefüggésből

$$(4.2.9) \quad \frac{d}{d\lambda} D_{ji}(\lambda) \geq 0,$$

a (4.2.5) és (4.2.9) összefüggésből pedig $D_{ji}(\lambda) \geq 0$ következik. Ezzel a tételt bebizonyítottuk. ■

A 4.2.1 tétel segítségével bebizonyíthatók az alábbi tételek.

4.2.2 tétel. *Ha \mathbf{A} nemnegatív elemű mátrix, melynek legnagyobb abszolút értékű sajátértéke r , és $\mathbf{F}(\lambda)$ a redukált adjungáltja, akkor*

$$(4.2.10) \quad \mathbf{F}(\lambda) \geq \mathbf{0}, \quad \text{ha } \lambda \geq r.$$

Bizonyítás. A redukált adjungált 3.2.1 definíciója szerint

$$(4.2.11) \quad \mathbf{F}(\lambda) = \frac{\text{adj}(\lambda \mathbf{E} - \mathbf{A})}{\theta(\lambda)},$$

ahol $\theta(\lambda)$ az adjungált elemeinek legnagyobb közös osztója. Mivel pedig $\theta(\lambda)$ a $D(\lambda)$ karakterisztikus polinomnak is osztója, és legmagasabb fokú tagjának együtthatója pozitív, ezért

$$(4.2.12) \quad \theta(\lambda) > 0, \quad \text{ha } \lambda > r.$$

A (4.2.3), (4.2.11) és a (4.2.12) összefüggésből már következik (4.2.10). Ezzel a tételt bizonyítottuk. ■

4.2.3 tétel. *Ha \mathbf{A} nemnegatív elemű, irreducibilis mátrix, amelynek legnagyobb abszolút értékű sajátértéke r , akkor*

$$(4.2.13) \quad \text{adj}(\lambda \mathbf{E} - \mathbf{A}) > \mathbf{0} \quad \text{és} \quad \mathbf{F}(\lambda) > \mathbf{0}, \quad \text{ha} \quad \lambda \geq r,$$

azaz minden olyan λ -ra, amelynél nagyobb sajátérték nem létezik, a karakterisztikus mátrix adjungáltja és redukált adjungáltja pozitív elemű mátrix.

Bizonyítás. A (4.1.46) összefüggés szerint $D_{ji}(r) > 0$, és a 4.2.1 tétel szerint $\frac{d}{d\lambda} D_{ji}(\lambda) \geq 0$, ha $\lambda \geq r$, tehát $D_{ji}(\lambda) > 0$, ha $\lambda \geq r$, vagyis

$$(4.2.14) \quad \text{adj}(\lambda \mathbf{E} - \mathbf{A}) > \mathbf{0}, \quad \text{ha} \quad \lambda \geq r.$$

Az $\mathbf{F}(\lambda) > \mathbf{0}$ egyenlőtlenség a (4.2.11), (4.2.12) és (4.2.14) összefüggésekből következik. ■

4.2.4 tétel. *Ha \mathbf{A} nemnegatív elemű, irreducibilis mátrix, amelynek legnagyobb abszolút értékű sajátértéke r , akkor*

$$(4.2.15) \quad (\lambda \mathbf{E} - \mathbf{A})^{-1} > \mathbf{0}, \quad \text{ha} \quad \lambda > r.$$

Bizonyítás. Az állítás közvetlenül belátható az inverz előállítására vonatkozó (1.2.34) összefüggés felhasználásával:

$$(\lambda \mathbf{E} - \mathbf{A})^{-1} = \frac{\text{adj}(\lambda \mathbf{E} - \mathbf{A})}{D(\lambda)}.$$

Ugyanis $\lambda > r$ esetén a 4.2.3 tétel értelmében $\text{adj}(\lambda \mathbf{E} - \mathbf{A})$ pozitív, a $D(\lambda)$ polinom pedig azért pozitív, mert a legmagasabb fokú tagjának együtthatója pozitív. ■

4.2.2 Reducibilis nemnegatív elemű mátrix normálalakja

A következőkben a reducibilis mátrixok normálalakjával fogunk megismerkedni. E fejezet bevezetésében megadott 4.0.1 definíció szerint akkor nevezünk egy mátrixot reducibilisnek, ha sorainak és oszlopainak ugyanolyan permutációjával az alábbi, szimmetrikusan négy blokkra particionált hipermátrixra hozható: $\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$, ahol tehát \mathbf{A}_{11} és \mathbf{A}_{22} kvadratikus

mátrixok. Ha ezek valamelyike szintén reducibilis, akkor az eljárást folytatva, végül is az

$$(4.2.16) \quad \mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{A}_{31} & \mathbf{A}_{32} & \mathbf{A}_{33} & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{A}_{s1} & \mathbf{A}_{s2} & \dots & \dots & \mathbf{A}_{ss} \end{bmatrix}$$

szimmetrikusan particionált hipermátrixra jutunk; ezt *hiper-háromszögmátrixnak* fogjuk nevezni.

4.2.2 definíció. Az $\mathbf{A}_{kk} \neq \mathbf{0}$ **diagonálblokkot izoláltnak** nevezzük, ha $\mathbf{A}_{kj} = \mathbf{0}$ ($j = 1, 2, \dots, k-1, k+1, \dots, s$), azaz ha a diagonálisban levő \mathbf{A}_{kk} blokkon kívül a k -adik blokk sor összes többi blokkja zérus.

A (4.2.16) alakra hozott reducibilis mátrix a blokk sorok és blokkoszlopok ugyanolyan permutációjával úgy rendezhető át, hogy izolált blokkjai – ha vannak – kerüljenek a fődiagonális első blokkjai helyére. Így jutunk a reducibilis mátrixok normálalakjának fogalmához.

4.2.3 definíció. Az

$$(4.2.17) \quad \mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{A}_g & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{A}_{g+1,1} & \mathbf{A}_{g+1,2} & \dots & \mathbf{A}_{g+1,g} & \mathbf{A}_{g+1} & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{A}_{s1} & \mathbf{A}_{s2} & \dots & \mathbf{A}_{sg} & \mathbf{A}_{s,g+1} & \dots & \mathbf{A}_s \end{bmatrix}$$

alakú hipermátrixot, ahol $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_s$ irreducibilis mátrixok és az

$$\mathbf{A}_{i1}, \mathbf{A}_{i2}, \dots, \mathbf{A}_{i,i-1} \quad (i = g+1, g+2, \dots, s)$$

blokkok közül legalább az egyik zérustól különböző, a **reducibilis mátrix normálalakjának** nevezzük.

Megjegyzés. A definícióban előforduló g és s számok a nemnegatív elemű mátrix invariánsai. (Irreducibilis mátrixok esetén nyilván $g = s = 1$.)

A normálalak felhasználásával bizonyítható a reducibilis mátrixok alábbi spektrális tulajdonsága.

4.2.5 tétel. A nemnegatív elemű \mathbf{A} mátrix legnagyobb abszolút értékű sajátértékéhez akkor és csak akkor tartozik pozitív elemű sajátvektor, ha a (4.2.17) normálalakban

(a) valamennyi $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_g$ blokk sajátértéke r , és

(b) az $\mathbf{A}_{g+1}, \mathbf{A}_{g+2}, \dots, \mathbf{A}_s$ blokkok közül egyiknek sem sajátértéke r .

Bizonyítás. A feltételek szükségesek. Legyen ugyanis a mátrix legnagyobb abszolút értékű r sajátértékéhez tartozó \mathbf{u} sajátvektor pozitív, és particionáljuk a mátrixot a (4.2.17) normálalaknak megfelelően. Ekkor az

$$(4.2.18) \quad \mathbf{A}\mathbf{u} = r\mathbf{u} \quad (\mathbf{u} > \mathbf{0})$$

egyenlet szétesik az

$$(4.2.19) \quad \mathbf{A}_i\mathbf{u}_i = r\mathbf{u}_i \quad (i = 1, 2, \dots, g)$$

és a

$$(4.2.20) \quad \sum_{k=1}^{j-1} \mathbf{A}_{jk}\mathbf{u}_k + \mathbf{A}_j\mathbf{u}_j = r\mathbf{u}_j \quad (j = g+1, \dots, s)$$

egyenletekre. A (4.2.19) egyenletekből következik, hogy az $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_g$ blokkok mindegyikének sajátértéke r , a (4.2.20) egyenletekből pedig következik

$$(4.2.21) \quad \mathbf{A}_j\mathbf{u}_j \leq r\mathbf{u}_j \quad \text{és} \quad \mathbf{A}_j\mathbf{u}_j \neq r\mathbf{u}_j \quad (j = g+1, \dots, s).$$

Jelölje r_j az \mathbf{A}_j blokk legnagyobb sajátértékét ($j = g+1, \dots, s$). A (4.1.111) és a (4.2.21) egyenlőtlenség alapján ekkor felírható

$$(4.2.22) \quad r_j \leq \max_i \frac{(\mathbf{A}_j\mathbf{u}_j)_i}{u_{ij}} \leq r.$$

A 4.1.11 tétel alapján azonban az $r_j = r$ egyenlőség ellentmond a (4.2.21) alatti $\mathbf{A}_j\mathbf{u}_j \neq r\mathbf{u}_j$ egyenlőtlenségnek, ezért

$$(4.2.23) \quad r_j < r \quad (j = g+1, \dots, s).$$

A feltételek elégségesek is. Legyen r az \mathbf{A}_i mátrixok legnagyobb sajátértéke $i = 1, 2, \dots, g$ esetén, és legyen a (4.2.23) egyenlőtlenség érvényes $i = g+1, \dots, s$ esetén. Ha a (4.2.18) egyenletet a (4.2.19) és (4.2.20) egyenletekre bontjuk, akkor az \mathbf{A}_i mátrixokhoz a (4.2.19) egyenletek szerint $i = 1, 2, \dots, g$ esetén pozitív \mathbf{u}_i sajátvektorok tartoznak. Így a (4.2.20) egyenletből

$$(4.2.24) \quad \mathbf{u}_j = (r\mathbf{E} - \mathbf{A}_j)^{-1} \sum_{k=1}^{j-1} \mathbf{A}_{jk}\mathbf{u}_k \quad (j = g+1, \dots, s).$$

Mivel $r_j < r$ ($j = g+1, \dots, s$), a (4.2.15) összefüggés alapján

$$(4.2.25) \quad (r\mathbf{E} - \mathbf{A}_j)^{-1} > \mathbf{0} \quad (j = g+1, \dots, s).$$

Megmutatjuk, hogy a (4.2.24) összefüggéssel meghatározott $\mathbf{u}_{g+1}, \dots, \mathbf{u}_s$ vektorok pozitívak. Ezek ugyanis pozitívak, ha az $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{j-1}$ vektorok pozitívak, hiszen akkor $\sum_{k=1}^{j-1} \mathbf{A}_{jk} \mathbf{u}_k \geq \mathbf{0}$, $\sum_{k=1}^{j-1} \mathbf{A}_{jk} \mathbf{u}_k \neq \mathbf{0}$, és így a (4.2.25) egyenlőtlenség felhasználásával a (4.2.24) egyenletből csakugyan következik $\mathbf{u}_j > \mathbf{0}$ ($j = g+1, \dots, s$). Az $\mathbf{u} = [\mathbf{u}_i]$ ($i = 1, 2, \dots, s$) hipervektor tehát valóban az \mathbf{A} mátrix r sajátértékéhez tartozó sajátvektor. ■

4.2.6 tétel. *A nemnegatív elemű \mathbf{A} mátrix legnagyobb abszolút értékű r sajátértékéhez akkor és csak akkor tartozik pozitív elemű jobb és bal oldali sajátvektor, ha a mátrixot sorainak és oszlopainak ugyanolyan permutációjával az*

$$(4.2.26) \quad \mathbf{A} = \langle \mathbf{A}_1 \quad \mathbf{A}_2 \dots \mathbf{A}_s \rangle$$

hiperdiagonál alakra lehet hozni; ekkor az \mathbf{A}_k blokkok ($k = 1, 2, \dots, s$) olyan irreducibilis mátrixok, amelyek mindegyikének sajátértéke r .

Bizonyítás. A szükségességet indirekt úton bizonyítjuk. Tegyük fel, hogy az \mathbf{A} mátrix r legnagyobb abszolút értékű sajátértékéhez pozitív jobb és bal oldali sajátvektor tartozik. Ekkor a 4.2.5 tétel értelmében az \mathbf{A} mátrix a (4.2.17) alakú normálalakra hozható, ahol az $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_g$ blokkok legnagyobb sajátértéke r , és az $\mathbf{A}_{g+1}, \mathbf{A}_{g+2}, \dots, \mathbf{A}_s$ ($g < s$) blokkok legnagyobb sajátértéke kisebb, mint r . Az \mathbf{A} mátrix bal oldali sajátvektora azonban az \mathbf{A}^\top mátrix jobb oldali sajátvektora, tehát a feltevés alapján az

$$(4.2.27) \quad \mathbf{A}^\top = \begin{bmatrix} \mathbf{A}_1^\top \dots \mathbf{0} & \mathbf{A}_{g+1,1}^\top \dots \mathbf{A}_{s,1}^\top \\ \dots & \dots \\ \mathbf{0} \dots \mathbf{A}_g^\top & \mathbf{A}_{g+1,g}^\top \dots \mathbf{A}_{s,g}^\top \\ \mathbf{0} \dots \mathbf{0} & \mathbf{A}_{g+1}^\top \dots \mathbf{A}_{s,g+1}^\top \\ \dots & \dots \\ \mathbf{0} \dots \mathbf{0} & \mathbf{0} \dots \mathbf{A}_s^\top \end{bmatrix}$$

mátrix (r sajátértékéhez tartozó) jobb oldali sajátvektora pozitív. Az \mathbf{A}^\top mátrix szerkezetéből viszont látható, hogy \mathbf{A}_s izolált blokk, legnagyobb sajátértéke tehát nem lehet kisebb, mint r ; vagyis a $g < s$ feltevés ellentmondásra vezetett.

A feltétel elégségessége nyilvánvaló.

Ezek alapján kimondhatjuk a következő tételt.

4.2.7 tétel. *Ha egy nemnegatív elemű mátrix legnagyobb abszolút értékű sajátértéke egyszeres, és a hozzá tartozó jobb és bal oldali sajátvektor pozitív, akkor a mátrix irreducibilis.*

4.3 PRIMITÍV ÉS IMPRIMITÍV MÁTRIXOK

Frobenius II. tétele – mint láttuk – azt mondja ki, hogy minden nemnegatív elemű, irreducibilis mátrixot sorainak és oszlopainak ugyanolyan permutációjával olyan hiperciklikus mátrix alakra lehet hozni, amelynek – mint hipermátrixnak – a rendszáma megegyezik a mátrix legnagyobb abszolút értékű különböző sajátértékeinek számával. Ezzel kapcsolatosan bevezetjük a primitív és az imprimitív mátrix fogalmát, majd néhány ezekre vonatkozó fontos tételt bizonyítunk be.

4.3.1 definíció. *Ha a nemnegatív elemű, irreducibilis mátrixnak egyetlen legnagyobb abszolút értékű sajátértéke van, akkor **primitív mátrixnak** nevezzük; ha $h > 1$ számú legnagyobb abszolút értékű sajátértéke van, akkor **h indexű imprimitív mátrixnak** nevezzük.*

Az imprimitív mátrix indexének meghatározására vonatkozik a következő tétel.

4.3.1 tétel. *Legyen a nemnegatív elemű, irreducibilis mátrix karakterisztikus polinomja*

$$(4.3.1) \quad D(\lambda) = \lambda^n + a_1\lambda^{n_1} + a_2\lambda^{n_2} + \dots + a_t\lambda^{n_t},$$

ahol $n > n_1 > n_2 > \dots > n_t$ és $a_1 \neq 0, a_2 \neq 0, \dots, a_t \neq 0$. Ekkor az imprimitív mátrix h indexe az

$$(4.3.2) \quad n - n_1, \quad n_1 - n_2, \quad \dots, \quad n_{t-1} - n_t$$

különbségek legnagyobb közös osztója.

Bizonyítás. Frobenius II. tétele értelmében az adott mátrix spektruma (vagyis az összes sajátértéke) a komplex számsíkon a 0 pont körüli $\frac{2\pi}{h}$ nagyságú szöggel való elforgatással önmagába megy át. Léteznie kell tehát egy olyan $g(u)$ polinomnak, amellyel a (4.3.1) karakterisztikus polinom előállítható $D(\lambda) = g(\lambda^h)\lambda^{n_t}$ alakban. Innen következik, hogy h a (4.3.2) különbségek közös osztója. Egyúttal azonban e különbségek d -vel jelölt legnagyobb közös osztója, mert a spektrum a $\frac{2\pi}{d}$ nagyságú szöggel való elforgatással nem változik, ami $h < d$ esetén lehetetlen. ■

A következő tétel a primitív mátrixok egy fontos tulajdonságát fejezi ki; ez egyúttal szükséges és elégséges feltételt ad arra, hogy egy mátrix primitív legyen.

4.3.2 tétel. *A nemnegatív elemű \mathbf{A} mátrix akkor és csak akkor primitív, ha valamely hatványa pozitív elemű, vagyis ha létezik olyan p természetes szám, amelyre*

$$(4.3.3) \quad \mathbf{A}^p > \mathbf{0} \quad (p \geq 1).$$

Bizonyítás. A feltétel elégséges. Ugyanis ha $\mathbf{A}^p > \mathbf{0}$, akkor \mathbf{A} irreducibilis, hiszen ha \mathbf{A} reducibilis volna, akkor ebből az következne, hogy \mathbf{A}^p is reducibilis volna. Továbbá az \mathbf{A} mátrixra $h = 1$, mivel $h > 1$ esetén a pozitív elemű \mathbf{A}^p mátrixnak h számú maximális abszolút értékű $\lambda_1^p, \lambda_2^p, \dots, \lambda_h^p$ sajátértéke lenne – ezek abszolút értéke r^p –, ami azonban ellentmondana a pozitív elemű mátrixokra vonatkozó 4.1.8 Perron-féle tételnek.

A feltétel szükséges is. Legyen \mathbf{A} primitív mátrix és legyen az egyetlen legnagyobb abszolút értékű sajátértéke $\lambda_1 = r$; így a többi sajátértékre $|\lambda_i| < r$ ($i = 2, \dots, s$). Jelölje a minimálpolinomot

$$\Delta(\lambda) \equiv (\lambda - \lambda_1)(\lambda - \lambda_2)^{\gamma_2} \dots (\lambda - \lambda_s)^{\gamma_s},$$

és alkalmazzuk a mátrixfüggvények előállítására vonatkozó (3.4.7) képletet. Ekkor a gyökhelyekhez tartozó Hermite-féle alappolinomok segítségével felírható

$$(4.3.4) \quad \mathbf{A}^p = p! \sum_{k=1}^s \left\{ \frac{\lambda_k^p}{p!} H_{k0}(\mathbf{A}) + \frac{\lambda_k^{p-1}}{(p-1)!} H_{k1}(\mathbf{A}) + \dots + \frac{\lambda_k^{p-\gamma_k+1}}{(p-\gamma_k+1)!} H_{k,\gamma_k-1}(\mathbf{A}) \right\}, \quad \text{ha } p \geq \gamma_k - 1.$$

Figyelembe véve, hogy $\lambda_1 = r$ egyszeres gyöke a karakterisztikus egyenletnek, tehát a (3.4.19) képlet szerint

$$(4.3.5) \quad H_{10}(\mathbf{A}) = \frac{1}{\Delta'(r)} \left\{ \frac{\text{adj}(\lambda \mathbf{E} - \mathbf{A})}{\theta(\lambda)} \right\}_{\lambda=r} = \frac{1}{\Delta'(r)} \mathbf{F}(r),$$

a (4.3.4) összefüggés

$$(4.3.6) \quad \mathbf{A}^p = \frac{1}{\Delta'(r)} \mathbf{F}(r) \cdot r^p + p! \sum_{k=2}^s \sum_{\nu=1}^{\gamma_k-1} \frac{\lambda_k^{p-\nu}}{(p-\nu)!} H_{k\nu}(\mathbf{A})$$

alakban írható. Mivel a $\lambda_1 = r$ sajátértéken kívül az összes többi sajátérték abszolút értéke kisebb, mint r : $|\lambda_k| < r$ ($k = 2, 3, \dots, s$), a (4.3.6) összefüggésből

$$(4.3.7) \quad \lim_{p \rightarrow \infty} \frac{\mathbf{A}^p}{r^p} = \frac{1}{\Delta'(r)} \mathbf{F}(r).$$

A 4.2.3 tétel értelmében azonban $\mathbf{F}(r) > \mathbf{0}$, továbbá, mivel a minimálpolinom legnagyobb valós gyöke r , tehát $\Delta'(r) > 0$; innen következik

$$(4.3.8) \quad \lim_{p \rightarrow \infty} \frac{\mathbf{A}^p}{r^p} > \mathbf{0},$$

és így valóban található olyan p természetes szám, amelyre már $\mathbf{A}^p > \mathbf{0}$. Ezzel a tételt bebizonyítottuk. ■

4.3.3 tétel. *Ha \mathbf{A} nemnegatív elemű irreducibilis mátrix, és e mátrixnak valamely q egész kitevős hatványa reducibilis, akkor \mathbf{A}^q a sorainak és oszlopainak ugyanolyan permutációjával hiperdiagonál alakra hozható:*

$$(4.3.9) \quad \mathbf{P}\mathbf{A}^q\mathbf{P}^\top = \langle \mathbf{A}_1 \quad \mathbf{A}_2 \dots \mathbf{A}_d \rangle;$$

itt az $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_d$ blokkok irreducibilis mátrixok, amelyeknek maximális abszolút értékű sajátértékei megegyeznek. Ha az \mathbf{A} mátrix h indexű imprimitív mátrix, akkor a d szám q és h legnagyobb közös osztója.

Bizonyítás. Mivel \mathbf{A} irreducibilis, ezért a Frobenius-tételek értelmében az r maximális abszolút értékű sajátértékéhez tartozó jobb és bal oldali sajátvektorok pozitívak. Ekkor azonban az \mathbf{A}^q mátrix r^q maximális abszolút értékű sajátértékéhez tartozó sajátvektorok is pozitívak. Ezért a 4.2.6 tétel alkalmazásával az \mathbf{A}^q mátrix a (4.3.9) alakra hozható, ahol az $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_d$ blokkok olyan irreducibilis mátrixok, amelyeknek r^q maximális abszolút értékű sajátértékei megegyeznek. Az \mathbf{A} mátrixnak viszont h számú maximális abszolút értékű sajátértéke van:

$$r, \quad r\varepsilon, \dots, r\varepsilon^{h-1}, \quad \text{ahol } \varepsilon = e^{\frac{2\pi}{h}i} \quad (\text{itt } i \text{ a képzetes egység}).$$

Ebből következik, hogy az \mathbf{A}^q mátrixnak is h maximális abszolút értékű sajátértéke van: $r^q, \quad r^q\varepsilon^q, \dots, r^q\varepsilon^{q(h-1)}$; ezek közül d darab megegyezik egymással, ami csak akkor lehetséges, ha d a q és a h számok legnagyobb közös osztója. Ezzel a tételt bebizonyítottuk. ■

Speciális esetként adódik a következő két tétel.

4.3.4 tétel. *Primitív mátrix minden pozitív kitevőjű hatványa irreducibilis és primitív.*

A tétel közvetlenül adódik az előzőből, $h = 1$ helyettesítéssel. ■

4.3.5 tétel. *Ha \mathbf{A} egy h indexű imprimitív mátrix, akkor h -adik hatványa, \mathbf{A}^h , szétesik h számú primitív mátrixra, amelyeknek maximális abszolút értékű sajátértékeik megegyeznek.*

A tétel a 4.3.3 tételből $q = h$ helyettesítéssel adódik. ■

4.4 SZTOCHASZTIKUS MÁTRIXOK

Ebben a szakaszban speciális nemnegatív elemű mátrixoknak, a sztochasztikus mátrixoknak a tulajdonságait vizsgáljuk.

4.4.1 Alapfogalmak és alapvető tételek

4.4.1 definíció. Sztochasztikus mátrixnak nevezzük az olyan nemnegatív elemű mátrixot, amelyben minden sor elemeinek összege 1. Tehát $\mathbf{P} = [p_{\mu\nu}]$ sztochasztikus mátrix, ha

$$p_{\mu\nu} \geq 0 \quad \text{és} \quad \sum_{\nu=1}^n p_{\mu\nu} = 1 \quad (\mu = 1, 2, \dots, n)$$

teljesül.*

A sztochasztikus mátrixok jelentőségét a valószínűségi számításban, közelebbről a homogén Markov-láncok elméletében való alkalmazásuk adja meg (lásd [4], [16]). A következőképpen jutunk a (véges) Markov-lánc fogalmához. Tekintsünk egy fizikai rendszert, amely a t_0, t_1, t_2, \dots időpontokban n különböző lehetséges állapot,

$$S_1, S_2, \dots, S_n$$

valamelyikében található. Jelölje $p_{\mu}^{(k)}$ annak valószínűségét, hogy a rendszer a t_k időpontban az S_{μ} állapotban van, $p_{\mu\nu}(t_k)$ pedig annak a valószínűségét, hogy ha a rendszer a t_{k-1} időpontban az S_{μ} állapotban volt, akkor a t_k időpontban az S_{ν} állapotba kerül. A $p_{\mu\nu}(t_k)$ átmenetvalószínűségek mátrixa: $\mathbf{P}(t_k) = [p_{\mu\nu}(t_k)]$, meghatározza az n lehetséges állapothoz tartozó Markov-láncot. Ha az átmenetvalószínűségek függetlenek a t_k időponttól, vagyis $p_{\mu\nu}(t_k) = p_{\mu\nu}$ ($k = 1, 2, \dots$), akkor a Markov-láncot homogénnek nevezzük. Ha tudjuk, hogy a kezdeti időpontban a rendszer $p_{\mu}^{(0)}$ ($\mu = 1, 2, \dots, n$) valószínűséggel található az S_{μ} állapotban, akkor annak valószínűsége, hogy

t_1 időpontban az S_{ν} állapotban legyen, nyilván $p_{\nu}^{(1)} = \sum_{\mu=1}^n p_{\mu}^{(0)} p_{\mu\nu}$, azaz –

$$(4.4.1) \quad \mathbf{p}^{(1)\top} = \mathbf{p}^{(0)\top} \mathbf{P}.$$

Ha a $\mathbf{p}^{(N)}$ vektor $p_{\nu}^{(N)}$ eleme jelöli annak a valószínűségét, hogy a rendszer a t_N időpontban az S_{ν} állapotban található, akkor a (4.4.1) összefüggés ismételt alkalmazásával a

$$(4.4.2) \quad \mathbf{p}^{(N)\top} = \mathbf{p}^{(n-1)\top} \mathbf{P} = \mathbf{p}^{(N-2)\top} \mathbf{P}^2 = \dots = \mathbf{p}^{(0)\top} \mathbf{P}^N$$

összefüggésre jutunk. A mondottakból következik, hogy az n számú állapothoz tartozó homogén Markov-lánc egyértelműen meghatározott, ha ismerjük

*A teljesség kedvéért megjegyezzük, hogy a sztochasztikus mátrix definíciójában olykor azt is megkövetelik, hogy a mátrixnak nem lehet csupa zérus elemet tartalmazó oszlopa.

az átmenetvalószínűségeket \mathbf{P} mátrixát és a $\mathbf{p}^{(0)}$ kezdeti valószínűségeloszlást. A \mathbf{P}^N mátrix az N lépéses átmenetvalószínűségeket mátrixa. Ugyanis homogén Markov-lánc esetében a \mathbf{P}^N mátrixhatvány $\mu\nu$ indexű eleme annak valószínűsége, hogy a rendszer az S_μ állapotból N „lépés” után (vagyis bármely t_k időponttól a t_{k+N} időpontig) az S_ν állapotba kerül.

A sztochasztikus mátrixoknak ez a tulajdonsága indokolja, hogy részletesebben foglalkozzunk velük. Mivel a sztochasztikus mátrixok speciális, nemnegatív elemű mátrixok, ezért a nemnegatív mátrixok tulajdonságai rájuk is érvényesek, tehát vizsgálatukban ezekből a tulajdonságokból indulhatunk ki.

Mindenekelőtt bizonyítjuk a sztochasztikus mátrixok legnagyobb abszolút értékű sajátértékére és a hozzá tartozó jobb és bal oldali sajátvektorra vonatkozó alábbi tételt.

4.4.1 tétel. *Adott nemnegatív elemű mátrix akkor és csak akkor sztochasztikus, ha $r = 1$ sajátértéke, és ha az ehhez a sajátértékhez tartozó (jobb oldali) sajátvektor minden eleme 1. Az $r = 1$ sajátérték a sztochasztikus mátrix legnagyobb abszolút értékű sajátértéke.*

Bizonyítás. A sztochasztikus mátrix 4.4.1 definíciójából közvetlenül adódik, hogy a feltétel szükséges. Jelölje ugyanis

$$(4.4.3) \quad \mathbf{e}^T = [1 \quad 1 \quad 1 \dots 1]$$

a csupa 1 elemből álló vektort; ekkor a definíció értelmében nyilván

$$(4.4.4) \quad \mathbf{P}\mathbf{e} = \mathbf{e},$$

és innen kiolvasható, hogy \mathbf{e} az $r = 1$ sajátértékhez tartozó sajátvektor. Az $r = 1$ sajátérték egyúttal a legnagyobb abszolút értékű sajátérték, ugyanis a 4.1.7 tétel értelmében a nemnegatív elemű mátrixok legnagyobb abszolút értékű sajátértéke a sorelemek összegének maximumánál nem nagyobb és a minimumánál nem kisebb. Mivel pedig a sorelemek összege sztochasztikus mátrixok esetén 1, tehát a legnagyobb abszolút értékű sajátérték szükségképpen 1. A feltétel azonban elégséges is: a (4.4.4) összefüggésből ugyanis azonnal következik, hogy

$$\sum_{\nu=1}^n p_{\mu\nu} = 1 \quad (\mu = 1, 2, \dots, n),$$

tehát \mathbf{P} sztochasztikus mátrix. ■

Megjegyzés. Irreducibilis sztochasztikus mátrix $r = 1$ sajátértékéhez tartozó \mathbf{v}^T bal oldali sajátvektorra – amely a Frobenius-tétel értelmében szintén pozitív elemű vektor – a sajátvektorok biortogonalitása folytán fennáll a $\mathbf{v}^T\mathbf{e} = 1$ összefüggés, vagyis (4.4.3) figyelembevételével

$$(4.4.5) \quad \sum_{\mu=1}^n v_\mu = 1;$$

ez azt jelenti, hogy a \mathbf{v}^T vektor elemei valószínűségeloszlást alkotnak. Ezt az eloszlást a \mathbf{P} sztochasztikus mátrixszal jellemzett Markov-lánc stacionárius eloszlásának nevezzük. A bal oldali sajátvektor definíciója értelmében a stacionárius eloszlást az

$$(4.4.6) \quad (\mathbf{E} - \mathbf{P}^T)\mathbf{v} = \mathbf{0}$$

homogén lineáris egyenletrendszer $\sum_{\mu=1}^n v_{\mu} = 1$ feltétellel normált megoldása szolgáltatja.

A (4.4.6) összefüggésből az is kiolvasható, hogy ha a kezdeti eloszlás éppen a stacionárius eloszlás, akkor (4.4.1), ill. (4.4.2) alapján ez az eloszlás akárhány lépés után sem változik, vagyis $\mathbf{p}^{(N)} = \mathbf{p}^{(0)} = \mathbf{v}$ ($N = 1, 2, \dots$). (Ez indokolja a „stacionárius” elnevezést.)

A következő tételben azt mutatjuk meg, hogy minden nemnegatív elemű mátrix – állandó szorzótól eltekintve – egy sztochasztikus mátrixhoz hasonló.

4.4.2 tétel. *Ha \mathbf{A} nemnegatív elemű mátrix, amelynek legnagyobb abszolút értékű r pozitív sajátértékéhez tartozó pozitív elemű jobb oldali sajátvektor*

$$(4.4.7) \quad \mathbf{z} = [z_i] \quad (z_i > 0),$$

akkor az r sajátértékkel osztott \mathbf{A} mátrix a z_i elemekből alkotott \mathbf{Z} diagonál-mátrix segítségével végzett hasonlósági transzformációval egy \mathbf{P} sztochasztikus mátrixba megy át:

$$(4.4.8) \quad \frac{1}{r}\mathbf{A} = \mathbf{Z}\mathbf{P}\mathbf{Z}^{-1}, \quad \text{ahol } \mathbf{Z} = \langle z_k \rangle.$$

Bizonyítás. Legyen (4.4.7) az \mathbf{A} mátrix $r > 0$ maximális abszolút értékű pozitív sajátértékéhez tartozó pozitív elemű sajátvektor, és képezzük a

$$(4.4.9) \quad \mathbf{P} = \frac{1}{r}\mathbf{Z}^{-1}\mathbf{A}\mathbf{Z}$$

mátrixot, amely a feltételek értelmében szintén nemnegatív elemű. Vegyük figyelembe, hogy a (4.4.3) alatti \mathbf{e} vektor felhasználásával

$$(4.4.10) \quad \mathbf{Z}\mathbf{e} = \mathbf{z},$$

továbbá

$$(4.4.11) \quad \mathbf{A}\mathbf{z} = r\mathbf{z},$$

és

$$(4.4.12) \quad \mathbf{Z}^{-1}\mathbf{z} = \mathbf{e}.$$

Ezért a (4.4.9) összefüggést jobbról szorozva az \mathbf{e} vektorral, és behelyettesítve a (4.4.10), (4.4.11) és (4.4.12) összefüggéseket,

$$\mathbf{P}\mathbf{e} = \frac{1}{r}\mathbf{Z}^{-1}\mathbf{A}\mathbf{Z}\mathbf{e} = \frac{1}{r}\mathbf{Z}^{-1}\mathbf{A}\mathbf{z} = \frac{1}{r}\mathbf{Z}^{-1}r\mathbf{z} = \mathbf{e}$$

adódik, tehát P valóban sztochasztikus mátrix. ■

Mielőtt a sztochasztikus mátrixokra vonatkozó következő fontos tételt kimondanánk, előbb bebizonyítunk egy tételt, amely reducibilis mátrixok elemi osztóival kapcsolatos, és amelyet a következőkben fel fogunk használni.

4.4.3 tétel. *Tekintsük az*

$$(4.4.13) \quad \mathbf{A} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{S} & \mathbf{Q}_2 \end{bmatrix}$$

hiper-háromszögmátrixot, ahol a \mathbf{Q}_1 és \mathbf{Q}_2 blokk kvadratikus. Ha az \mathbf{A} mátrix λ_0 sajátértéke egyúttal a \mathbf{Q}_1 blokknak is sajátértéke, de nem sajátértéke a \mathbf{Q}_2 blokknak, akkor az \mathbf{A} és \mathbf{Q}_1 mátrix λ_0 sajátértékének megfelelő elemi osztók megegyeznek.

Bizonyítás. Tegyük fel, hogy a \mathbf{Q}_1 és \mathbf{Q}_2 blokknak nincs közös sajátértéke. Megmutatjuk, hogy ebben az esetben \mathbf{Q}_1 és \mathbf{Q}_2 elemi osztói együttesen az \mathbf{A} mátrix elemi osztóit szolgáltatják, azaz található olyan nonszinguláris \mathbf{T} mátrix, amellyel végzett hasonlósági transzformáció az \mathbf{A} mátrixot a \mathbf{Q}_1 és \mathbf{Q}_2 blokkokból álló hiperdiagonál alakba viszi át:

$$(4.4.14) \quad \mathbf{TAT}^{-1} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2 \end{bmatrix}.$$

Keressük a \mathbf{T} mátrixot a (4.4.13) szerint particionált alakban:

$$\mathbf{T} = \begin{bmatrix} \mathbf{E}_1 & \mathbf{0} \\ \mathbf{U} & \mathbf{E}_2 \end{bmatrix},$$

ahol \mathbf{E}_1 és \mathbf{E}_2 megfelelő rendszámú egységmátrix. Ekkor

$$(4.4.15) \quad \mathbf{TAT}^{-1} = \begin{bmatrix} \mathbf{E}_1 & \mathbf{0} \\ \mathbf{U} & \mathbf{E}_2 \end{bmatrix} \begin{bmatrix} \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{S} & \mathbf{Q}_2 \end{bmatrix} \begin{bmatrix} \mathbf{E}_1 & \mathbf{0} \\ -\mathbf{U} & \mathbf{E}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{UQ}_1 - \mathbf{Q}_2\mathbf{U} + \mathbf{S} & \mathbf{Q}_2 \end{bmatrix}.$$

Az így nyert, (4.4.15) jobb oldalán álló mátrix akkor egyezik meg a (4.4.14) alatti hiperdiagonál alakkal, ha az \mathbf{U} blokkot úgy választjuk meg, hogy az kielégítse a

$$(4.4.16) \quad \mathbf{Q}_2\mathbf{U} - \mathbf{UQ}_1 = \mathbf{S}$$

mátrixegyenletet. A mátrixegyenletek elméletéből ismeretes, hogy ha a \mathbf{Q}_1 és \mathbf{Q}_2 mátrixnak nincs közös sajátértéke, a (4.4.16) egyenlet tetszőleges \mathbf{S} mátrix esetén egyértelműen megoldható (lásd pl. [6]). Ez könnyen belátható, ha a \mathbf{Q}_1 és \mathbf{Q}_2 mátrixot Jordan-féle normálalakra transzformáljuk:

$$\mathbf{Q}_1 = \mathbf{Y} \langle \mu_l \mathbf{E}_{\beta_l} + \mathbf{J}_{\beta_l} \rangle \mathbf{Y}^{-1}$$

$$\mathbf{Q}_2 = \mathbf{X} \langle \lambda_k \mathbf{E}_{\alpha_k} + \mathbf{J}_{\alpha_k} \rangle \mathbf{X}^{-1}$$

és behelyettesítjük a (4.4.16) egyenletbe:

$$\mathbf{X} \langle \lambda_k \mathbf{E}_{\alpha_k} + \mathbf{J}_{\alpha_k} \rangle \mathbf{X}^{-1} \mathbf{U} - \mathbf{U} \mathbf{Y} \langle \mu_l \mathbf{E}_{\beta_l} + \mathbf{J}_{\beta_l} \rangle \mathbf{Y}^{-1} = \mathbf{S},$$

ahol α_k és β_l az egyes Jordan-blokkok rendszáma. Szorozzuk meg a kapott egyenletet balról az \mathbf{X}^{-1} , jobbról pedig az \mathbf{Y} mátrixszal

$$\langle \lambda_k \mathbf{E}_{\alpha_k} + \mathbf{J}_{\alpha_k} \rangle (\mathbf{X}^{-1} \mathbf{U} \mathbf{Y}) - (\mathbf{X}^{-1} \mathbf{U} \mathbf{Y}) \langle \mu_l \mathbf{E}_{\beta_l} + \mathbf{J}_{\beta_l} \rangle = \mathbf{X}^{-1} \mathbf{S} \mathbf{Y}$$

és írjuk át ezt a mátrixegyenletet egyetlen lineáris egyenletrendszerre, amelyben az $\tilde{\mathbf{U}} = \mathbf{X}^{-1} \mathbf{U} \mathbf{Y}$ oszlopaiból alkotott $[\tilde{\mathbf{u}}]$ hipervektor az ismeretlenek vektora, az $\tilde{\mathbf{S}} = \mathbf{X}^{-1} \mathbf{S} \mathbf{Y}$ oszlopaiból alkotott $[\tilde{\mathbf{s}}]$ a jobb oldali hipervektor (ugyanazt az eljárást alkalmazzuk, mint a 3.6.8 pont 38. példájában):

$$[\langle \lambda_k \mathbf{E}_{\alpha_k} + \mathbf{J}_{\alpha_k} \rangle \otimes \mathbf{E}_{\beta_l} - \mathbf{E}_{\alpha_k} \otimes \langle \mu_l \mathbf{E}_{\beta_l} + \mathbf{J}_{\beta_l} \rangle^T] [\tilde{\mathbf{u}}] = [\tilde{\mathbf{s}}],$$

azaz

$$[\{(\lambda_k - \mu_l) \mathbf{E}_{\alpha_k} + \mathbf{J}_{\alpha_k}\} \otimes \mathbf{E}_{\beta_l} - \mathbf{E}_{\alpha_k} \otimes \mathbf{J}_{\beta_l}^T] [\tilde{\mathbf{u}}] = [\tilde{\mathbf{s}}].$$

Mivel az együtthatómátrix blokkjai kommutatívák, alkalmazható a 3.3.1 tétel, és így az együtthatómátrix determinánsa egyenlő a hiperdetermináns determinánsával:

$$\left| \{(\lambda_k - \mu_l) \mathbf{E}_{\alpha_k} + \mathbf{J}_{\alpha_k}\}^{\beta_l} \right| = (\lambda_k - \mu_l)^{\alpha_k \beta_l},$$

amiből következik, hogy az adott egyenletrendszernek akkor és csak akkor van egyértelmű megoldása, ha $\lambda_k - \mu_l \neq 0$ minden k, l indexpárra.

Most tegyük fel, hogy λ_0 az \mathbf{A} és a \mathbf{Q}_1 mátrix közös sajátértéke. Helyettesítsük a (4.4.13) mátrix \mathbf{Q}_1 blokkját Jordan-féle normálalakjával:

$$\mathbf{V}^{-1} \mathbf{Q}_1 \mathbf{V} = \langle \mathbf{J}_0 \quad \mathbf{J}_1 \rangle,$$

ahol \mathbf{J}_0 a λ_0 sajátértékhez tartozó Jordan-blokkokból áll, \mathbf{J}_1 pedig a többi sajátértéket tartalmazza. A megfelelő hasonlósági transzformációt az \mathbf{A} mátrixon elvégezve, az alábbi adódik:

$$\begin{bmatrix} \mathbf{V}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{E} \end{bmatrix} \begin{bmatrix} \mathbf{A} \end{bmatrix} \begin{bmatrix} \mathbf{V} & \mathbf{0} \\ \mathbf{0} & \mathbf{E} \end{bmatrix} = \begin{bmatrix} \mathbf{J}_0 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{S}_{11} & \mathbf{S}_{12} & & \\ \mathbf{S}_{21} & \mathbf{S}_{22} & \mathbf{Q}_2 & \end{bmatrix} = \begin{bmatrix} \mathbf{J}_0 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & & & \\ \mathbf{S}_{11} & & & \\ \mathbf{S}_{21} & & \tilde{\mathbf{Q}}_2 & \end{bmatrix}.$$

Mivel a \mathbf{J}_0 és a \mathbf{Q}_2 blokknak nincs közös sajátértéke, az előzőekből következik, hogy megfelelő hasonlósági transzformációval a kapott mátrix a \mathbf{J}_0 és $\tilde{\mathbf{Q}}_2$

blokkból álló hiperdiagonál alakba vihető át. Ebből viszont már következik, hogy \mathbf{A} és \mathbf{J}_0 elemi osztói – és így \mathbf{A} és \mathbf{Q}_1 λ_0 sajátértékének megfelelő elemi osztói is – megegyeznek. ■

4.4.4 tétel. *Egy \mathbf{P} sztochasztikus mátrix 1 sajátértékéhez mindig kizárólag lineáris elemi osztók tartoznak.*

Bizonyítás. A 4.2.3 definíció alapján a \mathbf{P} mátrixot sorainak és oszlopainak ugyanolyan permutációjával a reducibilis mátrixok (4.2.17) alatti normálalakjára lehet hozni, ahol az $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_s$ blokkok irreducibilis mátrixok és az

$$\mathbf{A}_{r1}, \quad \mathbf{A}_{r2}, \quad \dots, \quad \mathbf{A}_{r,r-1} \quad (r = g + 1, g + 2, \dots, s)$$

blokkok közül legalább az egyik zérustól különböző. Mivel az $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_g$ blokkok izolált blokkok, így azok irreducibilis sztochasztikus mátrixok, ezért az 1 mindegyiküknek egyszeres sajátértéke. A többi diagonál-blokknak legalább egy sorelemösszege kisebb, mint 1; ezért a 4.1.7 tétel értelmében az $\mathbf{A}_{g+1}, \mathbf{A}_{g+2}, \dots, \mathbf{A}_s$ blokkok legnagyobb sajátértéke kisebb 1-nél. A \mathbf{P} mátrix tehát az alábbi particionált alakban állítható elő: $\mathbf{P} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{S} & \mathbf{Q}_2 \end{bmatrix}$, ahol a

$$\mathbf{Q}_1 = \langle \mathbf{A}_1 \quad \mathbf{A}_2 \dots \mathbf{A}_g \rangle$$

blokk 1 sajátértékéhez lineáris elemi osztók tartoznak, a \mathbf{Q}_2 blokknak pedig 1 nem sajátértéke. Így a 4.4.3 tétel alapján azonnal adódik, hogy a \mathbf{P} mátrix 1 sajátértékéhez kizárólag lineáris elemi osztók tartoznak. ■

Ha λ_0 valamely \mathbf{P} irreducibilis sztochasztikus mátrixnak egységnyi abszolút értékű komplex sajátértéke, akkor a (4.1.65) összefüggés szerint \mathbf{P} és $\lambda_0 \mathbf{P}$ hasonlóak, és így a 4.4.4 tétel alapján a λ_0 sajátértékhez tartozó elemi osztók lineárisak. A (4.2.17) normálalak és a 4.4.3 tétel segítségével ez az állítás reducibilis sztochasztikus mátrixokra is könnyen kiterjeszthető, így fennáll a következő tétel.

4.4.5 tétel. *Bármely sztochasztikus mátrix egységnyi abszolút értékű sajátértékéhez tartozó elemi osztók mind lineárisak.*

Ugyancsak a (4.1.65) összefüggésből, valamint a 4.2.5, a 4.4.2 és a 4.4.4 tételből következik az alábbi általánosabb tétel.

4.4.6 tétel. *Ha nemnegatív elemű mátrix legnagyobb abszolút értékű r sajátértékéhez pozitív sajátvektor tartozik, akkor a mátrix valamennyi r abszolút értékű sajátértékéhez tartozó elemi osztók lineárisak.*

4.4.2 Markov-láncok ergodicitása; a sztochasztikus mátrixok osztályozása

A következőkben a homogén Markov-láncok elméletének egyik legfontosabb kérdésével, a határeloszlás létezésének kérdésével foglalkozunk.

Mint már említettük, a homogén Markov-láncot meghatározza az (egylépéses) átmenetvalószínűségekből alkotott $\mathbf{P} = [p_{\mu\nu}]$ sztochasztikus mátrix (lásd a (4.4.1) definíciót). A (4.4.2) összefüggés pedig megmutatta, hogy az N -lépéses átmenetvalószínűségek mátrixa \mathbf{P}^N . Azt a kérdést vizsgáljuk az alábbiakban, hogy ha az N lépésszámot növeljük, akkor milyen feltételek mellett létezik a \mathbf{P}^N mátrixok sorozatának határértéke.

4.4.2 definíció. A

$$(4.4.17) \quad \lim_{N \rightarrow \infty} \mathbf{P}^N = \mathbf{P}^\infty = [p_{\mu\nu}^{(\infty)}]$$

összefüggéssel definiált \mathbf{P}^∞ mátrix $p_{\mu\nu}^{(\infty)}$ elemeit – ha a (4.4.17) határérték létezik – **határvalószínűségeknek** nevezzük.

A határvalószínűség létezésének kérdésével kapcsolatban szükséges néhány új fogalom bevezetése.

4.4.3 definíció. Egy sztochasztikus mátrixot **gyengén regulárisnak** nevezzünk, ha az 1 sajátértéken kívül nincs más egységnyi abszolút értékű sajátértéke.

4.4.4 definíció. Egy sztochasztikus mátrixot **regulárisnak** nevezzünk, ha 1 az egyetlen egységnyi abszolút értékű sajátértéke, és ez egyszeres sajátérték.

E definíciókból következik, hogy gyengén reguláris sztochasztikus mátrix (4.2.17) alatti normálalakjában az $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_g$ blokkok primitív mátrixok, reguláris sztochasztikus mátrix esetén pedig ezen felül még $g = 1$, tehát a normálalak egyetlen izolált blokkja primitív.

Adott sztochasztikus mátrix által meghatározott Markov-láncot – a mátrixnak megfelelően – irreducibilisnak, reducibilisnek, regulárisnak, ill. gyengén regulárisnak nevezhetünk. Ha a mátrix imprimitív, akkor a Markov-láncot *ciklikusnak*, ha a mátrix primitív, akkor a Markov-láncot *aciklikusnak* nevezzük.

Tekintsünk most egy Markov-láncot és tegyük fel, hogy létezik a (4.4.17) összefüggéssel definiált határvalószínűségek \mathbf{P}^∞ mátrixa. Különös figyelmet érdemel az az eset, amikor \mathbf{P}^∞ valamennyi sora megegyezik. Ez azt jelenti, hogy \mathbf{P}^∞ egyetlen diádként írható fel: $\mathbf{P}^\infty = \mathbf{e}\mathbf{w}^\top$, ahol \mathbf{e} a (4.4.3) alatti, csupa 1 elemet tartalmazó oszlopvektor, \mathbf{w}^\top pedig \mathbf{P}^∞ bármely sorának elemeiből alkotott sorvektor. Ha most tetszőleges $\mathbf{p}^{(0)}$ kezdeti eloszlásra felírjuk a (4.4.2) összefüggést: $\mathbf{p}^{(N)\top} = \mathbf{p}^{(0)\top} \mathbf{P}^N$, és elvégezzük az $N \rightarrow \infty$ határát-

menetet, akkor bevezetve a $\lim_{N \rightarrow \infty} \mathbf{p}^{(N)} = \mathbf{p}^{(\infty)} = [p_k^{(\infty)}]$ összefüggéssel definiált $p_k^{(\infty)}$ abszolút határvalószínűségeket, a következő összefüggést kapjuk: $\mathbf{p}^{(\infty)\top} = \mathbf{p}^{(0)\top} \mathbf{P}^\infty = \mathbf{p}^{(0)\top} \mathbf{e} \mathbf{w}^\top$. Tekintettel arra, hogy $\mathbf{p}^{(0)}$ valószínűségeloszlás, tehát $\mathbf{p}^{(0)\top} \mathbf{e} = 1$, így $\mathbf{p}^{(\infty)\top} = \mathbf{w}^\top$. Ez azt fejezi ki, hogy ha a határvalószínűségek mátrixa létezik és minden sora megegyezik, akkor az abszolút határvalószínűségek is léteznek, függetlenek a kezdeti valószínűségeloszlástól, és egyenlők a $\mathbf{P}^{(\infty)}$ mátrix (bármely) sorának az elemeivel. Ezzel kapcsolatban bevezetjük az ergodicitás fogalmát.

4.4.5 definíció. Egy homogén Markov-láncot **ergodikusnak** nevezünk, ha az abszolút határvalószínűségek léteznek, és függetlenek a kezdeti valószínűségeloszlástól. Az ergodikus Markov-láncok abszolút határvalószínűségeinek eloszlását **határeloszlásnak** nevezzük.

A következőkben Markov-láncok határvalószínűségeinek létezésére, ill. ergodicitására vonatkozó néhány tételt bizonyítunk be.

4.4.7 tétel. Homogén Markov-lánc határvalószínűségei akkor és csak akkor léteznek, ha a Markov-lánc gyengén reguláris. A határvalószínűségek mátrixa

$$(4.4.18) \quad \mathbf{P}^\infty = \frac{1}{\Delta'(1)} \mathbf{F}(1),$$

ahol

$$(4.4.19) \quad \Delta(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{\gamma_k}$$

a \mathbf{P} sztochasztikus mátrix minimálpolinomja, $\Delta'(1)$ a $\Delta(\lambda)$ polinom λ szerinti deriváltjának értéke a $\lambda = 1$ helyen és

$$(4.4.20) \quad \mathbf{F}(\lambda) = \frac{\text{adj}(\lambda \mathbf{E} - \mathbf{P})}{\theta(\lambda)}$$

a redukált adjungált $\theta(\lambda)$ az adjungált elemeinek legnagyobb közös osztója).

Bizonyítás. Először azt bizonyítjuk be, hogy a feltétel elégséges. Ugyanis, ha \mathbf{P} gyengén reguláris; akkor a 4.4.4 tétel alapján a (4.4.19) minimálpolinomban

$$(4.4.21) \quad \lambda_1 = 1, \quad \gamma_1 = 1,$$

mivel az 1 sajátértékhez tartozó elemi osztók lineárisak és így 1 a minimálpolinom egyszeres gyöke. Állítsuk elő a \mathbf{P}^N hatványt a minimálpolinom gyökhelyeihez tartozó Hermite-féle mátrixpolinomok segítségével (lásd a (4.3.6) összefüggést).

$$(4.4.22) \quad \mathbf{P}^N = \frac{1}{\Delta'(1)} \mathbf{F}(1) + N! \sum_{k=2}^s \sum_{\nu=1}^{\gamma_k-1} \frac{\lambda_k^{N-\nu}}{(N-\nu)!} H_{k\nu}(\mathbf{P}).$$

Mivel \mathbf{P} gyengén reguláris, ezért $\lambda_1 = 1$ kivételével az összes többi sajátértéke

$$|\lambda_k| < 1 \quad (k = 2, 3, \dots, s),$$

és így a (4.4.22) összefüggés jobb oldalán $N \rightarrow \infty$ esetén az első tag kivételével az összes többi zérushoz tart. Tehát gyengén reguláris sztochasztikus mátrixok esetén létezik a \mathbf{P}^∞ határérték, és (4.4.18) alakban állítható elő.

A feltétel azonban szükséges is. Tegyük fel ugyanis, hogy létezik a

$$(4.4.23) \quad \mathbf{P}^\infty = \lim_{N \rightarrow \infty} \mathbf{P}^N$$

határérték. Ekkor azonban nem lehet a \mathbf{P} mátrixnak 1-től különböző olyan sajátértéke, amely az egységkörön fekszik, ellenkező esetben ugyanis a $\lim_{N \rightarrow \infty} \lambda_k^N$ ($\lambda_k \neq 1, |\lambda_k| = 1$) határérték nem létezne, ami viszont a (4.4.23) határérték létezésének szükséges feltétele. Ezzel a tételt bebizonyítottuk. ■

Abban az esetben, amikor a \mathbf{P} mátrix reguláris – ami speciális esete a gyengén reguláris mátrixoknak –, az 1 sajátérték a $D(\lambda)$ karakterisztikus polinomnak is egyszeres gyöke, tehát a (4.4.18) összefüggés helyett

$$(4.4.24) \quad \mathbf{P}^\infty = \frac{\text{adj}(\mathbf{E} - \mathbf{P})}{D'(1)}$$

írható.

Tekintsünk most egy gyengén reguláris Markov-lánchoz tartozó sztochasztikus mátrixot a (4.2.17) szerinti normálalakban:

$$(4.4.25) \quad \left[\begin{array}{ccc|c} \mathbf{Q}_1 & & & \\ & \mathbf{Q}_2 & & \\ & & \ddots & \\ & & & \mathbf{Q}_g & \mathbf{0} \\ \hline & \mathbf{U} & & \mathbf{W} \end{array} \right].$$

Itt $\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_g$ primitív sztochasztikus mátrixok, a \mathbf{W} blokk pedig olyan alsó hiper-háromszögmátrix, melynek sajátértékei 1-nél kisebb abszolút értékűek. Képezve a (4.4.17) határértéket,

$$\mathbf{P}^\infty = \lim_{N \rightarrow \infty} \mathbf{P}^N = \left[\begin{array}{ccc|c} \mathbf{Q}_1^\infty & & & \\ & \mathbf{Q}_2^\infty & & \\ & & \ddots & \\ & & & \mathbf{Q}_g^\infty & \mathbf{0} \\ \hline & \mathbf{U}_\infty & & \mathbf{W}^\infty \end{array} \right]$$

adódik, ahol – mivel \mathbf{W} sajátértékei 1-nél kisebb abszolút értékűek –

$$\mathbf{W}^\infty = \lim_{N \rightarrow \infty} \mathbf{W}^N = \mathbf{0},$$

\mathbf{U}_∞ pedig a \mathbf{P}^N mátrixsorozat határértékeként adódó, négy blokkra partícionált \mathbf{P}^∞ hipermátrix bal alsó blokkja, amelyre $\mathbf{U}_\infty \geq \mathbf{0}$ és $\mathbf{U}_\infty \neq \mathbf{0}$.

Így a határvalószínűségek mátrixa

$$(4.4.26) \quad \mathbf{P}^\infty = \left[\begin{array}{ccc|c} \mathbf{Q}_1^\infty & & & \\ & \mathbf{Q}_2^\infty & & \\ & & \ddots & \\ & & & \mathbf{Q}_g^\infty \\ \hline & & & \mathbf{U}_\infty \end{array} \right].$$

A $\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_g$ primitív blokkokból képezett $\mathbf{Q}_1^\infty, \mathbf{Q}_2^\infty, \dots, \mathbf{Q}_g^\infty$ mátrixok meghatározásához a (4.4.24) összefüggést felhasználva, (4.1.44) és (4.1.46) alapján

$$\mathbf{Q}_1^\infty > \mathbf{0}, \dots, \mathbf{Q}_g^\infty > \mathbf{0}$$

adódik.

A redukált adjungált (4.4.20) kifejezésének segítségével felírható az alábbi azonosság:

$$(\lambda \mathbf{E} - \mathbf{P})\mathbf{F}(\lambda) \equiv \Delta(\lambda)\mathbf{E};$$

innen $\lambda = 1$ helyettesítéssel és (4.4.18) figyelembevételével

$$(4.4.27) \quad (\mathbf{E} - \mathbf{P})\mathbf{P}^\infty = \mathbf{0},$$

tehát a (4.4.26) mátrix bármely oszlopa a \mathbf{P} sztochasztikus mátrix 1 sajátértékéhez tartozó jobb oldali sajátvektor.

4.4.8 tétel. *Homogén Markov-lánc akkor és csak akkor ergodikus, ha reguláris.*

Bizonyítás. A feltétel elégséges. Egy reguláris sztochasztikus mátrixnak ugyanis 1 egyszeres sajátértéke, a hozzá tartozó egyetlen jobb oldali sajátvektor pedig a csupa 1 elemet tartalmazó (4.4.3) alatti \mathbf{e} vektor. A határvalószínűségek (4.4.24) mátrixa tehát egyetlen diád, amelynek oszlopvektora \mathbf{e} , sorvektora pedig az 1 sajátértékhez tartozó \mathbf{w}^\top bal oldali sajátvektor. Mivel így a határvalószínűségek \mathbf{P}^∞ mátrixának minden sora megegyezik, az független a kezdeti állapottól, a Markov-lánc tehát ergodikus, és a határeloszlást \mathbf{w}^\top szolgáltatja. (Megjegyezzük, hogy a határeloszlás egyúttal a stacionárius eloszlást is megadja, vö. (4.4.6)).

A feltétel szükséges is. Tegyük fel ugyanis, hogy egy gyengén reguláris Markov-láncnak van – a kezdeti valószínűségeloszlástól független – határeloszlása, azaz a határvalószínűségek \mathbf{P}^∞ mátrixának a sorai megegyeznek. Ez azonban csak akkor lehetséges, ha \mathbf{P}^∞ (4.4.26) alakú előállításában $g = 1$, és ebből következik, hogy a \mathbf{P} sztochasztikus mátrixnak egyszeres sajátértéke az 1, vagyis a Markov-lánc reguláris. ■

Végül megadjuk annak szükséges és elégséges feltételét, hogy egy Markov-lánc határvalószínűségei mind pozitívak legyenek.

4.4.9 tétel. *Homogén Markov-lánc valamennyi határvalószínűsége akkor és csak akkor pozitív, ha a Markov-lánc aciklikus.*

Bizonyítás. Az aciklikus Markov-lánc a reguláris Markov-lánc speciális esete, a hozzá tartozó \mathbf{P} sztochasztikus mátrix primitív. A 4.3.2 tétel értelmében tehát található olyan $q > 0$ természetes szám, amelyre $\mathbf{P}^q > \mathbf{0}$. Ha most a $\mathbf{P}^m = \mathbf{P}^{m-q}\mathbf{P}^q$ ($m > q$) azonosságban elvégezzük az $m \rightarrow \infty$ határátmenetet, akkor a $\mathbf{P}^\infty = \mathbf{P}^\infty\mathbf{P}^q$ összefüggésre jutunk. Mivel \mathbf{P}^∞ sztochasztikus mátrix, ezért $\mathbf{P}^\infty \geq \mathbf{0}$, és minden sorában találhatóak zérustól különböző elemek. Innen következik, hogy $\mathbf{P}^\infty\mathbf{P}^q > \mathbf{0}$, és ezért $\mathbf{P}^\infty > \mathbf{0}$.

A feltétel szükséges is: ha feltesszük, hogy $\mathbf{P}^\infty > \mathbf{0}$, akkor valamely $q > 0$ természetes számra $\mathbf{P}^q > \mathbf{0}$.

A 4.3.2 tétel alapján innen következik, hogy a \mathbf{P} mátrix primitív, és így a megfelelő homogén Markov-lánc aciklikus. ■

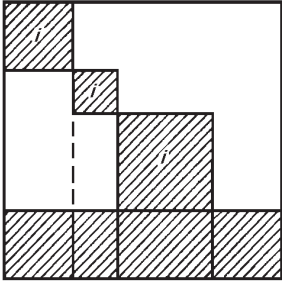
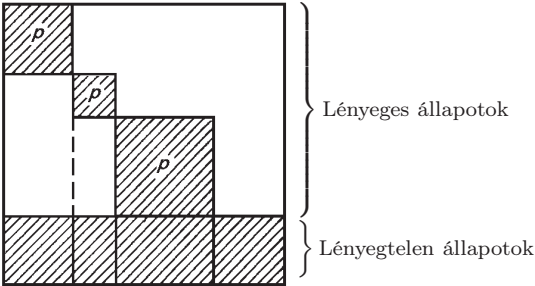
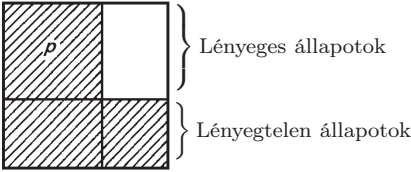
Megjegyzés. Sztochasztikus mátrix (4.4.25) normálalakja alapján az adott rendszer S_1, S_2, \dots, S_n állapotait olyan

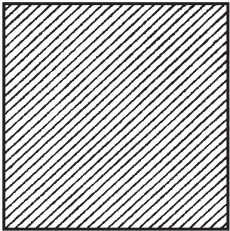
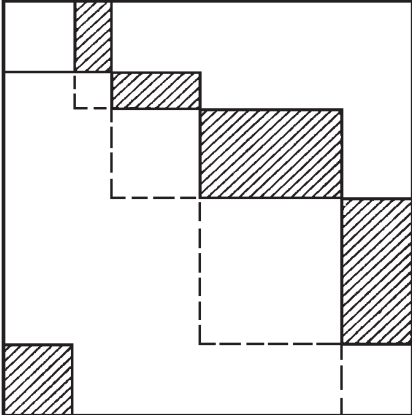
$$\Sigma_1, \Sigma_2, \dots, \Sigma_g, \Sigma_{g+1}$$

csoportokba sorolhatjuk, amelyek a (4.4.25) normálalak egyes blokksorainak felelnek meg. A rendszer azon állapotait, amelyeket a $\Sigma_1, \Sigma_2, \dots, \Sigma_g$ csoportok foglalnak magukban, *lényeges állapotoknak*, a többi pedig, amelyeket a Σ_{g+1} csoport foglal magában, *lényegtelen állapotoknak* nevezzük. Ha hatványozzuk a (4.4.25) mátrixot, akkor látható, hogy bármely véges számú lépés után a rendszer a következő átmeneteket végezheti: lényeges állapotból csak ugyanazon csoportbeli lényeges állapotba, lényegtelen állapotból vagy ugyancsak lényegtelen állapotba, vagy pedig bármely csoportbeli lényeges állapotba mehet át. Ha a lépések számát minden határon túl növeljük, akkor \mathbf{P}^∞ (4.4.26) alatti alakjából az olvasható ki, hogy határátmenetben tetszőleges állapotból csakis lényeges állapotba lehetséges az átmenet, vagyis a lényegtelen állapotba való átmenet valószínűsége zérushoz tart, ha a lépések száma a végtelenhez tart. Ezért a lényeges állapotokat *határállapotoknak* is nevezzük.

A 4.4.9 tétel szerint aciklikus Markov-lánc esetén valamennyi állapot lényeges állapot, a határeloszlás csupa pozitív elemet tartalmaz. A 4.4.8 tétel szerint reguláris Markov-láncnak – feltéve, hogy nem aciklikus – vannak lényeges és lényegtelen állapotai is, a lényeges állapotok azonban egyetlen csoportot alkotnak; a határeloszlásnak vannak pozitív elemei (ezek a lényeges állapotokhoz tartoznak) és vannak zérus elemei is (ezek a lényegtelen állapotokhoz tartoznak).

A különböző sztochasztikus mátrixok szerkezetét összefoglalóan illusztráltuk az 442–443. oldalakon.

Sztochasztikus mátrixok osztályozása I.	
Reducibilis mátrixok	Főbb tulajdonságok
<p>Általános:</p> 	<p>i-vel jelölt blokkok imprimitívek;</p> <p>Határvalószínűségek nem léteznek</p>
<p>Gyengén reguláris:</p> 	<p>p-vel jelölt blokkok primitívek;</p> <p>Határvalószínűségek léteznek</p>
<p>Reguláris:</p> 	<p>p-vel jelölt blokk primitív;</p> <p>Határvalószínűségek léteznek;</p> <p>Határeloszlás létezik és 0 elemei is vannak</p>

Sztochasztikus mátrixok osztályozása II.	
Irreducibilis mátrixok	Főbb tulajdonságok
<p>Primitív:</p>  <p>Minden állapot lényeges</p>	<p>Egyetlen primitív blokk;</p> <p>Határvalószínűségek léteznek;</p> <p>Határeloszlás létezik és minden eleme pozitív</p>
<p>Imprimitív:</p> 	<p>Határvalószínűségek nem léteznek</p>

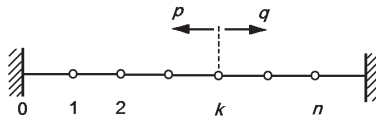
4.4.3 Bolyongási feladatok

A Markov-láncokra példaként vizsgáljunk meg néhány *diszkrét bolyongási feladatot*. Ezeket általában az jellemzi, hogy egy tömegpont (anyagi részecske) az egyenes, sík vagy tér egész koordinátájú pontjaiban – a rácspontokban – tartózkodhat, és minden lépésben valamelyik ilyen rácspontból adott átmenetvalószínűségekkel valamely ezzel szomszédos rácspontba kerül. A részecske végezheti a bolyongást az egyenes, sík vagy tér egy zárt tartományának rácspontjain, vagy az egész egyenes, ill. a teljes sík vagy tér rácspontjain. Ennek megfelelően beszélhetünk egy-, két-, háromdimenziós bolyongásról, ill.

korlátos vagy korlátlan bolyongásról. Korlátos bolyongás esetén megkülönböztethetünk „elnyelő falú” és „visszaverő falú” tartományt: ha a részecske a tartomány határára érve, ott 1 valószínűséggel helyben marad (elnyelődik), akkor a tartomány elnyelő falú, ha pedig a tartomány határával szomszédos pontból a határpontot érintve pozitív valószínűséggel tér oda vissza, akkor visszaverő falú tartományról van szó. A bolyongást homogénnek nevezzük, ha az átmenetvalószínűségek függetlenek az időtől, azaz a lépések számától. Ha ezenkívül a részecske tartózkodási helyétől sem függnek, csupán egy átmenet előtti és utáni állapot sorszámanak a különbségétől, akkor az átmenetvalószínűségek mátrixa egydimenziós bolyongás esetén – a peremfeltételek hatását kifejező sarokelemektől eltekintve – Toeplitz-típusú mátrix (lásd az 1.4.3 definíciót). Ha ezenfelül a részecske csak a szomszédos rácpontokba léphet, akkor az átmenetvalószínűségek mátrixa – a sarokelemektől eltekintve – egyenletes kontinuáns mátrix, kétdimenziós bolyongás esetén pedig hiperkontinuáns mátrix. Feltesszük, hogy a kiindulási helyzetben adott valószínűségeloszlás szerint található a részecske annak a tartománynak a rácpontjaiban, amelyben a bolyongást végzi, a bolyongás tehát egyértelműen jellemezhető az átmenetvalószínűségek mátrixával (bővebben lásd [33], [34]).

Most bemutatunk néhány egyszerűbb bolyongási feladatot.

1. Példa. *Egydimenziós bolyongás elnyelő falakkal határolt zárt intervallumban.* Tegyük fel, hogy a részecske a számegyenes $0, 1, 2, \dots, n, n+1$ koordinátájú rácpontjain végez bolyongást. Ha a részecske az intervallum bármely belső pontjában van, akkor q valószínűséggel lép az eggyel nagyobb sorszámú (koordinátájú) pontba, és p valószínűséggel az eggyel kisebb sorszámú pontba ($p + q = 1$). Ha pedig a részecske az intervallum 0 , ill. $n+1$ koordinátájú határára kerül, akkor itt „elnyelődik”, vagyis 1 valószínűséggel helyben marad (lásd az ábrát).



Írjuk fel az átmenetvalószínűségek mátrixát:

$$(4.4.28) \quad \mathbf{P} = \begin{matrix} \begin{matrix} \underbrace{0} & \underbrace{1} & \underbrace{2} & \dots & \underbrace{n} & \underbrace{n+1} \end{matrix} \\ \begin{bmatrix} 1 & 0 & & & & \\ p & 0 & q & & & \\ & p & 0 & q & & \\ & & \ddots & \ddots & \ddots & \\ & & & p & 0 & q \\ & & & & 0 & 1 \end{bmatrix} \end{matrix}.$$

A μ -edik állapot most azt jelenti, hogy a részecske a μ koordinátájú pontban van; a $p_{\mu\nu}$ átmenetvalószínűség pedig annak valószínűségét, hogy ebből a pontból egy lépésben a ν koordinátájú pontba kerül.

Az így kapott reducibilis sztochasztikus mátrix normálalakja:

$$\hat{\mathbf{P}} = \begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & p & 0 & q & & \\ & & p & 0 & q & \\ & & & \ddots & \ddots & \\ & & & & 0 & q \\ q & & & & p & 0 \end{bmatrix} = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \mathbf{q} & \mathbf{p} & \tilde{\mathbf{P}} \end{bmatrix},$$

$$\text{ahol } \mathbf{q} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ q \end{bmatrix} \text{ és } \mathbf{p} = \begin{bmatrix} p \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Látható, hogy \mathbf{P} gyengén reguláris, mivel az 1 kétszeres sajátértéke (két izolált blokkja van), de nincs több 1 abszolút értékű sajátértéke. Tehát létezik a határvalószínűségek mátrixa, de a Markov-lánc nem ergodik.

Határozzuk most meg az N -lépéses átmenetvalószínűségek mátrixát. A (4.4.28) mátrixot írjuk fel az alábbi particionált alakban:

$$(4.4.29) \quad \mathbf{P} = \begin{bmatrix} 1 & \mathbf{0} & 0 \\ \mathbf{p} & \tilde{\mathbf{P}} & \mathbf{q} \\ 0 & \mathbf{0} & 1 \end{bmatrix}.$$

Ennek N -edik hatványa, ugyancsak particionált alakban felírva:

$$(4.4.30) \quad \mathbf{P}^N = \begin{bmatrix} 1 & \mathbf{0} & 0 \\ \mathbf{x} & \tilde{\mathbf{P}}^N & \mathbf{y} \\ 0 & \mathbf{0} & 1 \end{bmatrix};$$

az \mathbf{x} és \mathbf{y} vektorok elemeit később külön kiszámítjuk. A $\tilde{\mathbf{P}}$ mátrix N -edik hatványát spektrálfelbontása segítségével határozzuk meg. Meg kell jegyezni, hogy bár $\tilde{\mathbf{P}}$ egyenletes kontinuáns mátrix, mégsem alkalmazható közvetlenül az ismert spektrálfelbontás, mivel $\tilde{\mathbf{P}}$ nem szimmetrikus. Könnyen belátható azonban, hogy a

$$(4.4.31) \quad \mathbf{D} = \left\langle 1 \quad \sqrt{\frac{q}{p}} \quad \left(\sqrt{\frac{q}{p}}\right)^2 \quad \dots \quad \left(\sqrt{\frac{q}{p}}\right)^{n-1} \right\rangle$$

diagonálmátrixszal végzett hasonlósági transzformációval szimmetrizálható. Ugyanis

$$\begin{aligned}
 (4.4.32) \quad \tilde{\mathbf{D}}\mathbf{P}\mathbf{D}^{-1} &= \\
 &= \begin{bmatrix} 1 & & & \\ \sqrt{\frac{q}{p}} & & & \\ \left(\sqrt{\frac{q}{p}}\right)^2 & & & \\ & \ddots & & \\ & & \left(\sqrt{\frac{q}{p}}\right)^{n-1} & \end{bmatrix} \begin{bmatrix} 0 & q & & \\ p & 0 & q & \\ & & \ddots & \\ & & & p & 0 \end{bmatrix} \begin{bmatrix} 1 & & & \\ \sqrt{\frac{p}{q}} & & & \\ \left(\sqrt{\frac{p}{q}}\right)^2 & & & \\ & \ddots & & \\ & & \left(\sqrt{\frac{p}{q}}\right)^{n-1} & \end{bmatrix} = \\
 &= \begin{bmatrix} 0 & \sqrt{pq} & & \\ \sqrt{pq} & 0 & \sqrt{pq} & \\ & & \ddots & \\ & & & \sqrt{pq} & 0 \end{bmatrix} = \sqrt{pq} \begin{bmatrix} 0 & 1 & & \\ 1 & 0 & 1 & \\ & & \ddots & \\ & & & 1 & 0 & 1 \\ & & & & 1 & 0 \end{bmatrix} = \sqrt{pq} \mathbf{K}.
 \end{aligned}$$

Mivel \mathbf{K} spektrálfelbontása ismert: $\mathbf{K} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$, ahol (3.2.63) és (3.2.69) szerint

$$\begin{aligned}
 (4.4.33) \quad \lambda_k &= 2 \cos \frac{k\pi}{n+1} \quad (k = 1, 2, \dots, n), \\
 u_{\mu k} &= \sqrt{\frac{2}{n+1}} \sin \frac{\mu k\pi}{n+1} \quad (\mu = 1, 2, \dots, n),
 \end{aligned}$$

így a keresett $\tilde{\mathbf{P}}^N$ hatvány az alábbi alakban írható:

$$\tilde{\mathbf{P}}^N = (\sqrt{pq})^N \mathbf{D}^{-1} \mathbf{K}^N \mathbf{D} = (\sqrt{pq})^N (\mathbf{D}^{-1} \mathbf{U}) \mathbf{\Lambda}^N (\mathbf{U}^T \mathbf{D});$$

ennek tetszőleges $\mu\nu$ indexű eleme:

$$(4.4.34) \quad \{\tilde{\mathbf{P}}^N\}_{\mu\nu} = (\sqrt{pq})^N \sum_{k=1}^n \left(\sqrt{\frac{p}{q}}\right)^{\mu-1} u_{\mu k} \lambda_k^N u_{\nu k} \left(\sqrt{\frac{q}{p}}\right)^{\nu-1},$$

azaz

$$\begin{aligned}
 (4.4.35) \quad \{\tilde{\mathbf{P}}^N\}_{\mu\nu} &= \\
 &= (\sqrt{pq})^N \left(\sqrt{\frac{p}{q}}\right)^{\mu-\nu} \frac{2^{N+1}}{n+1} \sum_{k=1}^n \cos^N \frac{k\pi}{n+1} \sin \frac{\mu k\pi}{n+1} \sin \frac{\nu k\pi}{n+1}.
 \end{aligned}$$

A (4.4.30) mátrixban szereplő \mathbf{x} és \mathbf{y} vektor elemeit a következő megfontolás segítségével határozhatjuk meg. Az \mathbf{x} vektor μ -edik eleme annak valószínűsége, hogy a részecske, amely kezdetben a μ koordinátájú pontban volt, N lépés után a 0 pontban van. Ez azonban csakis úgy következhet be, hogy a részecske μ, \dots, K, \dots, N lépés után jut a 0 koordinátájú pontba. A részecske csak akkor kerülhet éppen K lépés után a 0 pontba, ha $K - 1$ lépés után az 1 pontban volt. Ennek megfelelően

$$\begin{aligned} x_\mu &= \{\mathbf{P}^N\}_{\mu 0} = \sum_{K=\mu}^N p\{\tilde{\mathbf{P}}^{K-1}\}_{\mu 1} = \\ &= \sum_{K=\mu}^N p(\sqrt{pq})^{K-1} \left(\sqrt{\frac{p}{q}}\right)^{\mu-1} \frac{2^K}{n+1} \sum_{k=1}^n \cos^{K-1} \frac{k\pi}{n+1} \sin \frac{\mu k\pi}{n+1} \sin \frac{k\pi}{n+1}. \end{aligned}$$

Hasonló megfontolással adódnak az \mathbf{y} vektor elemei is.

Abban a speciális esetben amikor $p = q = \frac{1}{2}$, a bolyongást szimmetrikusnak nevezzük.

* * *

Megjegyzés. A (4.4.32) alatti transzformációhoz hasonlóan, bármilyen nonszimmetrikus kontinúans mátrix is egy alkalmasan választott diagonálmátrixszal végzett hasonlósági transzformációval szimmetrizálható. Tekintsük ezért az \mathbf{A} kontinúans mátrixot:

$$\mathbf{A} = \begin{bmatrix} a_1 & b_1 & & & \\ c_1 & a_2 & b_2 & & \\ & c_2 & a_3 & \ddots & \\ & \ddots & \ddots & \ddots & b_{n-1} \\ & & c_{n-1} & a_n & \end{bmatrix}.$$

A

$$(4.4.36) \quad \mathbf{D} = \left\langle 1 \quad \sqrt{\frac{b_1}{c_1}} \quad \sqrt{\frac{b_1 b_2}{c_1 c_2}} \quad \sqrt{\frac{b_1 b_2 b_3}{c_1 c_2 c_3}} \dots \sqrt{\frac{b_1 b_2 \dots b_{n-1}}{c_1 c_2 \dots c_{n-1}}} \right\rangle$$

diagonálmátrixszal elvégezve a hasonlósági transzformációt,

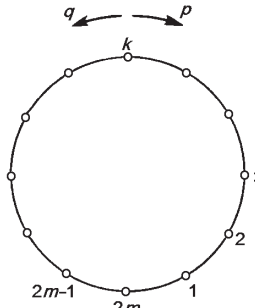
$$\mathbf{DAD}^{-1} = \begin{bmatrix} a_1 & \sqrt{b_1 c_1} & & & \\ \sqrt{b_1 c_1} & a_2 & \sqrt{b_2 c_2} & & \\ & \sqrt{b_2 c_2} & a_3 & \ddots & \\ & & & \ddots & \sqrt{b_{n-1} c_{n-1}} \\ & & & \sqrt{b_{n-1} c_{n-1}} & a_n \end{bmatrix},$$

és ezzel az \mathbf{A} mátrixot valóban szimmetrizáltuk. Látható, hogy egyenletes kontinúans mátrix esetén a (4.4.36) diagonálmátrix a (4.4.31) mátrixba megy át.

2. Példa. *Egydimenziós korlátlan bolyongás.* Tekintsük először azt a feladatot, amikor a részecske egy kör kerületének egyenlő távolságban elhelyezkedő

pontjain végzi a bolyongást. Ha a kör sugarát úgy növeljük minden határon túl, hogy a pontok száma nőjön és távolságuk ne változzék, akkor határátmennel a korlátlan egyenes rácsponjtjain végzett bolyongáshoz jutunk (lásd [42]).

Tegyük fel tehát, hogy a kör kerületén egyenlő távolságokban elhelyezünk $2m$ pontot, amelyeket mondjuk pozitív forgási irányban 1-től $2m$ -ig megszámozzunk. Legyen q annak a valószínűsége, hogy a részecske egy lépésben a pozitív irányban szomszédos, p pedig annak a valószínűsége, hogy a negatív irányban szomszédos pontba kerül. Az átmenetvalószínűségek mátrixa ekkor a következő:

$$(4.4.37) \quad \mathbf{P} = \begin{matrix} & \underbrace{1} & \underbrace{2} & & \dots & & \underbrace{2m} \\ \begin{matrix} (1) \\ \vdots \\ (2m) \end{matrix} & \begin{bmatrix} 0 & q & 0 & 0 & \dots & 0 & 0 & p \\ p & 0 & q & 0 & \dots & 0 & 0 & 0 \\ 0 & p & 0 & q & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & p & 0 & q \\ q & 0 & 0 & \dots & 0 & p & 0 & 0 \end{bmatrix} \end{matrix}.$$


Ezután a \mathbf{P} mátrixot másodrendű blokkokra particionáljuk, majd az elemeket egy olyan permutáló mátrixszal végzett ortogonális transzformációnak vetjük alá, amely a másodrendű blokkokból álló m -edrendű hipermátrixot m -edrendű blokkokból álló másodrendű hipermátrixba viszi át. Így adódik

$$\tilde{\mathbf{P}} = \begin{matrix} & \underbrace{1} & & \dots & & \underbrace{m} \\ \begin{matrix} (1) \\ \vdots \\ (m) \end{matrix} & \begin{bmatrix} & & q & & & p \\ & & p & & q & \\ & & & & p & \ddots \\ & & & & & p & q \\ p & q & & & & & \\ p & & q & & & & \\ & & & p & & & \\ q & & & & q & & \\ & & & & & p & \end{bmatrix} \end{matrix}.$$

$\underbrace{1} \quad \dots \quad \underbrace{m}$

Ebből következik, hogy $\tilde{\mathbf{P}}$ egy 2 indexű imprimitív sztochasztikus mátrix, vagyis a Markov-lánc ciklikus, ezért határeloszlás nincs, tehát nem ergodik.

Térjünk vissza a (4.4.37) mátrixhoz, és határozzuk meg az N -lépéses átmenetvalószínűségek mátrixát. A \mathbf{P}^N hatványt ismét \mathbf{P} spektrálfelbontása segítségével fejezzük ki. Látható, hogy a (4.4.37) sztochasztikus mátrix ciklikus,

tehát felírható az Ω elemi ciklikus mátrix polinomjaként: $\mathbf{P} = q\Omega + p\Omega^{n-1}$ ($n = 2m$), vagy – az $\Omega^{n-1} = \Omega^{-1}$ összefüggés figyelembevételével –

$$(4.4.38) \quad \mathbf{P} = q\Omega + p\Omega^{-1}.$$

Felhasználva az elemi ciklikus mátrix ismert spektrálfelbontását (lásd a 3.2 szakasz 5. példáját),

$$(4.4.39) \quad \Omega = \mathbf{U}\langle\omega_k\rangle\mathbf{U}^H,$$

ahol (3.2.48) és (3.2.51) szerint

$$(4.4.40) \quad \begin{aligned} \omega_k &= e^{\frac{2k\pi}{n}i} \quad (k = 0, 1, \dots, n-1), \\ u_{\mu k} &= \frac{1}{\sqrt{n}} e^{\frac{2\mu k\pi}{n}i} \quad (\mu = 0, 1, \dots, n-1), \end{aligned}$$

ezért a \mathbf{P}^N hatvány elemei

$$(4.4.41) \quad \{\mathbf{P}^N\}_{\mu\nu} = \frac{1}{n} \sum_{k=0}^{n-1} \left(qe^{\frac{2k\pi}{n}i} + pe^{-\frac{2k\pi}{n}i} \right)^N e^{(\mu-\nu)\frac{2k\pi}{n}i}.$$

Szimmetrikus bolyongás esetén, vagyis ha $p = q = \frac{1}{2}$,

$$(4.4.42) \quad \{\mathbf{P}_{\text{szim}}^N\}_{\mu\nu} = \frac{1}{n} \sum_{k=0}^{n-1} \cos^N \frac{2k\pi}{n} \cos(\mu-\nu) \frac{2k\pi}{n}.$$

A kapott eredményből leolvasható az az egyébként magától értetődő tulajdonság, hogy a μ sorszámú pontból a ν sorszámú pontba az N -lépéses átmenetvalószínűség csak a $\mu - \nu$ különbségtől függ.

Ha most a korlátlan bolyongás vizsgálatára kívánunk rátérni, akkor a (4.4.41) képletben oly módon végzünk határátmenetet, hogy k és n ugyanolyan gyorsan tartson a ∞ -hez a következő értelemben:

$$(4.4.43) \quad \frac{2k\pi}{n} \rightarrow \varphi; \quad \frac{2\pi}{n} = \Delta\varphi \rightarrow 0.$$

Elvégezve a helyettesítést, egy olyan integrálközelítő összeget kapunk, amelyből határátmenettel

$$(4.4.44) \quad \lim_{\Delta\varphi \rightarrow 0} \{\mathbf{P}^N\}_{\mu\nu} = \frac{1}{2\pi} \int_0^{2\pi} (qe^{\varphi i} + pe^{-\varphi i})^N e^{(\mu-\nu)\varphi i} d\varphi$$

adódik.

Jelölje $p_{\mu-\nu}^N$ annak valószínűségét, hogy a részecske N lépésben az egyenes μ pontjából az egyenes ν pontjába kerül, és végezzük el a (4.4.44) integrálban a $z = e^{\varphi i}$, azaz $\varphi = \frac{1}{i} \ln z$, $d\varphi = \frac{1}{i} \frac{dz}{z}$ helyettesítést. Így a következő (komplex változós) integrálhoz jutunk:

$$(4.4.45) \quad p_{\mu-\nu}^{(N)} = \frac{1}{2\pi i} \oint_{|z|=1} \left(qz + p\frac{1}{z} \right)^N z^{\mu-\nu-1} dz.$$

Alakítsuk át az integranduszt, amelyet jelöljön $g(z)$:

$$(4.4.46) \quad g(z) \equiv \left(qz + p\frac{1}{z} \right)^N z^{\mu-\nu-1} = \sum_{k=0}^N \binom{N}{k} (qz)^k \left(\frac{p}{z} \right)^{N-k} z^{\mu-\nu-1} = \\ = \sum_{k=0}^N \binom{N}{k} q^k p^{N-k} z^{2k-N+\mu-\nu-1}.$$

Mivel a (4.4.45) integrálást a komplex síkon, az egységgör mentén kell elvégezni, és a $g(z)$ integrandusznak (4.4.46) szerint az egységgör belsejében csak a $z = 0$ helyen van pólusa, a reziduum-tételt alkalmazva, $p_{\mu-\nu}^{(N)} = \operatorname{Res}_{z=0} g(z)$

adódik. Mivel pedig a $g(z)$ függvény reziduuma a $g(z) = \sum_{k=-m}^{\infty} c_k z^k$ alakú

Laurent-sor* c_{-1} együtthatója, ahol m a $z = 0$ pólus rendszáma, a (4.4.46) előállításból közvetlenül leolvasható, hogyan számítható ki c_{-1} . Ha ugyanis a $2k - N + \mu - \nu = 0$ feltételből kifejezett $k = \frac{N - \mu + \nu}{2}$ és $N - k = \frac{N + \mu - \nu}{2}$ értékeket a (4.4.46) jobb oldalán álló kifejezésbe behelyettesítjük, akkor a keresett valószínűségekre a következőket kapjuk:

$$p_{\mu-\nu}^{(N)} = c_{-1} = \left(\frac{N}{N - \mu + \nu} \right) q^{\frac{N-\mu+\nu}{2}} p^{\frac{N+\mu-\nu}{2}},$$

ha $N - \mu + \nu$ páros és $p_{\mu-\nu}^{(N)} = 0$, ha $N - \mu + \nu$ páratlan (lásd [4], [16]).

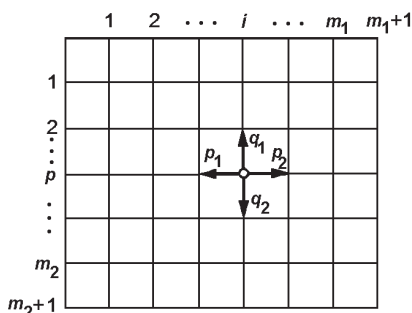
* * *

3. Példa. *Bolyongás elnyelő falakkal határolt, téglalap alakú tartomány rácspontjain.* Tekintsünk egy kétdimenziós bolyongási feladatot, amikor a

*A. Laurent (1813–1854) francia matematikus.

4.4 SZTOCHASZTIKUS MÁTRIXOK

részecske egy téglalap alakú tartomány rácspontjaiban tartózkodhat, bármely rácspontból p_1 valószínűséggel léphet balra, p_2 valószínűséggel jobbra, q_1 valószínűséggel felfelé és q_2 valószínűséggel lefelé ($p_1 + p_2 + q_1 + q_2 = 1$). A tartománynak a 0 és $m_1 + 1$, ill. 0 és $m_2 + 1$ sorszámú határvonalára jutva, a részecske ott elnyelődik, vagyis 1 valószínűséggel helyben marad. Az átmenetvalószínűségek mátrixa eszerint a következő:



$$P =$$

Átrendezve a sorokat és oszlopokat, azonnal látható, hogy a kapott mátrix gyengén reguláris sztochasztikus mátrix, mivel az 1 egy $2(m_1 + m_2 + 2)$ multiplicitású sajátérték, és nincs több 1 abszolút értékű sajátérték. Tehát a Markov-lánc nem ergodikus, bár létezik a határvalószínűségek mátrixa. Határozzuk meg az N -lépéses átmenetvalószínűségek mátrixát, azaz \mathbf{P}^N elemeit. Tekintettel arra, hogy \mathbf{P} reducibilis, normálalakja pedig $\hat{\mathbf{P}} = \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{S} & \tilde{\mathbf{P}} \end{bmatrix}$, a keresett hatványra az egydimenziós esethez hasonlóan (lásd (4.4.30)) most is $\hat{\mathbf{P}}^N = \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{U} & \tilde{\mathbf{P}}^N \end{bmatrix}$ adódik. Az \mathbf{U} mátrix elemei a (4.4.30) alatti mátrixban előforduló \mathbf{x} és \mathbf{y} vektorok elemeihez hasonlóan, egyszerű valószínűségi számításai megfontolással nyerhetők.

A $\tilde{\mathbf{P}}^N$ hatvány elemeit $\tilde{\mathbf{P}}$ spektrálfelbontásának segítségével számítjuk ki. A spektrálfelbontás könnyen meghatározható, ha észrevesszük, hogy $\tilde{\mathbf{P}}$ direkt polinomként írható fel. Bevezetve ugyanis az alábbi jelölést:

$$(4.4.47) \quad \mathbf{K}_{m_1}(p_1, p_2) = \begin{bmatrix} 0 & p_2 & & & \\ p_1 & 0 & p_2 & & \\ & p_1 & 0 & p_2 & \\ & & & \ddots & \\ & & & & p_2 \\ & & & & p_1 & 0 \end{bmatrix} \begin{matrix} (1 \\ (2 \\ \vdots \\ (m_1 \end{matrix},$$

$\tilde{\mathbf{P}}$ particionált alakban a következőképpen írható fel:

$$(4.4.48) \quad \tilde{\mathbf{P}} = \begin{bmatrix} \mathbf{K}_{m_1}(p_1, p_2) & q_2 \mathbf{E}_{m_1} & \mathbf{0} & \dots & \mathbf{0} \\ q_1 \mathbf{E}_{m_1} & \mathbf{K}_{m_1}(p_1, p_2) & q_2 \mathbf{E}_{m_1} & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & q_2 \mathbf{E}_{m_1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{K}_{m_1}(p_1, p_2) \end{bmatrix}.$$

Innen már következik a $\tilde{\mathbf{P}}$ mátrixra

$$(4.4.49) \quad \tilde{\mathbf{P}} = \mathbf{K}_{m_1}(p_1, p_2) \otimes \mathbf{E}_{m_2} + \mathbf{E}_{m_1} \otimes \mathbf{K}_{m_2}(q_1, q_2).$$

Mivel pedig $\mathbf{K}_{m_1}(p_1, p_2)$ és $\mathbf{K}_{m_2}(q_1, q_2)$ spektrálfelbontása ismert, a direkt polinomok spektrálfelbontására vonatkozó 3.3.4 tétel alkalmazásával azonnal felírható a (4.4.49) mátrix tetszőleges hatványának spektrálfelbontása. Felhasználva a (4.4.32), (4.4.33) és (4.4.34) összefüggéseket, és μ, ν indexpárral jelölve $\mathbf{K}_{m_1}(p_1, p_2)$ elemeit, ill. σ, ϱ indexpárral a $\mathbf{K}_{m_2}(q_1, q_2)$ mátrix elemeit, ezek spektrálfelbontása az alábbi alakban írható:

$$\mathbf{K}_{m_1}(p_1, p_2) = \left[\left(\sqrt{\frac{p_1}{p_2}} \right)^{\mu-1} \sqrt{\frac{2}{m_1+1}} \sin \frac{\mu k_1 \pi}{m_1+1} \right] \left\langle \sqrt{p_1 p_2} 2 \cos \frac{k_1 \pi}{m_1+1} \right\rangle \times$$

$$\times \left[\left(\sqrt{\frac{p_2}{p_1}} \right)^{\nu-1} \sqrt{\frac{2}{m_1+1}} \sin \frac{\nu k_1 \pi}{m_1+1} \right],$$

illetve

$$\begin{aligned} \mathbf{K}_{m_2}(q_1, q_2) = & \left[\left(\sqrt{\frac{q_1}{q_2}} \right)^{\sigma-1} \sqrt{\frac{2}{m_2+1}} \sin \frac{\sigma k_2 \pi}{m_2+1} \right] \left\langle \sqrt{q_1 q_2} 2 \cos \frac{k_2 \pi}{m_2+1} \right\rangle \times \\ & \times \left[\left(\sqrt{\frac{q_2}{q_1}} \right)^{\varrho-1} \sqrt{\frac{2}{m_2+1}} \sin \frac{\varrho k_2 \pi}{m_2+1} \right]. \end{aligned}$$

Ezek segítségével a (4.4.49) direkt polinom N -edik hatványának $\sigma\varrho$ indexű blokkjában a $\mu\nu$ indexű elem – a direkt polinomok spektrálfelbontásának segítségével – a következő kettős szumma alakjában adódik:

$$\begin{aligned} (4.4.50) \quad \{\mathbf{P}^N\}_{\sigma\varrho, \mu\nu} = & \frac{2^{N+2}}{(m_1+1)(m_2+1)} \left(\sqrt{\frac{p_1}{p_2}} \right)^{\mu-\nu} \left(\sqrt{\frac{q_1}{q_2}} \right)^{\sigma-\varrho} \times \\ & \times \sum_{k_2=1}^{m_2} \sum_{k_1=1}^{m_1} \left(\sqrt{p_1 p_2} \cos \frac{k_1 \pi}{m_1+1} + \sqrt{q_1 q_2} \cos \frac{k_2 \pi}{m_2+1} \right)^N \times \\ & \times \sin \frac{\mu k_1 \pi}{m_1+1} \sin \frac{\nu k_1 \pi}{m_1+1} \sin \frac{\sigma k_2 \pi}{m_2+1} \sin \frac{\varrho k_2 \pi}{m_2+1}, \end{aligned}$$

és ennek valószínűségszámítási jelentése a következő: a tartomány $\mu\sigma$ rácspontjából $\nu\varrho$ rácspontjába N lépésben a (4.4.50) szerinti valószínűséggel jut el a részecske.

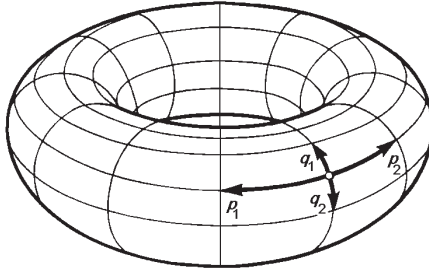
* * *

4. Példa. *Korlátlan bolyongás a sík rácspontjain.* E feladat tárgyalásakor a megfelelő egydimenziós feladat – vagyis az egyenesen folytatott korlátlan bolyongás – esetén alkalmazott gondolatmenetet általánosítjuk két dimenzióra. Először azt a feladatot vizsgáljuk meg, amikor a részecske egy tórusz felületén elhelyezkedő rácspontokon végez bolyongást. Abban az esetben, amikor a tórusz paralelköreinek és meridiánköreinek sugara egyaránt a végtelenhez tart, határártmenettel kapjuk a síkon való korlátlan bolyongásra vonatkozó eredményt.*

Jelölje p_1 annak a valószínűségét, hogy a részecske egy lépésben a paralelkörön a szomszédos rácspontra kerül pozitív irányban, p_2 pedig hogy negatív

* Ahhoz hasonlóan, ahogyan a körvonalra illeszkedő rácspontok a kör sugarának növelésével az egydimenziós, lineáris rácspontrendszerhez konvergálnak, ugyanúgy belátható, hogy egy tórusz meridiánköreire és paralelköreire illeszkedő rácspontok mindkét körrendszer sugarának növelésével a kétdimenziós, síkbeli rácspontrendszerhez konvergálnak.

irányban; jelölje továbbá q_1 annak valószínűségét, hogy a részecske a meridiánkörön kerül pozitív irányban a szomszédos rácspontra, q_2 pedig annak valószínűségét, hogy negatív irányban ($p_1 + p_2 + q_1 + q_2 = 1$) (lásd az ábrát).



Az átmenetvalószínűségek mátrixa a következő:

(4.4.51)

$$\mathbf{P} = \begin{bmatrix}
 \begin{array}{cc|cc|c|c|c|c}
 0 & p_2 & & p_1 & q_2 & & & \\
 p_1 & 0 & p_2 & & & q_2 & & \\
 & & \ddots & & p_2 & & \ddots & \\
 p_2 & & & p_1 & 0 & & & q_2 \\
 \hline
 q_1 & & & & 0 & p_2 & p_1 & q_2 \\
 & q_1 & & & p_1 & 0 & p_2 & \\
 & & \ddots & & & & p_2 & \\
 & & & q_1 & p_2 & & p_1 & 0 \\
 & & & & & & & q_2 \\
 \hline
 & & & & & & & \\
 & & & & & & & \\
 & & & & & & & \\
 & & & & & & & \\
 \hline
 q_2 & & & & & & q_1 & \\
 & q_2 & & & & & & q_1 \\
 & & \ddots & & & & & \\
 & & & q_2 & & & q_1 & p_2 \\
 & & & & & p_1 & 0 & p_1 \\
 & & & & & p_2 & & p_2 \\
 & & & & & & p_1 & 0
 \end{array}
 \end{bmatrix}.$$

Mint látható, a tórusz rácspontjain végzett bolyongásnak megfelelő Markov-láncot olyan sztochasztikus mátrix írja le, amely ciklikus blokkokból álló

ciklikus hipermátrix. Ezért a \mathbf{P} mátrix most az $\mathbf{\Omega}$ elemi ciklikus mátrix segítségével az alábbi direkt polinom alakjában írható fel:

$$(4.4.52) \quad \mathbf{P} = (p_2\mathbf{\Omega} + p_1\mathbf{\Omega}^{-1})_{m_1} \otimes \mathbf{E}_{m_2} + \mathbf{E}_{m_1} \otimes (q_2\mathbf{\Omega} + q_1\mathbf{\Omega}^{-1})_{m_2}.$$

Az $\mathbf{\Omega}$ mátrix (3.2.52) alatti spektrálfelbontásának és a direkt polinomok spektrálfelbontásának segítségével közvetlenül felírhatók a \mathbf{P}^N hatvány elemei. Ezek annak valószínűségét adják, hogy a tórusz $\mu\sigma$ indexű pontjából a részecske N lépés után a tórusz $\nu\varrho$ indexű pontjába kerül. Ezt a valószínűséget tehát a \mathbf{P}^N mátrix $\sigma\varrho$ indexű blokkjában a $\mu\nu$ indexű elem szolgáltatja. Az idézett 3.3.4 tétel értelmében így

$$(4.4.53) \quad \{\mathbf{P}^N\}_{\sigma\varrho, \mu\nu} = \frac{1}{m_1 m_2} \sum_{k_2=0}^{m_2-1} \sum_{k_1=0}^{m_1-1} \left\{ p_2 e^{\frac{2k_1\pi}{m_1}i} + p_1 e^{-\frac{2k_1\pi}{m_1}i} + q_2 e^{\frac{2k_2\pi}{m_2}i} + q_1 e^{-\frac{2k_2\pi}{m_2}i} \right\}^N \left(e^{(\mu-\nu)\frac{2k_1\pi}{m_1}i} e^{(\sigma-\varrho)\frac{2k_2\pi}{m_2}i} \right).$$

Szimmetrikus bolyongás esetén a keresett valószínűségeket megkapjuk, ha a $p_1 = p_2 = q_1 = q_2 = \frac{1}{4}$ behelyettesítése után vesszük a (4.4.53) kettős szumma valós részét, – hiszen a \mathbf{P}^N hatvány elemei biztosan valósak*:

$$\begin{aligned} \{\mathbf{P}_{\text{szimm}}^N\}_{\sigma\varrho, \mu\nu} &= \frac{1}{m_1} \frac{1}{m_2} \frac{1}{2^N} \sum_{k_2=0}^{m_2-1} \sum_{k_1=0}^{m_1-1} \left(\cos \frac{2k_1\pi}{m_1} + \cos \frac{2k_2\pi}{m_2} \right)^N \times \\ &\quad \times \cos \left[(\mu - \nu) \frac{2k_1\pi}{m_1} + (\sigma - \varrho) \frac{2k_2\pi}{m_2} \right]. \end{aligned}$$

Most térjünk rá a korlátlan bolyongás feladatára. Egyszerűség kedvéért ezt csak a szimmetrikus bolyongás esetére vizsgáljuk meg.

Tegyük fel, hogy k_1 és k_2 , ill. m_1 és m_2 úgy tart a végtelenhez, hogy az alábbi összefüggések érvényesek: $\frac{2\pi k_1}{m_1} \rightarrow \varphi_1$; $\frac{2\pi k_2}{m_2} \rightarrow \varphi_2$, továbbá $\frac{2\pi}{m_1} = \Delta\varphi_1 \rightarrow 0$; $\frac{2\pi}{m_2} = \Delta\varphi_2 \rightarrow 0$. A határátmenetet elvégezve, annak valószínűségét, hogy a részecske N lépésben a sík $\mu\sigma$ indexpárral jellemzett pontjából a sík $\nu\varrho$ indexű pontjába kerül, az alábbi kettős integrál alakjában

*Mivel a keresett valószínűség valós szám, ezért a (4.4.53) kettős szumma képzetes részét – ami biztosan zérust ad – nem írjuk ki.

kapjuk meg*:

$$\begin{aligned}
 p_{\mu-\nu, \sigma-\varrho}^{(N)} &= \lim_{\substack{\Delta\varphi_1 \rightarrow 0 \\ \Delta\varphi_2 \rightarrow 0}} \{ \mathbf{P}_{\text{szimm}}^N \}_{\sigma\varrho, \mu\nu} = \\
 &= \left(\frac{1}{2\pi} \right)^2 \frac{1}{2N} \int_0^{2\pi} \int_0^{2\pi} (\cos \varphi_1 + \cos \varphi_2)^N \cos(\mu - \nu) \varphi_1 \cos(\sigma - \varrho) \varphi_2 d\varphi_1 d\varphi_2.
 \end{aligned}$$

* * *

4.4.4 Spektrálfelbontás meghatározása generátor-függvény segítségével

Ebben a pontban egy nevezetes tridiagonális sztochasztikus mátrix spektrálfelbontását határozzuk meg. Tekintsük a

$$(4.4.54) \quad \mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ \frac{1}{n} & 0 & \frac{n-1}{n} & \dots & 0 \\ 0 & \frac{2}{n} & 0 & \frac{n-2}{n} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & 0 \\ & & & & & 0 & \frac{1}{n} \\ & & & & & 1 & 0 \end{bmatrix}$$

$n+1$ -edrendű sztochasztikus mátrixot. Már J. J. Sylvester felismerte, hogy a $\mathbf{Q} = n\mathbf{P}$ mátrix sajátértékei a 0-ra szimmetrikusan helyezkednek el és olyan számtani sorozatot alkotnak, amelynek különbsége 2. J. J. Sylvester ezt az állítást bizonyítás nélkül közölte [40]. M. Kac** *generátorfüggvény* segítségével meghatározta a (4.4.54) mátrix sajátvektorait, eljárása automatikusan szolgáltatja Sylvester állításának a bizonyítását [31]. Ezért az irodalomban a (4.4.54) mátrixot gyakran Sylvester–Kac-mátrixnak nevezik, amelynek tulajdonságaival és alkalmazásával azóta is számos szerző foglalkozott (lásd pl. [35], [43], [41], [22]). A következőkben megadjuk a generátorfüggvény definícióját, és ennek segítségével meghatározzuk a Sylvester–Kac-mátrix spektrálfelbontását.

*Itt felhasználjuk a trigonometrikus függvények addíciós tételét és az ortogonalitásból közvetlenül adódó

$$\int_0^{2\pi} \int_0^{2\pi} (\cos \varphi_1 + \cos \varphi_2)^N \sin(\mu - \nu) \varphi_1 \sin(\sigma - \varrho) \varphi_2 d\varphi_1 d\varphi_2 = 0$$

összefüggést.

**M. Kac (1914–1984) USA-beli matematikus.

4.4.6 definíció. Egy c_0, c_1, c_2, \dots valós számokból alkotott sorozat **generátorfüggvényén** az x valós változó

$$g(x) = c_0 + c_1x + c_2x^2 + \dots$$

függvényét értjük, feltéve, hogy a jobb oldalon álló sor valamilyen $-r < x < r$ intervallumban konvergens.

A spektrálfelbontást a következő gondolatmenet alapján határozzuk meg.

1. Először a \mathbf{P} mátrix bal oldali sajátvektorainak elemeit meghatározó algebrai egyenletrendszert írjuk fel. Mivel a mátrix tridiagonális, így az egyes egyenletekben három, egymást követő elem szerepel, ezért az egyenletrendszert egy másodrendű lineáris *differenciaegyenletnek* tekinthetjük, amelynek bizonyos *peremfeltételeket* kielégítő megoldását keressük. Ez azt jelenti, hogy az első és utolsó ismeretlen értéke adott.

2. Az egyenleteket rendre t hatványaival szorozva és az egyenleteket összeadva olyan generátorfüggvényt vezethetünk be, amelynek együtthatói a bal oldali sajátvektorok elemei. Így a generátorfüggvényre egy elsőrendű homogén lineáris differenciálegyenletet kapunk. A peremfeltételek figyelembevételével a generátorfüggvényekre *polinom* alakú megoldást kapunk, és innen közvetlenül következik a sajátértékekre vonatkozó Sylvester-féle állítás.

3. A sajátértékek behelyettesítésével megkapjuk a generátorfüggvényeket, amelyek – egy állandó szorzótól eltekintve – az ún. *Kravcsuk-polinomok*. Ezek együtthatóiból alkotott mátrixokat *Kravcsuk-mátrixoknak** nevezzük.

4. Az állandó szorzók megfelelő megválasztásával bebizonyítjuk, hogy a Sylvester–Kac-mátrix – egyetlen normálós tényezőtől eltekintve – a Kravcsuk-mátrixszal végzett involutórius transzformációval diagonalizálható.

A keresett spektrálfelbontáshoz tehát a következőképpen jutunk.

1. Jelölje $\mathbf{y}_i^T = [y_{i0} \ y_{i1} \ \dots \ y_{ik} \ \dots \ y_{in}]$ a \mathbf{P} mátrix λ_i sajátértékéhez tartozó bal oldali sajátvektorait ($i = 0, 1, 2, \dots, n$). Valamennyi bal oldali sajátvektorra felírva az őket definiáló egyenleteket, a következő egyenletrendszert nyerjük (n -nel szorozva az egyenleteket):

$$\begin{aligned} -n\lambda_i y_{i0} + y_{i1} &= 0 \\ ny_{i0} - n\lambda_i y_{i1} + 2y_{i2} &= 0 \\ (n-1)y_{i1} - n\lambda_i y_{i2} + 3y_{i3} &= 0 \\ \dots\dots\dots &= 0 \\ (n-k+1)y_{i,k-1} - n\lambda_i y_{ik} + (k+1)y_{i,k+1} &= 0 \\ \dots\dots\dots &= 0 \\ y_{i,n-1} - \lambda_i y_{in} &= 0 \end{aligned}$$

*M. F. Kravcsuk (Krawtchouk) (1892–1942) ukrán matematikus.

Ha az első egyenletet kiegészítjük az $(n+1)y_{i,-1}$ taggal, az utolsót pedig az $(n+1)y_{i,n+1}$ taggal, ahol $y_{i,-1} = y_{i,n+1} = 0$, akkor az

$$(4.4.56) \quad (n-k+1)y_{i,k-1} - n\lambda_i y_{ik} + (k+1)y_{i,k+1} = 0 \quad (k = 0, 1, 2, \dots, n)$$

másodrendű lineáris differenciaegyenletnek az

$$(4.4.57) \quad y_{i,-1} = y_{i,n+1} = 0 \quad (i = 0, 1, 2, \dots, n)$$

peremfeltételeket kielégítő megoldását kell meghatározni.

2. Egészítsük ki a (4.4.55) egyenletrendszert oly módon, hogy a (4.4.56) differenciaegyenletet felírjuk $k = n+1, n+2, \dots$ értékekre, így egy végtelen sok ismeretlenes egyenletrendszert kapunk. Ha ezeket az egyenleteket rendre megszorozzuk az $1, t, t^2, \dots, t^n, \dots$ hatványokkal és az egyenleteket összeadjuk, akkor bevezethetjük a

$$(4.4.58) \quad g_i(t) = y_{i0} + y_{i1}t + y_{i2}t^2 + \dots + y_{in}t^n + \dots \quad (i = 0, 1, \dots, n)$$

generátorfüggvényeket. Ha most figyelembe vesszük, hogy a (4.4.56) egyenletből $k = n+1$ esetén az $y_{i,n+1} = 0$ peremfeltétel miatt $y_{i,n+2} = 0$, és hasonlóképpen $y_{ik} = 0$ minden $k > n+1$ esetén is, akkor látható, hogy a $g_i(t)$ generátorfüggvények n -edfokú polinomok, amelyekre a

$$-n\lambda_i g_i(t) + g'_i(t) + ntg_i(t) - t^2 g'_i(t) = 0$$

elsőrendű lineáris differenciálegyenletet kapjuk. Rendezés után

$$\frac{dg_i(t)}{g_i(t)} = n \cdot \frac{\lambda_i - t}{1 - t^2} dt.$$

Ennek megoldása

$$(4.4.59) \quad g_i(t) = c_i(1+t)^{\frac{n}{2}(1+\lambda_i)}(1-t)^{\frac{n}{2}(1-\lambda_i)}.$$

A generátorfüggvények viszont csak akkor lehetnek polinomok, ha a kitevők egész számok, ebből következik, hogy

$$(4.4.60) \quad \lambda_i = 1 - \frac{2i}{n} \quad (i = 0, 1, 2, \dots, n).$$

Ezzel Sylvesternek a sajátértékekre vonatkozó állítása bizonyítást nyert.

3. Ha a sajátértékekre kapott (4.4.60) eredményt behelyettesítjük a generátorfüggvények (4.4.59) kifejezésébe, akkor ezekre

$$(4.4.61) \quad g_i(t) = c_i(1+t)^{n-i}(1-t)^i$$

adódik. Ha egyelőre eltekintünk a c_i állandóktól és az $(1+t)^{n-i}(1-t)^i$ kifejezéseket t polinomjaként írjuk fel, akkor

$$(4.4.62) \quad (1+t)^{n-i}(1-t)^i \equiv \sum_{k=0}^n a_{ik}^{(n)} t^k.$$

Ezeket a polinomokat nevezzük *Kravcsuk-polinomoknak* [32], az együtthatókból alkotott

$$(4.4.63) \quad \mathbf{A} = [a_{ik}^{(n)}]$$

mátrixokat pedig *Kravcsuk-mátrixoknak* [29]. Tehát azt az eredményt kaptuk, hogy a Sylvester–Kac-mátrix bal oldali sajátvektorai az n -edfokú Kravcsuk-polinom együtthatói, vagy – ami ezzel ekvivalens – az $n+1$ -edrendű Kravcsuk-mátrix sorai.

4.4.10 Tétel. *A Sylvester–Kac-mátrix involutórius transzformációval diagonalizálható* [35].

Ezt a következő lépésekben látjuk be:

(a) Először a nemszimmetrikus tridiagonális Sylvester–Kac-mátrixot egy diagonálmátrix segítségével végzett hasonlósági transzformációval szimmetrizáljuk.

(b) Az így nyert szimmetrikus mátrixot egy ortogonális transzformációval diagonalizáljuk.

(c) Ezután bebizonyítjuk, hogy a Kravcsuk-mátrixból hasonlósági transzformációval nyerhető a transzponáltja.

(d) Végül teljes indukcióval bebizonyítjuk, hogy a Kravcsuk-mátrix egy állandóval szorozva involutórius.

Bizonyítás. (a) A 4.4.3 pont 1. példája után tett megjegyzésből tudjuk, hogy bármely nemszimmetrikus tridiagonális mátrix egy diagonálmátrix segítségével végzett hasonlósági transzformációval szimmetrizálható. A (4.4.36) képlet alkalmazásával megmutatjuk, hogy a

$$(4.4.64) \quad \mathbf{D}^{\frac{1}{2}} = \left\langle \sqrt{\binom{n}{k}} \right\rangle$$

diagonálmátrix segítségével a $\mathbf{Q} = n\mathbf{P}$ mátrix hasonlósági transzformációval szimmetrikus alakra hozható:

$$(4.4.65) \quad \mathbf{S} = \mathbf{D}^{\frac{1}{2}} \mathbf{Q} \mathbf{D}^{-\frac{1}{2}}.$$

Valóban, mivel $q_{i,i-1} = i-1$ és $q_{i-1,i} = n-i+2$, tehát

$$s_{i,i-1} = \sqrt{\binom{n}{i-1}}(i-1) \frac{1}{\sqrt{\binom{n}{i-2}}} = \sqrt{(i-1)(n-i+2)}$$

és

$$s_{i-1,i} = \sqrt{\binom{n}{i-2}}(n-i+2) \frac{1}{\sqrt{\binom{n}{i-1}}} = \sqrt{(i-1)(n-i+2)}.$$

(b) Felhasználva a generátorfüggvényre kapott (4.4.61), a Kravcsuk-mátrixra kapott (4.4.62) és (4.4.63) összefüggéseket, valamint bevezetve a c_i állandókból képzett $\mathbf{C} = \langle c_i \rangle$ diagonálmátrixot, a \mathbf{P} mátrix bal oldali sajátvektoraiból (mint sorvektorokból) képzett mátrix

$$(4.4.66) \quad \mathbf{Y}^\top = \mathbf{C}\mathbf{A}$$

alakban írható. Felhasználva a \mathbf{P} mátrix spektrálfelbontására a $\mathbf{P} = \mathbf{X}\mathbf{\Lambda}\mathbf{Y}^\top$ összefüggést és a sajátvektorok biortogonalitását:

$$(4.4.67) \quad \mathbf{Y}^\top \mathbf{X} = \mathbf{E},$$

a szimmetrizált \mathbf{S} mátrix spektrálfelbontását (4.4.65) segítségével a következő alakban írhatjuk fel: $\mathbf{S} = (\mathbf{D}^{\frac{1}{2}} \mathbf{X} \mathbf{D}^{-\frac{1}{2}}) \mathbf{\Lambda} (\mathbf{D}^{\frac{1}{2}} \mathbf{Y}^\top \mathbf{D}^{-\frac{1}{2}})$. Mivel tudjuk, hogy \mathbf{S} ortogonális transzformációval diagonalizálható, a \mathbf{C} mátrixot abból a feltételből határozhatjuk meg, hogy \mathbf{S} jobb oldali és bal oldali sajátvektorai megegyeznek, azaz $\mathbf{D}^{\frac{1}{2}} \mathbf{X} \mathbf{D}^{-\frac{1}{2}} = (\mathbf{D}^{\frac{1}{2}} \mathbf{Y}^\top \mathbf{D}^{-\frac{1}{2}})^\top$. Innen (4.4.66) behelyettesítésével

$$(4.4.68) \quad \mathbf{X} = \mathbf{D}^{-1} \mathbf{A}^\top \mathbf{C} \mathbf{D},$$

majd (4.4.67) segítségével $\mathbf{C} \mathbf{A} \mathbf{D}^{-1} \mathbf{A}^\top \mathbf{C} \mathbf{D} = \mathbf{E}$ adódik. Mivel \mathbf{C} és \mathbf{D} kommutatívak, ezért a normáló tényezőkre azt kapjuk, hogy

$$(4.4.69) \quad \mathbf{C}^{-2} = \mathbf{A} \mathbf{D}^{-1} \mathbf{A}^\top \mathbf{D}.$$

(c) Most azt látjuk be, hogy

$$(4.4.70) \quad \mathbf{D}^{-1} \mathbf{A}^\top \mathbf{D} = \mathbf{A},$$

vagyis a Kravcsuk-mátrix hasonló a transzponáltjához. Fejezzük ki a (4.4.62) összefüggés alapján az $a_{ik}^{(n)}$ számokat a binomiális együtthatók segítségével:

$$(4.4.71) \quad a_{ik}^{(n)} = \sum_{\nu} (-1)^{\nu} \binom{i}{\nu} \binom{n-i}{k-\nu},$$

ahol ν értéke

$$k \leq n-i \text{ esetén } \nu = \begin{cases} 0, 1, \dots, k, & \text{ha } k \leq i, \\ 0, 1, \dots, i, & \text{ha } k \geq i, \end{cases}$$

$$k \geq n - i \text{ esetén } \nu = \begin{cases} k - n + i, \dots, k, & \text{ha } k \leq i, \\ k - n + i, \dots, i, & \text{ha } k \geq i. \end{cases}$$

Szorozzuk meg a (4.4.71) egyenlőség mindkét oldalát $\binom{n}{i}$ értékével:

$$\binom{n}{i} a_{ik}^{(n)} = n! \sum_{\nu} \frac{(-1)^{\nu}}{\nu! (i - \nu)! (k - \nu)! (n - i - k - \nu)!}.$$

A jobb oldalon i és k szimmetrikus kifejezése áll, tehát

$$\binom{n}{i} a_{ik}^{(n)} = a_{ki}^{(n)} \binom{n}{k},$$

ahonnan közvetlenül adódik a (4.4.70) összefüggés.

Ezzel beláttuk, hogy $\mathbf{A}^2 = \mathbf{C}^{-2}$, azaz a Kravcsuk-mátrix négyzete diagonálmátrix.

(d) Végül bebizonyítjuk, hogy

$$(4.4.72) \quad \mathbf{C} = \frac{1}{\sqrt{2^n}} \mathbf{E}.$$

Ehhez szükségünk van egy, az $a_{ik}^{(n)}$ számok között fennálló egyszerű összefüggésre, ami önmagában is érdekes, mert ennek alapján azok gyakorlatilag igen egyszerűen, rögtön táblázatosan írhatók fel. Tekintsük az

$$(1+t)\{(1+t)^{n-i}(1-t)^i\} \equiv \{(1+t)^{n-i+1}(1-t)^{i-1}\}(1-t)$$

azonosságot, ami (4.4.62) felhasználásával

$$(1+t) \sum_{k=0}^n a_{ik}^{(n)} t^k \equiv \sum_{k=0}^n a_{i-1,k}^{(n)} t^k \cdot (1-t)$$

alakban írható. Innen rögtön következik az

$$(4.4.73) \quad a_{ik}^{(n)} + a_{i,k-1}^{(n)} = a_{i-1,k}^{(n)} - a_{i-1,k-1}^{(n)}$$

összefüggés. Az $a_{ik}^{(n)}$ együtthatók (4.4.62) definíciójából nyilvánvaló, hogy $a_{ok}^{(n)} = \binom{n}{k}$ és $a_{io}^{(n)} = 1$. Ezért az $a_{ik}^{(n)}$ számok rendszerét – a Pascal-háromszöghöz hasonlóan, de nem egyetlen elemből, hanem az első sor és oszlop elemeiből kiindulva – (4.4.73) segítségével igen egyszerűen felépíthetjük.

Például $n = 4$ esetén az $a_{ik}^{(4)}$ számok \mathbf{A} táblázata:

$$\mathbf{A}_4 = \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 1 & 2 & 0 & -2 & -1 \\ 1 & 0 & -2 & 0 & 1 \\ 1 & -2 & 0 & 2 & -1 \\ 1 & -4 & 6 & -4 & 1 \end{bmatrix}.$$

Most megmutatjuk, hogy $\mathbf{A}^2 = 2^n \mathbf{E}$, vagyis

$$(4.4.74) \quad \sum_{k=0}^n a_{ik}^{(n)} a_{ki}^{(n)} = 2^n, \quad i = 0, 1, \dots, n.$$

Állításunkat teljes indukcióval bizonyítjuk: $n = 1$ esetén, amint ez a (4.4.62) összefüggésből nyilvánvaló, $\mathbf{A}_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, innen $\mathbf{A}_1^2 = 2\mathbf{E}$, azaz

$$\sum_{k=0}^1 a_{ik}^{(1)} a_{ki}^{(1)} = 2. \text{ Most – feltéve, hogy (4.4.74) teljesül – meghatározzuk}$$

$$\sum_{k=0}^{n+1} a_{ik}^{(n+1)} a_{ki}^{(n+1)}$$

értékét. Az $a_{ik}^{(n)}$ számokra érvényes (4.4.71) összefüggésből következik, hogy

$$a_{ik}^{(n+1)} = a_{ik}^{(n)} + a_{i,k-1}^{(n)}, \quad \text{ha } k = 1, 2, \dots, n,$$

$$a_{io}^{(n+1)} = a_{io}^{(n)} \quad \text{és} \quad a_{i,n+1}^{(n+1)} = a_{i,n}^{(n)},$$

valamint

$$a_{ki}^{(n+1)} = a_{ki}^{(n)} + a_{k,i-1}^{(n)}, \quad \text{ha } k = 0, 1, \dots, n,$$

illetve – felhasználva a (4.4.73) összefüggést is –

$$a_{ki}^{(n+1)} = a_{k-1,i}^{(n)} - a_{k-1,i-1}^{(n)}, \quad \text{ha } k = 1, 2, \dots, n+1.$$

Ezzel

$$\begin{aligned} \sum_{k=0}^{n+1} a_{ik}^{(n+1)} a_{ki}^{(n+1)} &= \sum_{k=0}^n a_{ik}^{(n)} (a_{ki}^{(n)} + a_{k,i-1}^{(n)}) + \sum_{k=1}^{n+1} a_{i,k-1}^{(n)} (a_{k-1,i}^{(n)} - a_{k-1,i-1}^{(n)}) = \\ &= 2 \sum_{k=0}^n a_{ik}^{(n)} a_{ki}^{(n)} = 2^{n+1}. \end{aligned}$$

(Itt kihasználtuk azt a körülményt, hogy \mathbf{A}^2 diagonálmátrix, vagyis $\sum_{k=0}^n a_{ik}^{(n)} a_{kj}^{(n)} = 0$, ha $i \neq j$.)

Ezzel bebizonyítottuk, hogy érvényes (4.4.74), tehát (4.4.72) is. Így (4.4.66), (4.4.68) és (4.4.70) figyelembevételével azt kapjuk, hogy

$$\mathbf{Y}^T = \mathbf{X} = \frac{1}{\sqrt{2^n}} \mathbf{A},$$

vagyis a Sylvester–Kac-mátrix involutórius transzformációval diagonalizálható. ■

Megjegyezzük hogy a sajátvektorokat úgy is normálhatjuk, ahogy az a Markov-láncok elméletében szokásos. A Kravcsuk-mátrix szerkezetéből következik, hogy ha a Sylvester–Kac-mátrix spektrálfelbontását

$$\mathbf{P} = \mathbf{A} \mathbf{\Lambda} \left[\frac{1}{2^n} \mathbf{A} \right]$$

alakban írjuk fel, akkor a $\lambda_0 = 1$ sajátértékhez tartozó jobb oldali sajátvektor minden eleme 1, a bal oldali sajátvektor elemei pedig az $\frac{1}{2^n} \binom{n}{k}$ binomiális eloszlást alkotják, ami egyszersmind a Markov-lánc stacionárius eloszlása. A sajátértékek (4.4.60) alakjából látható, hogy a Sylvester–Kac-mátrixnak két, maximális abszolút értékű sajátértéke $+1$ és -1 , azaz a \mathbf{P} mátrix egy 2 indexű imprimitív mátrix, amelyet sorok és oszlopok egyidejű átrendezésével másodrendű hiperciklikus alakra transzformálhatunk (lásd a 4.1.5 tételt). Ebből egyúttal az is következik, hogy a Markov-lánc nem ergodikus, azaz nincs határeloszlása. Az átrendezés egyszerűen elvégezhető, ha először pl. a páratlan indexű oszlopokat (és sorokat), utána pedig a páros indexűeket tekintjük. Így egy másodrendű blokkokból álló hipermátrixból egy négy blokkból álló hipermátrixot kapunk, amelynek jobb felső blokkja alsó bidiagonális (sztochasztikus) mátrix, bal alsó blokkja pedig felső bidiagonális (sztochasztikus) mátrix, a főátlójában pedig 0 blokkok állnak. Ezt a következőképpen szemléltethetjük:

$$\mathbf{P} \sim \begin{array}{|c|c|} \hline \begin{array}{c} 0 \\ \hline \end{array} & \begin{array}{c} \diagup \\ \hline \end{array} \\ \hline \begin{array}{c} \diagdown \\ \hline \end{array} & \begin{array}{c} 0 \\ \hline \end{array} \\ \hline \end{array} = \begin{bmatrix} \mathbf{0} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{0} \end{bmatrix}.$$

Megjegyezzük még, hogy $n + 1 = 2m$ esetén \mathbf{P}_{12} és \mathbf{P}_{21} m -edrendű kvadratikussá mátrix, $n + 1 = 2m + 1$ esetén \mathbf{P}_{12} egy $(m + 1) \times m$ -es, \mathbf{P}_{21} pedig $m \times (m + 1)$ -es téglalap alakú blokk.

Ez az átrendezés megkönnyíti a szinguláris értékek számítását, ami a

$$\begin{bmatrix} \mathbf{0} & \mathbf{P}_{21}^T \\ \mathbf{P}_{12}^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{P}_{21}^T \mathbf{P}_{21} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{12}^T \mathbf{P}_{12} \end{bmatrix}$$

hiperdiagonál-mátrix sajátérték-feladatára vezet. Ezzel a \mathbf{P} sztochasztikus mátrix szinguláris értékeinek a meghatározása a

$$\mathbf{P}_{21}^T \mathbf{P}_{21} \text{ és } \mathbf{P}_{12}^T \mathbf{P}_{12}$$

tridiagonális mátrixok sajátérték-feladatára vezethető vissza (lásd [22]).

IRODALOMJEGYZÉK

Könyvek

- [1] *Beke, M.*: Determinánsok. Athenaeum, Budapest, 1915.
- [2] *Bellman, R.*: Introduction to Matrix Analysis. McGraw-Hill, New York, 1960.
- [3] *Boullion, T. L. – Odell, P. L.*: Generalized inverse of matrices. Wiley, New York, 1971.
- [4] *Feller, W.*: An Introduction to Probability Theory and its Applications. Wiley, New York, 1950.
- [5] *Forsythe, G. E. – Moler, C. B.*: Lineáris algebrai problémák megoldása számítógéppel, Műszaki Könyvkiadó, Budapest, 1976.
- [6] *Гантмахер, Ф. П.*: Теория матриц. 2. kiadás. Nauka, Moszkva, 1966. Németül: *Gantmacher, F. R.*: Matrizenrechnung. I–II. 3. kiadás. Deutscher Verlag der Wissenschaften, Berlin, 1970, 1971.
- [7] *Гантмахер, Ф. П. – Крейн, М. Г.*: Осцилляционные матрицы и ядра и малые колебания механических систем. 2. kiadás. Gostehizdat, Moszkva, 1950. Németül: *Gantmacher, F. R. – Krein, M. G.*: Oszillationsmatrizen, Oszillationskerne und kleine Schwingungen mechanischer Systeme. Akademie Verlag, Berlin, 1960.
- [8] *Gelfand, I. M.*: Előadások a lineáris algebráról. Akadémiai Kiadó, Budapest, 1955.
- [9] *Golub, G. H. – Van Loan, A. F.*: Matrix computations. The John Hopkins Univ. Press, Baltimore, 1983.
- [10] *Householder, A. S.*: The Theory of Matrices in Numerical Analysis. Blaisdell, New York, 1964.
- [11] *Lancaster, P. – Tismenetsky, M.*: The Theory of Matrices. 2. kiadás. Academic Press, New York, 1985.
- [12] *MacDuffee, C. C.*: The Theory of Matrices. Chelsea, New York, 1956.
- [13] *Pascal, E.*: Die Determinanten. Teubner, Leipzig, 1900.
- [14] *Pringle, R. M. – Rayner, A. A.*: Generalized Inverse Matrices with Applications to Statistics. Griffen, London, 1971.
- [15] *Rao, C. R. – Mitra, S. K.*: Generalized Inverse of Matrices and its Application. Wiley, New York, 1971.
- [16] *Rényi A.*: Valószínűségszámítás. 5. kiadás. Tankönyvkiadó, Budapest, 1989.
- [17] *Schwarz, H. R.*: Numerische Mathematik. Teubner, Stuttgart, 1986.
- [18] *Szász P.*: A differenciál- és integrálszámítás elemei. II. Közoktatásügyi Kiadó, Budapest, 1951.
- [19] *Varga, R. S.*: Matrix Iterative Analysis. Prentice Hall, Englewood Cliffs, 1962.
- [20] *Zurmühl, R. – Falk, S.*: Matrizen und ihre Anwendungen. I–II. 5. kiadás. Springer, Berlin, 1984, 1986.

Dolgozatok

- [21] *Bevilacqua, R. – Codenotti, B. – Romani, F.*: Parallel solution of block tridiagonal linear systems. *Linear Algebra and its Applications* 104 (1988) 39–57.
- [22] *Boros T. – Rózsa P.*: An explicit formula for singular values of the Sylvester–Kac matrix. *Linear Algebra and its Applications* 421 (2007) 407–416.
- [23] *Egerváry, E.*: Verschärfung eines Harnackschen Satzes und anderer Abschätzungen für nichtnegative harmonische Polynome. *Mathematische Zeitschrift* 34 (1932) 741–757.
- [24] *Egerváry, E.*: On a property of the projector matrices and its application to the canonical representation of matrix functions. *Acta Scientiarum Mathematicarum* 15 (1953) 1–6.
- [25] *Egerváry J.*: Mátrixok diadikus előállításán alapuló módszer bilineáris alakok transzformációjára és lineáris egyenletrendszerek megoldására. *MTA Alkalmazott Matematikai Int. Közleményei* 2 (1953) 11–32.
- [26] *Egerváry J.*: Mátrix-függvények kanonikus előállításáról és annak néhány alkalmazásáról. *MTA III. (Matematikai és Fizikai) Osztályának Közleményei* 3 (1953) 417–458.
- [27] *Egerváry J.*: Páronként felcserélhető blokkokból álló hipermátrixokról és azok alkalmazásáról a rácsdinamikában. *MTA Alkalmazott Matematikai Int. Közleményei* 3 (1954) 31–47.
- [28] *Egerváry, J.*: Über eine konstruktive Methode zur Reduktion einer Matrix auf Jordansche Normalform. *Acta Mathematica Academiae Scientiarum Hungaricae* 10 (1959) 31–54.
- [29] *Feinsilver, Ph. – Kocik, J.*: Krawtchouk matrices from classical and quantum random walks. *Contemporary Mathematics* 287 (2001) 83–96.
- [30] *Golub, G. – Kahan, W.*: Calculating the Singular Values and Pseudo-Inverse of a Matrix. *SIAM Journal on Numerical Analysis* Ser. B 2 (1965) 205–224.
- [31] *Kac, M.*: Random walk and the theory of Brownian motion. *American Mathematical Monthly* 54 (1947) 369–390.
- [32] *Krawtchouk, M.*: Sur une généralisation des polynômes d’Hermite. *Comptes Rendus* 189 (1929) 620–622.
- [33] *McCrea, W. H. – Whipple, F. J. W.*: Random Paths in Two and Three Dimensions. *Proceedings of the Royal Society of Edinburgh* 60 (1940) 281–298.
- [34] *Reimann, J.*: Unsymmetrical Random Walk on the Plane and in the Space with Absorbing Barriers. *Acta Mathematica Academiae Scientiarum Hungaricae* 15 (1964) 339–354.
- [35] *Rózsa P.*: Megjegyzések egy sztochasztikus mátrix spektrálfelbontásához. *MTA III. (Matematikai és Fizikai) Osztályának Közleményei* 7 (1957) 199–206.
- [36] *Rózsa, P. – Bevilacqua, R. – Favati, P. – Romani, F.*: On the inverse of block tridiagonal matrices with applications to the inverses of band matrices and block band matrices. *Operator Theory: Advances and Applications* 40 (1989) 447–469. Birkhäuser, Basel.
- [37] *Rózsa, P.*: Kronecker polynomials and their applications. *Computers and Mathematics with Applications* 38 (1999) 1–10.

- [38] *Schur, L.*: Über die Abschnitte einer im Einheitskreise beschränkten Potenzreihe. *Journal für die Reine und Angewandte Mathematik* 147 (1917) 205–232.
- [39] *Sherman, J. – Morrison, W. J.*: Adjustment of an Inverse Matrix Corresponding to Changes in a Given Column or a Given Row of the Original Matrix. *The Annals of Mathematical Statistics* 21 (1949) 124.
- [40] *Sylvester, J. J.*: Théorème sur les déterminants de M. Sylvester. *Nouvelles Annales de Mathématiques* 13 (1854) 305.
- [41] *Taussky, O. – Todd, J.*: Another look at a matrix of Mark Kac. *Linear Algebra and its Applications* 150 (1991) 341–360.
- [42] *Toeplitz, O.*: Zur Theorie der quadratischen und bilinearen Formen von unendlich vielen Veränderlichen. *Mathematische Annalen* 70 (1911) 351–376.
- [43] *Vincze, I.*: Über das Ehrenfestsche Modell der Wärmeübertragung. *Archiv der Mathematik* 15 (1964) 394–400.
- [44] *Wielandt, H.*: Unzerlegbare, nicht negative Matrizen. *Mathematische Zeitschrift* 52 (1950) 642–648.
- [45] *Woodbury, M. A.*: Inverting modified matrices. Memorandum Report 42, Statistical Research Group, Institute for Advanced Study. Princeton, New Jersey, June 14, 1950.

AJÁNLOTT IRODALOM

Könyvek

- Abraham, Ph. B.*: Calculation of Functionals of Matrices Arising in Solid State Physics and Quantum Chemistry. Goddard Space Flight Center, Greenbelt, 1966.
- Amundson, N. R.*: Mathematical Methods in Chemical Engineering, Matrices and their Applications. Prentice Hall, Englewood Cliffs, 1966.
- Barnett, S. – Storey, C.*: Matrix Methods in Stability Theory. Nelson, London, 1970.
- Bellman, R. – Cooke, K. L.*: Differential-Difference Equations. Academic Press, New York, 1963.
- Berg, L.*: Lineare Gleichungssysteme mit Bandstruktur und ihr asymptotisches Verhalten. Deutscher Verlag der Wissenschaften, Berlin, 1986.
- Bunse, W. – Bunse-Gerstner, A.*: Numerische Lineare Algebra. Teubner, Stuttgart, 1985.
- Cullen, Ch.G.*: Linear Algebra and Differential Equations: An integrated approach. Prindle, Weber and Schmidt, Boston, 1979.
- Фаддеев, Д.К. – Фаддеева, В.Н.*: Вычислительные методы линейной алгебры. Fizmatgiz, Moszkva, 1960. Némethül: *Faddejew, D.K. – Faddejewa, W.N.*: Numerische Methoden der linearen Algebra. Deutscher Verlag der Wissenschaften, Berlin, 1973.
- Fagyejev, D. K. – Szominszkij, J. Sz.*: Felsőfokú algebrai feladatok. Műszaki Könyvkiadó, Budapest, 1973.
- Fiedler, M.*: Special matrices and their applications in numerical mathematics. Martinus Nijhoff, Dordrecht, 1986.
- Freud R.*: Lineáris algebra. 6. kiadás. ELTE Eötvös Kiadó, Budapest, 2007.
- Fried E.*: Klasszikus és lineáris algebra. Tankönyvkiadó, Budapest, 1977.

- Gáspár L.*: Lineáris algebra példatár. Takönyvkiadó, Budapest, 1970.
- Gohberg, I. – Lancaster, P. – Rodman, L.*: Matrix Polynomials. Academic Press, New York, 1982.
- Gohberg, I. – Lancaster, P. – Rodman, L.*: Invariant Subspaces of Matrices with Applications. SIAM, 2006.
- Gregory, R. T. – Karney, D. L.*: A Collection of Matrices for Testing Computational Algorithms. Wiley, New York, 1969.
- Halmos, P. R.*: Finite-Dimensional Vector Spaces. 2. kiadás. Van Nostrand, Princeton, 1958.
- Hamburger, H. L. – Grimshaw, M. E.*: Linear Transformations in n -Dimensional Vector Space. Cambridge Univ. Press, London–New York, 1956.
- Horn, R. A. – Johnson, Ch. R.*: Matrix Analysis. Cambridge Univ. Press, Cambridge, 1985.
- Higham, N. J.*: Functions of matrices: theory and computation. SIAM, 2006.
- Hogben, L.*: Handbook of linear algebra. Chapman & Hall/CRC, Boca Raton–London–New York, 2007.
- Израилов, Х. Д.*: Задачник по линейной алгебре. Nauka, Moszkva, 1975.
- Kirchner, I.*: Lineáris algebra és vektoralgebra. Nemzeti Tankönyvkiadó, Budapest, 2007.
- Marcus, M. – Minc, H.*: A Survey of Matrix Theory and Matrix Inequalities. Dover, New York, 1992.
- Mirsky, L.*: An introduction to linear algebra. Dover, New York, 1990.
- Noble, B.*: Applied Linear Algebra. Prentice Hall, Englewood Cliffs, 1969.
- Parlett, B. N.*: The symmetric eigenvalue problem. Prentice Hall, Englewood Cliffs, 1980.
- Parodi, M.*: La Localisation des Valeurs Caractéristiques des Matrices et ses Applications. Gauthiers-Villars, Paris, 1959.
- Prékopa A.*: Valószínűségelmélet műszaki alkalmazásokkal. 4. kiadás. Műszaki Könyvkiadó, Budapest, 1980.
- Проскуряков, И. В.*: Сборник задач по линейной алгебре. Nauka, Moszkva 1970.
- Strang, G.*: Linear algebra and its applications. 4. kiadás. Thomson Brooks, Cole, 2006.
- Usmani, R. A.*: Applied Linear Algebra. Marcel Dekker, New York, 1987.
- Varga, R. S.*: Geršgorin and his circles, Springer, Berlin, 2007.
- Воеводин, В. В.*: Линейная алгебра. Nauka, Moszkva, 1974.
- Воеводин, В. В.*: Вычислительные основы линейной алгебры. Nauka, Moszkva, 1977.
- Wedderburn, J. H. N.*: Lectures on Matrices. American Mathematical Society Colloqu. Publications Vol. XVII., New York, 1934.
- Wilkinson, J. H.*: The Algebraic Eigenvalue Problem. Clarendon Press, Oxford, 1965.
- Wilkinson, J. H. – Reinsch, C.*: Handbook for Automatic Computation. Vol. 2. Linear Algebra. Springer, Berlin, 1971.
- Young, D. M.*: Nagy lineáris rendszerek iterációs megoldása. Műszaki Könyvkiadó, Budapest, 1979.

Dolgozatok

- Bajcsay, P.*: Anwendung der Matrizenrechnung zur näherungsweisen Lösung eines Strömungsmechanischen Problems. *Periodica Polytechnica, Mechanical Engineering* 13 (1969) 139–158.
- Béres E. – Lovass-Nagy V. – Szabó J.*: Ciklikus szimmetriával bíró térbeli rácsos tartók rúderőinek meghatározása hipermátrixok alkalmazásával. *MTA Matematikai Kutató Int. Közleményei* 1 (1956) 559–575.
- Egerváry, E.*: On rank-diminishing operations and their applications to the solution of linear equations. *Zeitschrift für angewandte Mathematik und Physik* 11 (1960) 376–386.
- Elsner, L. – Redheffer, R. M.*: Remarks on Band Matrices. *Numerische Mathematik* 10 (1967) 153–161.
- Elsner, L. – Rózsa, P.*: On eigenvectors and adjoints of modified matrices. *Linear and Multilinear Algebra* 10 (1981) 235–247.
- Falk, S.*: Einschliessung für die Eigenwerte normaler Matrizenpaare. *Zeitschrift für Angewandte Mathematik und Mechanik* 44 (1964) 41–55.
- Falk, S.*: Einschliessungssätze für die Eigenvektoren normaler Matrizenpaare. *Zeitschrift für Angewandte Mathematik und Mechanik* 45 (1965) 47–56.
- Falk, S.*: Iterative Einschliessung der kleinsten (grössten) Eigenwerte eines hermiteschen Matrizenpaares. I–II. *Acta Mechanica* 46 (1983) 233–254, 49 (1983) 111–131.
- Farkas, A. – Rózsa, P. – Stubnya, E.*: Transitive matrices and their applications. *Linear Algebra and its Applications* 302–303 (1999) 423–433.
- Farkas, A. – Lancaster, P. – Rózsa, P.*: Consistency adjustments of pairwise comparison matrices. *Numerical Linear Algebra and its Applications* 10 (2003) 689–700.
- Gellai, B.*: On Hypermatrices with Blocks Commutable in Pairs in the Theory of Molecular Vibrations. *Studia Scientiarum Mathematicarum Hungarica* 6 (1971) 347–353.
- Lancaster, P. – Rózsa, P.*: The spectrum and stability of a vibrating rail supported by sleepers. *Computers and Mathematics with Applications* 31 (1996) 201–213.
- Lee, A.*: Normal Matrix Pencils. *Periodica Mathematica Hungarica* 1 (1971) 287–301.
- Lee, A.*: Hermitian and unitary matrix pencils. *Periodica Mathematica Hungarica* 5 (1974) 255–259.
- Lee, A.*: On regular polynomial matrices. I–II. *Publicationes Mathematicae* 23 (1976) 129–136, 327–333.
- Litzman, O. – Rózsa, P.*: Allgemeine Behandlung primitiver idealer und nichtidealer Kristallgitter mit Anwendung der Theorie der Hypermatrizen. *Physica Status Solidi* 2 (1962) 28–41.
- Lovass-Nagy, V. – Rózsa, P.*: Matrix Analysis of Transient Voltage Distributions in Alternating Ladder Networks. *Proceedings of the Institution of Electrical Engineering* 110 (1963) 1663–1670.
- Lynch, R. E. – Rice, J. R. – Thomas, D. H.*: Direct Solution of Partial Difference Equations by Tensor Product Methods. *Numerische Mathematik* 6 (1964) 185–195.

- Ma, Er-Chieh*: A Finite Series Solution of the Matrix Equation $AX - XB = C$. SIAM Journal on Applied Mathematics 14 (1966) 490–495.
- Nagy, T.*: Matrix Equation Analysis in the Finite Element Method. Periodica Polytechnica, Civil Engineering 14 (1970) 173–192.
- Perjés, Z.*: Anwendung der Hypermatrixen für die Untersuchung eines Widerstandnetzes. Studia Scientiarum Mathematicarum Hungarica 2 (1967) 275–283.
- Rózsa, P.*: On Periodic Continuants. Linear Algebra and its Applications 2 (1969) 267–274.
- Rózsa, P. – Sárkány, Gy. – Tettamanti, K.*: The analytical calculation of the number of theoretical plates. Periodica Polytechnica, Chemical Engineering 14 (1970) 321–331.
- Rutherford, D. E.*: Some continuant Determinants arising in Physics and Chemistry. I–II. Proceedings of the Royal Society Edinburgh 62 (1947) 229–236; 63 (1952) 232–241.
- Szabó, J.*: Ein neues Verfahren zur unmittelbaren numerischen Lösung der Dirichletschen Randwertaufgaben. Zeitschrift für Angewandte Mathematik und Mechanik 38 (1958) 280–284.
- Tassi, G. – Rózsa, P.*: Treatise of forces in cable-stayed and extradosed concrete bridges. Budapesti Műszaki Egyetem Építőmérnöki Kar Vasbetonszerkezetek Tanszéke Tudományos Közleményei (2000) 289–298.
- Wimmer, H. K. – Ziebur, A. D.*: Blockmatrizen und lineare Matrizengleichungen. Mathematische Nachrichten 59 (1974) 213–219.

NÉVMUTATÓ

- Beke Manó, 17
 Bellman, R., 310, 398
 Bevilacqua, R., 93
 Boros Tibor, 459, 464
 Boullion, T. L., 109
 Bunyakovszkij, V. J., 155–157
- Cauchy, A. L., 155–157
 Cayley, A., 272, 327
 Codenotti, B., 93
 Cramer, G., 128
 Crout, P. D., 128, 132, 137–139
 Csebisev, P. L., 71–72
- Egerváry Jenő, 39, 99, 103, 272, 299, 304, 326
 Eukleidész, 152–153
- Falk, S., 61, 130
 Favati, P., 93
 Feinsilver, Ph., 459
 Feller, W., 431, 459
 Forsythe, G., 267
 Frobenius, G., 401–402, 407, 412, 416, 418–419, 428, 430, 432
- Gantmacher, F. R., 75, 93, 145, 401–402, 434
 Gauss, K. F., 128, 132, 137–140
 Gelfand, I. M., 145, 151
 Golub, G. H., 136, 267
 Gram, J. P., 154, 155, 227
- Hamilton, W. R., 272, 327,
 Hermite, Ch., 15, 105–107, 166, 170–173, 311–326, 343, 345, 370, 372–373, 382–383, 388, 429, 438
 Hessenberg, G., 95
 Householder, A. S., 91
- Jacobi, C. G. J., 63, 90
 Jordan, C., 326–358, 377, 434–436
- Kac, M., 456, 459, 463
 Kahan, W., 136, 267
 Kocik, J., 459
 Kravcsuk (Krawtchouk), M., 459–461, 463
 Krein, M. G., 75, 93
 Kronecker, L., 15, 34, 307, 309–310, 397
- Lagrange, J. L., 276–290, 369, 375, 387, 389, 392
 Lancaster, P., 133, 402
 Laurent, P. A., 450
 Legendre, A.-M., 161
 L'Hospital, G. F. A., 74
- MacDuffee, C. C., 301
 Markov, A. A., 402, 431–433, 437–443, 445, 448, 452, 454, 463
 McCrea, W. H., 444
 Mitra, S. K., 109
 Moler, C. B., 267
 Moore, R. H., 109–112, 270
 Morrison, W. J., 86, 89
- Odell, P. L., 109
- Pascal, E., 71
 Penrose, R., 109–112, 270
 Perron, O., 402, 418, 429
 Picard, Ch. E., 364
 Pringle, R. M., 109
 Püthagorasz, 156
- Rao, C. R., 109
 Rayner, A. A., 109
 Rayleigh, J. W., 238–239
 Reimann József, 444
 Rényi Alfréd, 431, 450
 Romani, F., 93
 Rózsa Pál, 93, 307, 456, 459, 464

Schmidt, E., 158–160, 199
Schur, I., 81, 90, 92, 262
Schwarz, H. A., 155–157
Segre, B., 354, 356–357
Sherman, J., 86, 89
Sylvester, J. J., 61, 456–457, 459, 463

Szász Pál, 72

Taussky, O., 456
Taylor, B., 342, 366, 386
Tismenetsky, M., 133, 402
Todd, J., 456
Toeplitz, O., 64–65, 67, 99, 444

Van Loan, A. F., 136, 267
Varga, R. S., 402

Weyr, E., 356–357
Whipple, F. J. W., 444
Wielandt, H., 401
Woodbury, M. A., 85, 88

Zurmühl, R., 61, 130

TÁRGYMUTATÓ

abszolút határvalószínűségek, 438
 aciklikus Markov-lánc, 437
 adjungált mátrix, 34
 – transzformáció, 215
 alsó háromszögmátrix, 66
 – Jordan-blokk, 327
 általánosított inverz, mátrixé, 109
 – karakterisztikus egyenlet, 233
 – – mátrix, 233
 – – polinom, 233
 – sajátértékek, transzformációpáré, 233
 – sajátvektorok, transzformációpáré, 233
 altér, 151
 –, invariáns, 206

bázis, 149
 bázisvektor, 149
 bidiagonális mátrix, felső (alsó), 136
 bilineáris alak, 167
 biortogonális vektorrendszer, 103
 – –, teljes, 103
 blokkoszlop, 77
 blokkosor, 77

ciklikus Markov-lánc, 437
 – mátrix, 67
 – permutáló mátrix, 67

derogatórius nilpotens mátrix, 326
 determináns, 17
 determinánsosztó, 347
 diád, 28
 diád, hermitikus, 105
 diagonálblokk, izolált, 425
 diagonalizálható lineáris transzformáció,
 211
 direkt szorzat, 301, 304

egyenlet, karakterisztikus, 208
 egypárú mátrix, 75
 egyszerű struktúrájú mátrix, 211, 260
 ekvivalens transzformáció, mátrixé, 53
 elemi osztó, 348
 elsőfajú lineáris függvény, 166

ergodikus homogén Markov-lánc, 438
 értékkészlet, transzformációé, 238
 euklideszi tér, 152
 – –, izomorf, 165

felső (alsó) bidiagonális mátrix, 136
 – háromszögmátrix, 66
 – Jordan-blokk, 330
 félig involutórius mátrix, 69
 fővektor, 329

generátorfüggvény, 457
 gyengén reguláris sztochasztikus mátrix,
 437

hasonló mátrixok, 195
 hasonlósági transzformáció, 195
 határeloszlás, 438
 határvalószínűségek, 437
 –, abszolút, 438
 háromszögmátrix, alsó, 66
 –, felső, 66
 Hermite-félé (hermitikus) alak, 170
 hermitikus alak, 171
 – –, pozitív definit, 172
 – diád, 105
 hiperdetermináns, 299
 hiperdiád, 78
 hiperdiagonál-mátrix, 78
 hipermátrix inverze, 82
 –, $m \times n$ típusú, 77
 –, n -edrendű, 77
 –, perszimmetrikusan particionált, 91
 –, szimmetrikusan particionált, 77
 homogén lineáris egyenletrendszer, 114
 – – – általános megoldása, 117

idempotens mátrix, 101
 imprimitív mátrix, 428
 inhomogén lineáris egyenletrendszer, 117
 – – – megoldása, 119

invariáns altér, 206
 – faktor, 348

- invertálható mátrix, 36
- inverz mátrix, 36
 - transzformáció, 183
- involutórius mátrix, 68
- irreducibilis mátrix, 402
- izolált diagonálblokk, 425
- izomorf euklideszi tér, 165
 - lineáris tér, 150
- Jacobi-féle mátrix, 63
- Jordan-blokk, alsó, 327
 - , felső, 330
- Jordan-féle normálalak, 327
- karakterisztikus egyenlet, 208
 - mátrix, 208
 - polinom, 208
- képtér, 175
- komplex euklideszi tér, 152
 - lineáris tér, 147
- kongruens mátrixok, 192
 - transzformáció, 192
- kontinuáns (tridiagonális) mátrix, 63
- konvergens mátrixsorozat, 421
- koordináták, 150
- Kravcsuk-mátrixok, 457, 459
- Kravcsuk-polinomok, 457, 459
- Kronecker-polinom, 307
- kvadratikus alak, 170
- Legendre-polinomok, 161
- lineáris függetlenség, mátrix oszlopvektoraié, 48
 - függvény, 166
 - –, elsőfajú, 166
 - –, másodfajú, 166
 - tér, 146
 - –, n -dimenziós, 148
 - transzformáció, 175
 - – sajátértéke, 206
 - – sajátvektora, 206
 - –, diagonalizálható, 211
 - transzformációk szorzata, 180
- lineárisan összefüggő vektorok, 147
 - független vektorok, 46, 147
- magtér, 176
- Markov-lánc, 431
 - , aciklikus, 437
 - , ciklikus, 437
 - , ergodikus homogén, 438
- másodfajú lineáris függvény, 166
- mátrix, adjungált, 72
 - , α indexű nilpotens, 326
 - általánosított inverze, 109
 - , általánosított karakterisztikus, 233
 - , ciklikus, 67
 - , – permutáló, 67
 - defektusa, 60
 - , derogatórius nilpotens, 326
 - determinánsosztója, 347
 - , egypárú, 75
 - , egyszerű struktúrájú, 260
 - ekvivalens transzformációja, 53
 - elemi osztói, 348
 - , félig involutórius, 69
 - , felső (alsó) bidiagonális, 136
 - , idempotens, 101
 - , imprimitív, 428
 - invariáns faktora, 348
 - , invertálható, 36
 - inverze, 36
 - , involutórius, 68
 - , irreducibilis, 402
 - , Jacobi-féle, 63
 - , karakterisztikus, 208
 - karakterisztikus egyenlete, 208
 - – polinomja, 208
 - , kontinuáns (tridiagonális), 63
 - , $m \times n$ típusú, 19
 - minimálegyenlete, 273
 - minimális diadikus előállítása, 39
 - – felbontása, 39
 - minimálpolinomja, 273
 - Moore–Penrose-féle inverze, 109
 - , nemderogatórius nilpotens, 326
 - , nemszinguláris, 35
 - , nilpotens, 65
 - , normális, 69
 - normált általánosított inverze, 109
 - nullítása, 60
 - oszlopvektorainak lineáris függetlensége, 48
 - , permutáló, 66
 - , primitív, 428
 - pszeudoinverze, 109
 - rangja, 38, 44
 - , reducibilis, 401
 - redukált adjungáltja, 274
 - reflexív általánosított inverze, 109
 - Segre-féle karakterisztikája, 354
 - spurja (nyoma), 105
 - , Sylvester–Kac-féle, 456, 457, 459, 463
 - , szinguláris, 35
 - szinguláris értékei, 267
 - szorzása komplex számmal, 18
 - , sztochasztikus, 431, 437
 - , Toeplitz-típusú, 64
 - transzformációja, 52
 - , unitér, 68

- , valós ortogonális, 68
- Weyr-féle karakterisztikája, 356
- mátrixfüggvény, 275
- mátrixok direkt szorzata, 301
- , egyszerű struktúrájú, 211
- , hasonló, 195
- , kongruens, 192
- , Kravcsuk-féle, 457, 459
- Kronecker-polinomja, 307
- összege, 17
- mátrixsorozat határértéke, 421
- , konvergencia, 421
- minimálegyenlet, 273
- minimálpolinom, 273
- modálmátrix, 224
- Moore–Penrose-féle inverz, 109

- nemderogatórius nilpotens mátrix, 326
- nemszinguláris mátrix, 35
- transzformáció, 183
- nemvalódi ortogonális transzformáció, 244
- nilpotens mátrix, 65, 326
- – fővektorai, 329
- – indexe, 66
- normálalak, Jordan-féle, 327
- normális mátrix, 69
- transzformáció, 222
- normált általánosított inverz, mátrixé, 109
- nulltér, 176

- önadjungált transzformáció, 216
- ortogonális bázis, 158
- mátrix, 68
- transzformáció, 243, 244
- vektorok, 156
- ortonormált bázis, 158

- permutáló mátrix, 66
- perszimmetrikusan particionált hipermátrix, 91
- polinom, általánosított karakterisztikus, 233
- , karakterisztikus, 208
- polinomok, Kravcsuk-féle, 457, 459
- pozitív definit hermitikus alak, 172
- – transzformáció, 226
- szemidefinit transzformáció, 226
- primitív mátrix, 428
- projektor, 101
- projektormátrix, 101
- pszseudoinverz, 109

- Rayleigh-féle hányados, 238
- reducibilis mátrix, 401
- – normálalakja, 425
- redukált adjungált, 274
- reflexív általánosított inverz, 109
- reguláris sztochasztikus mátrix, 437

- sajátérték, lineáris transzformációé, 206
- sajátvektor, lineáris transzformációé, 206
- sajátvektor-rendszer, teljes, 211
- sávmátrix, 64
- Schur-komplementum, 81
- Segre-féle karakterisztika, 354
- Sherman–Morrison-formula, 86
- skaláris szorzat, 152
- –, standard, 154
- standard bázis, 172
- skaláris szorzat, 154
- Sylvester–Kac-mátrix, 456, 457, 459, 463
- szalagmátrix, 64
- szimmetrikus transzformáció, 242
- szimmetrikusan particionált hipermátrix, 77
- szinguláris értékek, 267
- mátrix, 35
- transzformáció, 183
- sztochasztikus mátrix, 431, 437
- –, gyengén reguláris, 437
- –, reguláris, 437

- teljes biortogonális vektorrendszer, 103
- sajátvektor-rendszer, 211
- tér lineáris transzformációja, 175
- Toeplitz-típusú mátrix, 64
- transzformáció adjungáltja, 215
- értékkészlete, 238
- modálmátrixa, 224
- , nemszinguláris, 183
- , normális, 222
- , ortogonális, 243, 244
- , önadjungált, 216
- , pozitív definit, 226
- , pozitív szemidefinit, 226
- , szimmetrikus, 242
- , szinguláris, 183
- , unitér, 196, 218
- transzformációk összege, 179
- szorzata, 180
- transzformációpár általánosított sajátértékei, 233
- – sajátvektorai, 233

- unitér, 234
 - mátrix, 68
 - tér, 153, 172
 - transzformáció, 218, 196
 - vektorrendszer, 105
 - –, nemteljes, 105
 - –, teljes, 105
- valódi ortogonális transzformáció, 244
- valós euklideszi tér, 153
 - lineáris tér, 147
 - ortogonális mátrix, 68
- vektor hossza, 154
 - koordinátái, 150
- vektorok lineáris függetlensége, 46
 - – kombinációja, 148
 - , lineárisan függetlenek, 147
 - , – összefüggőek, 147
 - , ortogonális, 156
- vektorrendszer, biortogonális, 103
- , teljes biortogonális, 103
- vektortér, 146
- Weyr-féle karakterisztika, 356