MATRIX-FÜGGVÉNYEK KANONIKUS ELŐÁLLÍTÁSÁRÓL ÉS ANNAK NÉHÁNY ALKALMAZÁSÁRÓL

EGERVÁRY JENŐ r. tag

Előadta az 1953. január 5-én tartott felolvasó ülésen

Ismeretes, hogy a matrix-algebra és analízis mindinkább jelentős szerepet nyer lineáris problémák: algebrai- és differenciálegyenletrendszerek gyakorlati megoldásában.

Ezzel szemben a matrix-elmélet tan- és kézikönyveinek túlnyomó többsége a gyakorlati alkalmazhatóság szempontjait kevéssé méltatja figyelemre. Kitűnik ez a tény abból, hogy a tárgyalás előterében mindig magának az adott matrixnak kanonikus előállítása áll és a matrix-függvénynek (mely tudvalevőleg differenciálegyenletek matrix-kalkulussal való megoldásának a kulcsa) az előállíthatósága csupán mint mellékeredmény szerepel. Továbbá, magának a matrixnak a kanonikus előállítására adott eljárás is számítástechnikailag hosszadalmas és fölösleges részleteket tartalmaz.

Így pl. egy n-edrendű valós szimmetrikus $\mathbf{A} = [a_{ij}] = [a_{ji}]$ matrix kanonikus előállítása (mechanikai fogalmazásban, normálkoordináták bevezetése) céljából a korszerű kézikönyvek[1] a következő műveletek elvégzését írják elő:

1. Meghatározandók a $|\mathbf{A}-\lambda\mathbf{E}| = \det[a_{ij}-\lambda\delta_{ij}] = 0$ karakterisztikus egyenlet $\lambda_1, \lambda_2, \ldots, \lambda_s$ gyökei.
2. Meghatározandó az $\mathbf{A}-\lambda_k\mathbf{E}$ matrix ϱ_k rangszáma.
3. Meghatározandó az $(\mathbf{A}-\lambda_k\mathbf{E})\mathbf{x} = 0$ homogén, lineáris egyenletrendszernek $n-\varrho_k$ számú lineárisan független megoldása.
4. Az így nyert megoldások rendszere ortogonalizálandó és normirozandó.

A jelen dolgozatban diagonál-alakra hozható matrixfüggvények kanonikus előállítására egy olyan direkt eljárást fogunk ismertetni, mely lényegileg a következő két lépésből áll:

I. Matrix hatványsorának minimális fokszámú polinomra való redukálása a karakterisztikus (minimál) egyenlet felhasználásával. Ennél a redukciónál a matrixfüggvény automatikusan Lagrange f. matrix-polinomok összegére redukálódik.[2]

II. A Lagrange f. matrix-polinomok felbontása a sajátvektorok diadikus szorzatainak összegére. Ez a felbontás — a Lagrange f. matrix-polinomoknak, mint projektoroknak eddig észre nem vett tulajdonsága folytán — a diádokban

[1] L. pl. Г. Е. Шилов, Введение в теорию линейных пространств. Moszkva, 1952. pp. 238—239. H. Jung, Matricen und Determinanten, Leipzig, 1951. pp. 92—99.

[2] A Lagrange f. interpolációs polinomokat a matrixelméletben először alkalmazta J. J. Sylvester, Philos. Mag. 1883. pp. 267—269.

szereplő sajátvektorokat automatikusan ortogonalizált és normált alakban szolgáltatja.

A vázolt eljárás második lépése tehát a fentebb 2. 3. és 4-gyel jelölt műveletsorozatokat egyesíti és kivitelezéséhez csupán másodrendű determinánsok szukcesszív számítása szükséges.

* * *

A dolgozat első része a matrix-aritmetika fontosabb tételeit foglalja össze. Ennek a fejezetnek a beiktatását azért véltük szükségesnek, mert egyrészt a matrix-elmélet a magyarnyelvű tan- és kézikönyvekben alig van képviselve, másrészt a matrix-elméletben döntő szerepe van egy konzekvens szimbolikának, melyet legcélszerűbben egy előzetes összefoglalás kapcsán lehet megismertetni.

Az első fejezet természetesen nem tart igényt teljességre. Benne elsősorban a matrix (és determináns-) elméletnek azokat az alaptételeit ismertettük, melyek a későbbi fejezetekben alkalmazásra kerülnek.

A második fejezet a matrix rangját evidenciába helyező diadikus felbontással foglalkozik. Ebben a fejezetben mutatkozik először a biortogonális diádokra való felbontás jelentősége a matrix sajátértékeinek és sajátvektorainak a meghatározása szempontjából. A második fejezetben bizonyítjuk továbbá, hogy ha egy projektor-matrix diádok összegére van bontva, akkor a felbontásban szereplő diádok automatikusan biortogonalizálva vannak.

A harmadik fejezet a matrix karakterisztikus és minimál-egyenletével foglalkozik. A hermitikus matrixoknak többszörös sajátérték mellett bekövetkező rangszámcsökkenését egy egyszerű algebrai identitás segélyével vezetjük le.

A negyedik fejezet élén matrix-függvénynek hatványsorral való definíciója áll. Innen (diagonál-alakra hozható matrix esetén) az előző fejezetekben levezetett tételek alkalmazása közvetlenül a Lagrange f. matrix-polinomokkal, majd a biortogonális diádokkal való kanonikus előállításra vezet. Ugyanezen fejezetben röviden érintjük az általános (diagonál-alakra nem hozható) matrixok függvényeinek Hermite f. matrix-polinomokkal való előállítását.

Az utolsó, ötödik fejezet a matrix-függvényeknek lineáris differenciálegyenletrendszerek megoldására való alkalmazását mutatja be. Az elsőrendű rendszerek megoldását ciklikus együttható-matrix esetén, egy valószínűségszámítási probléma tárgyalásával kapcsolatban ismertetjük. Másodrendű rendszerekhez vezetnek tudvalevőleg a végesszámú szabadsági fokkal bíró konzervatív mechanikai rendszerek kis rezgései. Ilyen típusú példaképpen a korpuszkuláris húrmodell sajátrezgéseinek meghatározását mutatjuk be a matrixkalkulus felhasználásával.

I. Fejezet

A matrix-aritmetika alaptételeinek áttekintése

1. §. Az alábbiakban fellépő matrixok elemei általában tetszőleges komplex számok.

Egyetlen oszlopból álló matrix neve: oszlop-vektor. Jele és részletes kiírása:

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad \text{vagy} \quad \mathbf{a}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix}. \tag{1}$$

Egyetlen sorból álló matrix neve: sorvektor. Jele és részletes kiírása:

$$\mathbf{b}^* = [b_1, b_2, \ldots, b_m] \quad \text{vagy} \quad \mathbf{b}^j = [b_{j1}, b_{j2}, \ldots, b_{jm}]. \tag{2}$$

n sorból és m oszlopból álló matrix jele általában $\mathbf{A}$; ha a sorok és oszlopok számát feltüntetni kívánjuk, akkor $\mathbf{A}_m^n$. Valamely matrix kifejezhető oszlop-, ill. sorvektoraival[3], valamint skalár-elemeivel:

$$\mathbf{A}_m^n = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m] = \begin{bmatrix} \mathbf{a}^1 \\ \mathbf{a}^2 \\ \vdots \\ \mathbf{a}^n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{bmatrix} = [a_{ij}]. \tag{3}$$

$\mathbf{A}$ minor-matrixait a következőképpen jelöljük:

$$\mathbf{A}_{\mu_1 \mu_2 \ldots \mu_r}^{\nu_1 \nu_2 \ldots \nu_s} = \begin{bmatrix} a_{\nu_1 \mu_1} & a_{\nu_1 \mu_2} & \ldots & a_{\nu_1 \mu_r} \\ a_{\nu_2 \mu_1} & a_{\nu_2 \mu_2} & \ldots & a_{\nu_2 \mu_r} \\ \ldots & \ldots & \ldots & \ldots \\ a_{\nu_s \mu_1} & a_{\nu_s \mu_2} & \ldots & a_{\nu_s \mu_r} \end{bmatrix}. \tag{4}$$

(Ha valamely matrix típusának $\mathbf{A}_m^n$ jelölése félreértést okozhat, akkor az $\mathbf{A}_{12\ldots m}^{12\ldots n}$ jelölés alkalmazandó.)

Két matrix definíció szerint akkor és csak akkor egyenlő egymással, ha azok ugyanannyi sort és oszlopot tartalmaznak, továbbá a megfelelő helyeken álló elemeik egyenlők. Egy $\mathbf{A}_m^n = \mathbf{B}_m^n$ típusú egyenlet tehát $n \cdot m$ skalár egyenlettel, valamint n, illetve m vektor-egyenlettel ekvivalens.

Valamely $\mathbf{A}$ matrix konjugáltja: $\bar{\mathbf{A}}$ az a matrix, melynek elemei $\mathbf{A}$ megfelelő elemeinek konjugáltjai.

Valamely $\mathbf{A}$ matrix transzponáltja: $\mathbf{A}^*$ az a matrix, mely $\mathbf{A}$-ból a sorok és oszlopok felcserélésével keletkezik. Ennek megfelelően a sorvektor $\mathbf{b}^*$ jele arra utal, hogy az a $\mathbf{b}$ oszlopvektor transzponáltja.

[3] Valamely matrixnak oszlop- v. sorvektoraival való kifejezése nem egyéb, mint annak speciális módon való particionálása. L. pl. A. C. Aitken, Determinants and Matrices. Edinburgh, 1948. pp. 24—25.

Az $\mathbf{A}_n^n$ típusú matrixot quadratikusnak nevezzük, n a matrix rendszáma. Diagonál-matrix az olyan quadratikus matrix, melynek a fődiagonálison kívüli elemei mind 0-sal egyenlők. Rövidebb jelölésére a

$$\begin{bmatrix} a_1 0 \ldots 0 \\ 0\ a_2 \ldots 0 \\ \ldots\ldots \\ 0\ 0 \ldots a_n \end{bmatrix} = \langle a_1, a_2, \ldots, a_n \rangle \tag{5}$$

írásmód alkalmazható.

Egység-matrix az olyan diagonálmatrix, melynek mindegyik fődiagonális eleme 1-gyel egyenlő. Jele: $\mathbf{E} = [\delta_{ij}]$.

Ha $\mathbf{A}$ minden eleme 0, akkor azt 0-matrixnak nevezzük, és 0-val jelöljük.

Ha $\mathbf{A} = \mathbf{A}^*$, azaz $a_{ij} = a_{ji}$, akkor $\mathbf{A}$ szimmetrikus.

Ha $\mathbf{A} = \bar{\mathbf{A}}^*$, azaz $a_{ij} = \bar{a}_{ji}$, akkor $\mathbf{A}$ hermitikus.

2. §. Az összeadás és kivonás csupán ugyanannyi sort és oszlopot tartalmazó matrixokra van definiálva. Ha $\mathbf{A}_m^n = [a_{ij}]$ és $\mathbf{B}_m^n = [b_{ij}]$, akkor $\mathbf{A}_m^n \pm \mathbf{B}_m^n = [a_{ij} \pm b_{ij}]$. Az így értelmezett összeadás asszociatív és kommutatív.

Matrixnak skalárral való szorzási szabálya (bizonyos folytonossági követelmények mellett) a matrix-összeadás definíciójának következménye: Ha $\mathbf{A} = [a_{ij}]$, akkor $c \cdot \mathbf{A} = \mathbf{A} \cdot c = [c \cdot a_{ij}]$.

A skalárral való szorzás kommutatív és disztributív.

A matrix-aritmetika legjellegzetesebb művelete a matrixok szorzása. Ezzel kapcsolatban indokoltnak látszik két körülményt kiemelve előrebocsátani:

Két matrixnak adott sorrendben vett szorzata csak akkor van értelmezve, ha a tényezők „konformábilisak", azaz ha az első (baloldali) tényezőnek ugyanannyi oszlopa van, mint ahány sora van a második (jobboldali) tényezőnek. $\mathbf{A}_m^n \cdot \mathbf{B}_r^s$ tehát csak akkor van értelmezve, ha $m = s$.

A matrixok szorzása általában *nem* kommutatív.

A konformábilis $\mathbf{A}_m^n = [a_{ij}]$ és $\mathbf{B}_r^m = [b_{kl}]$ matrixok szorzata definíció szerint a következő $\mathbf{C}_r^n$ típusú matrix:

$$\mathbf{C}_r^n = [c_{ij}] = \left[\sum_{\nu=1}^{m} a_{i\nu} b_{\nu j}\right] \quad \begin{matrix} (i = 1, 2, \ldots, n) \\ (j = 1, 2, \ldots, r). \end{matrix} \tag{6}$$

A matrixok szorzásának legegyszerűbb esete az $\mathbf{A}_n^1 \mathbf{B}_1^n$ ill. $\mathbf{a}^* \cdot \mathbf{b}$ típusú szorzat képzése (két vektor skaláris szorzata):

$$\mathbf{a}^* \mathbf{b} = \mathbf{b}^* \mathbf{a} = [a_1, a_2, \ldots, a_n] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \sum_{\nu=1}^{n} a_\nu b_\nu .^4 \tag{7}$$

[4] Uniter metrikájú térben két komplex vektor skalár szorzata az $a^* \bar{b}$ kifejezéssel van értelmezve.

Az $\mathbf{a}^*\mathbf{b}$ szorzatnak a fenti képlettel megadott képzési módját az $[a_1 \ldots a_n]$ sor és az $\begin{vmatrix} b_1 \\ \vdots \\ b_n \end{vmatrix}$ oszlop komponálásának nevezzük.

Eszerint általános (konformábilis) matrixok szorzásánál a szorzat i-dik sorában és j-dik oszlopában álló elemet úgy nyerjük, hogy a baloldali tényező i-dik sorát a jobboldali tényező j-dik oszlopával komponáljuk.

A mondottakból nyilvánvaló, hogy az $\mathbf{A}$ és $\mathbf{B}$ matrixok csak akkor szorozhatók úgy az $\mathbf{AB}$, mint a $\mathbf{BA}$ sorrendben, ha $\mathbf{A}_m^n$ és $\mathbf{B}_n^m$ típusúak. A szorzatok $\mathbf{C}_n^n$, ill. $\mathbf{D}_m^m$ típusú quadratikus matrixok. Ennek a megkülönböztetésnek figyelemreméltó példáját mutatják egy $\mathbf{A}_n^1 = \mathbf{a}^*$ sorvektornak és egy $\mathbf{B}_1^n = \mathbf{b}$ oszlopvektornak különböző sorrendben vett szorzatai.

Míg ugyanis az $\mathbf{A}_n^1\mathbf{B}_1^n = \mathbf{a}^*\mathbf{b}$ szorzat a $\mathbf{A}_n^1\mathbf{B}_1^n = \mathbf{C}_1^1$ összefüggésnek megfelelően a fentebb ismertetett $\sum_{\nu=1}^{n} a_\nu b_\nu$ skalárral egyenlő, addig az ellenkező sorrendben vett szorzás a[5]

$$\mathbf{B}_1^n\mathbf{A}_n^1 = \mathbf{b}\cdot\mathbf{a}^* = \begin{vmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{vmatrix} [a_1 a_2 \ldots a_n] = \begin{vmatrix} b_1a_1 & b_1a_2 \ldots b_1a_n \\ b_2a_1 & b_2a_2 \ldots b_2a_n \\ . & . \ldots . \\ b_na_1 & b_na_2 \ldots b_na_n \end{vmatrix} = \mathbf{C}_n^n \tag{8}$$

n-edrendű quadratikus matrixot eredményezi. A $\mathbf{ba}^*$-típusú matrixot az irodalomban kialakult szokásnak megfelelően két vektortényező diadikus szorzatának, vagy röviden *diád*-nak fogjuk nevezni.

3. §. Két ugyanolyan rendű quadratikus matrix szorzata mindkét sorrendben értelmezve van, azonban az így nyert szorzat-matrixok általában különbözők. Ha $\mathbf{AB} = \mathbf{BA}$, akkor az $\mathbf{A}$ és $\mathbf{B}$ matrixokat kommutabiliseknek vagy felcserélhetőknek nevezzük.

Az egységmatrix és a 0-matrix bármely vele egyező rendű matrix-szal felcserélhető és

$$\mathbf{AE} = \mathbf{EA} = \mathbf{A}; \quad \mathbf{A}\cdot 0 = 0\cdot\mathbf{A} = 0.$$

Miután $\langle c, c, \ldots, c\rangle\mathbf{A} = c\mathbf{EA} = c\mathbf{A}$, azért az n-edrendű matrixok aritmetikájában szorzás szempontjából a $\langle c, c, \ldots, c\rangle$ matrix és a c skalár ekvivalensek.

Két diagonális matrix mindig felcserélhető, mert nyilván

$$\langle a_1 a_2 \ldots a_n\rangle\langle b_1 b_2 \ldots b_n\rangle = \langle a_1b_1, a_2b_2, \ldots, a_nb_n\rangle.$$

Valamely $\mathbf{A}_n^n\mathbf{B}_n^n$ típusú matrix-szorzat két különböző alakban állítható elő aszerint, hogy a tényezőket oszlop- v. sorvektoraikból felépítettnek tekintjük.

[5] A ba^* szorzat nyilván különböző dimenziójú vektor-tényezők esetén is értelmezhető.

Az
$$\mathbf{A}_n^n=\begin{bmatrix}\mathbf{a}^1\\ \mathbf{a}^2\\ \vdots\\ \mathbf{a}^n\end{bmatrix},\ \mathbf{B}_n^n=[\mathbf{b}_1\mathbf{b}_2\ldots\mathbf{b}_n]$$
alakban írt tényezők esetén
$$\mathbf{A}_n^n\mathbf{B}_n^n=\begin{bmatrix}\mathbf{a}^1\\ \mathbf{a}^2\\ \vdots\\ \mathbf{a}^n\end{bmatrix}[\mathbf{b}_1\mathbf{b}_2\ldots\mathbf{b}_n]=\begin{bmatrix}\mathbf{a}^1\mathbf{b}_1 & \mathbf{a}^1\mathbf{b}_2\ldots\mathbf{a}^1\mathbf{b}_n\\ \mathbf{a}^2\mathbf{b}_1 & \mathbf{a}^2\mathbf{b}_2\ldots\mathbf{a}^2\mathbf{b}_n\\ . & . \ \ldots\ .\\ \mathbf{a}^n\mathbf{b}_1 & \mathbf{a}^n\mathbf{b}_2\ldots\mathbf{a}^n\mathbf{b}_n\end{bmatrix}, \tag{9}$$
tehát a szorzat egyetlen n-edrendű matrix, melynek elemei skalár szorzatok.

Az $\mathbf{A}_n^n=[\mathbf{a}_1\mathbf{a}_2\ldots\mathbf{a}_n]$, $\mathbf{B}_n^n=\begin{bmatrix}\mathbf{b}^1\\ \mathbf{b}^2\\ \vdots\\ \mathbf{b}^n\end{bmatrix}$ alakban írt tényezőkből kiindulva
$$\mathbf{A}_n^n\mathbf{B}_n^n=[\mathbf{a}_1\mathbf{a}_2\ldots\mathbf{a}_n]\begin{bmatrix}\mathbf{b}^1\\ \mathbf{b}^2\\ \vdots\\ \mathbf{b}^n\end{bmatrix}=\mathbf{a}_1\mathbf{b}^1+\mathbf{a}_2\mathbf{b}^2+\ldots+\mathbf{a}_n\mathbf{b}^n, \tag{10}$$
ekkor tehát a szorzat a tényezők közös rendszámával megegyező számú diád összegeként van előállítva.

Az alábbiakban többször fogjuk alkalmazni egy matrix-szorzat valamely minormatrixának a következő előállítását:

Ha $\mathbf{A}=\mathbf{A}_k^n$; $\mathbf{C}=\mathbf{C}_m^l$ és $\mathbf{ABC}=\mathbf{D}_m^n$, akkor
$$\mathbf{D}_{\mu_1\mu_2\ldots\mu_r}^{\nu_1\nu_2\ldots\nu_s}=\mathbf{A}_{12\ldots k}^{\nu_1\nu_2\ldots\nu_s}\mathbf{B}\,\mathbf{C}_{\mu_1\mu_2\ldots\mu_r}^{12\ldots l}\quad \begin{matrix}(s\leqq n)\\ (r\leqq m)\end{matrix}. \tag{11}$$
Ez az azonosság a szorzat definíciójának közvetlen következménye és arra redukálódik az $r=1, s=1, k=l, \mathbf{B}=\mathbf{E}$ esetben.

4. §. A skaláris szorzásnak további két alaptörvénye: az asszociativitás és a disztributivitás érvényes matrixok szorzásánál is. A szorzat definíciójának közvetlen következménye, hogy
$$(\mathbf{A}_m^n\mathbf{B}_r^m)\mathbf{C}_s^r=\mathbf{A}_m^n(\mathbf{B}_r^m\mathbf{C}_s^r)=\mathbf{D}_s^n,$$
ugyanis mindkét szorzat-matrix általános eleme: $d_{ij}=\sum_{\mu=1}^{m}\sum_{\nu=1}^{r}a_{i\mu}b_{\mu\nu}c_{\nu j}$, tehát többtényezős matrix-szorzat (a sorrend megtartása mellett) a zárójellel kijelölt csoportosítás módjától független. Eszerint a zárójelezés fölösleges és tetszőleges számú (szomszédos páronként konformábilis) matrix szorzata az első (szélsőbal) tényezővel egyező számú sort, és az utolsó (szélsőjobb) tényezővel egyező számú oszlopot tartalmaz:
$$\mathbf{A}_\beta^\alpha\mathbf{B}_\gamma^\beta\mathbf{C}_\delta^\gamma\cdots\mathbf{M}_\nu^\mu=\mathbf{N}_\nu^\alpha.$$

A tényezők közt természetesen sor- és oszlopvektorok is előfordulhatnak. Ennek legismertebb példája az $\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}x_iy_j$ bilineáris alaknak

$$\mathbf{x}^*\mathbf{A}\mathbf{y} = [x_1x_2\ldots x_n]\begin{bmatrix} a_{11}\ldots a_{1m} \\ \\ a_{n1}\ldots a_{nm}\end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m\end{bmatrix}$$

matrix-szorzat alakban való előállítása.

A matrix-szorzás disztributivitását a

$$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{A}\mathbf{B}+\mathbf{A}\mathbf{C}$$

reláció fejezi ki, mely az összeg és szorzat definíciójának közvetlen következménye. Alkalmazásánál mindig figyelembe kell lenni a tényezők sorrendjére. Pl.

$$\mathbf{A}(\mathbf{B}+\mathbf{C})\mathbf{D} = (\mathbf{A}\mathbf{B}+\mathbf{A}\mathbf{C})\mathbf{D} = \mathbf{A}\mathbf{B}\mathbf{D}+\mathbf{A}\mathbf{C}\mathbf{D}.$$

Többtényezős matrix-szorzat speciális esete valamely quadratikus matrix természetes egész kitevőjü hatványa:

$$\mathbf{A}^n = \underset{\mathrm{I}}{\mathbf{A}}\cdot\underset{\mathrm{II}}{\mathbf{A}}\cdots\underset{n}{\mathbf{A}}.$$

A szorzás asszociativitása folytán $\mathbf{A}^n$ és $\mathbf{A}^m$ (az $\mathbf{A}^0 = \mathbf{E}$ egyenlettel értelmezett 0-dik hatványt is beleértve) kommutabilisek. Innen következik, hogy egy- és ugyanazon $\mathbf{A}$ matrixból skalár együtthatókkal szerkesztett $c_0\mathbf{E}+c_1\mathbf{A}+\cdots+c_n\mathbf{A}^n$ alakú matrix-polinomok közti racionális egész műveletek ugyanazon szabályok szerint végezhetők, mint skalár-polinomok esetén, továbbá, hogy egy változót tartalmazó racionális egész skalár identitások érvényesek maradnak akkor, ha bennük a változót tetszőleges quadratikus matrix-szal, az egységet pedig $\mathbf{E}$-vel helyettesítjük.

5. §. Noha az n-edrendű matrixok összesége gyűrűt alkot, azaz olyan mennyiségrendszert, melyben a fentiek szerint az összeadás, kivonás és szorzás definiálva vannak, (oly módon, hogy az összeadás kommutatív és asszociatív, a szorzás asszociatív és mindkét oldalról disztributív) ennek dacára az egyes matrixokhoz eddig semilyen számérték nem lett hozzárendelve. A végből, hogy a skalár-aritmetika egyes tételei a matrix-aritmetikába átvihetők legyenek, kívánatos minden quadratikus $\mathbf{A}$ matrixhoz egy olyan, $|\mathbf{A}|$-val jelölendő számértéket hozzárendelni, mely — a skalár $|ab| = |a|\,|b|$ azonosság mintájára — bármely két matrix esetén eleget tesz az $|\mathbf{A}\mathbf{B}| = |\mathbf{A}|\,|\mathbf{B}|$ azonosságnak.

Ilyen hozzárendelés lehetőségét és egyértelműségét biztosítja a következő [6]

I. Tétel: *Ha* $|\mathbf{A}|$ *olyan minimális fokú racionális egész függvénye az* $\mathbf{A}$ *matrix (független változóknak tekintendő)* a_{ij} *elemeinek, mely a*

$$|\mathbf{A}\mathbf{B}| = |\mathbf{A}|\cdot|\mathbf{B}| \tag{12}$$

[6] Az I. Tétel bizonyítását illetőleg l. pl. C. C. Mac Duffee, The Theory of Matrices. Berlin, 1933. pp. 6—7.

azonosságot kielégíti, akkor $|\mathbf{A}|$ *az* $\mathbf{A}$ *matrix determinánsával egyenlő.* (Ennek megfelelően szükség esetén $|\mathbf{A}|$ helyett a $\det(\mathbf{A})$ jelet is fogjuk használni.)

Későbbi alkalmazásokra való tekintettel idézzük a determinánselmélet alábbi tételeit.[7]

1. Ha az

$$(-1)^{\alpha+\beta}|\mathbf{A}_{1\,2\ldots\beta-1,\,\beta+1,\ldots n}^{1\,2\ldots\alpha-1,\,\alpha+1,\ldots n}| = A_{\alpha\beta}$$

jelölést az $a_{\alpha\beta}$ elem komplementer minora számára bevezetjük, akkor a determinánsnak sor vagy oszlop szerinti kifejtésére vonatkozó identitások a következő alakban írhatók:

$$\sum_{\nu=1}^{n} a_{i\nu}A_{j\nu} = \sum_{\nu=1}^{n} a_{\nu i}A_{\nu j} = \delta_{ij}|\mathbf{A}| = \delta_{ij}|\mathbf{A}^*|. \tag{13}$$

2. Cauchy—Binet tétele: Ha $\mathbf{A} = \mathbf{A}_k^n$, $\mathbf{B} = \mathbf{B}_m^k$, akkor

$$|(\mathbf{AB})_{\mu_1\mu_2\ldots\mu_r}^{\nu_1\nu_2\ldots\nu_r}| = |\mathbf{A}_{1\,2\ldots k}^{\nu_1\nu_2\ldots\nu_r}\cdot\mathbf{B}_{\mu_1\mu_2\ldots\mu_r}^{1\,2\ldots k}| = \sum_{(\gamma)} |\mathbf{A}_{\gamma_1\gamma_2\ldots\gamma_r}^{\nu_1\nu_2\ldots\nu_r}|\,|\mathbf{B}_{\mu_1\mu_2\ldots\mu_r}^{\gamma_1\gamma_2\ldots\gamma_r}|, \tag{14}$$

ahol $\sum\limits_{(\gamma)}$ az $1, 2, \ldots, k$ számok összes r-edosztályú $\gamma_1, \gamma_2, \ldots \gamma_r$ kombinációira kiterjedő összegezést jelent. (Ennek speciális esete $k = m = n = r$ esetben (12) és $r = 1$ esetben (7).)

3. A $|\lambda\mathbf{E}_n - \mathbf{A}_n^n|$ determináns a λ skalár változónak n-edfokú polinomja, melynek explicit kifejtése:

$$|\lambda\mathbf{E} - \mathbf{A}| = \lambda^n - \lambda^{n-1}\sum_i a_{ii} + \lambda^{n-2}\sum_i\sum_j |\mathbf{A}_{ij}^{ij}| - \cdots \pm |\mathbf{A}|. \tag{15}$$

4.

$$\frac{d^k}{d\lambda^k}|\lambda\mathbf{E} - \mathbf{A}| = \sum_{(\alpha)} |(\lambda\mathbf{E} - \mathbf{A})_{\alpha_1,\ldots\alpha_{n-k}}^{\alpha_1\ldots\alpha_{n-k}}|, \tag{16}$$

ahol $\sum\limits_{(\alpha)}$ az $1, 2, 3 \ldots n$ számok összes $n-k$-ad osztályú $\alpha_1, \alpha_2, \ldots \alpha_{n-k}$ kombinációira vonatkozó összegezést jelent.

6. §. A determináns fogalmának a matrix-elmélet szempontjából való jelentősége először az osztás műveletének bevezetésénél mutatkozik. A skalár aritmetikában az a számmal való osztás ekvivalens a reciprok értékével való szorzással. Az a szám reciprokát pedig az $ax = xa = 1$ egyenlet definiálja, melyből nyilvánvaló, hogy csupán 0-tól különböző a számnak van reciprokja.

Ennek alapján közelfekvő, hogy a matrixok osztásának általános tárgyalása előtt azt a kérdést vizsgáljuk, van-e adott $\mathbf{A}$ matrixnak reciprokja, azaz van-e olyan $\mathbf{X}$ matrix, mely az

$$\mathbf{AX} = \mathbf{E}, \text{ ill. } \mathbf{XA} = \mathbf{E} \tag{17}$$

egyenletek valamelyikét kielégíti. Ezen egyenletek megoldhatóságára azonnal

[7] L. pl. A. C. Aitken, Detereminants and Matrices. Edinburgh, 1948.

nyerünk egy szükséges feltételt, ha a bal- és jobboldali matrixok determinánsaira térünk át és a (12) alaptulajdonságot felhasználjuk. Ekkor az $|\mathbf{AX}| = |\mathbf{E}|$ és $|\mathbf{XA}| = |\mathbf{E}|$ egyenletek mindegyikéből következik

$$|\mathbf{A}|\cdot|\mathbf{X}| = |\mathbf{X}|\cdot|\mathbf{A}| = 1.$$

Tehát a (17) egyenletek bármelyikének csak akkor lehet megoldása, ha az $\mathbf{A}$ matrix determinánsa 0-tól különböző.

Ki fogjuk mutatni, hogy a reciprok-matrix létezéséhez ez a feltétel elegendő, a feltétel teljesülése esetén a (17) egyenleteknek egyetlen közös megoldásuk van, tehát jobb- és baloldali reciprok-matrixok megkülönböztetése fölösleges, és az $\mathbf{A}^{-1}$ reciprok-matrix racionális műveletekkel megszerkeszthető $\mathbf{A}$ elemeiből.

7. §. A reciprok-matrix megszerkesztése céljából először az adjungált matrix fogalmát vezetjük be. Jelöljük most is az $\mathbf{A} = [a_{ij}]$ quadratikus matrix determinánsát $|\mathbf{A}|$-val és az a_{ij} elemhez tartozó előjeles aldeterminánst (az a_{ij} elem komplementer minorát) A_{ij}-vel. Ekkor az $[A_{ji}]$ matrixot $\mathbf{A} = [a_{ij}]$ adjungáltjának nevezzük és adj. $\mathbf{A}$-val jelöljük. Eszerint:

$$\text{adj. } \mathbf{A} = \text{adj.} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdot & \cdot & \cdots & \cdot \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{vmatrix}. \tag{18}$$

Külön kiemeljük, hogy $\mathbf{A}$ adjungáltja az $|\mathbf{A}|$ minoraiból alkotott matrixnak a *transzponált*ja.

A determináns fentebb idézett (13) tulajdonsága alapján

$$\sum_{\nu=1}^{n} a_{i\nu}A_{j\nu} = \sum_{\nu=1}^{n} a_{\nu i}A_{\nu j} = 0, \quad \text{ha} \quad i \neq j \tag{19$_1$}$$

$$\sum_{\nu=1}^{n} a_{i\nu}A_{i\nu} = \sum_{\nu=1}^{n} a_{\nu i}A_{\nu i} = |\mathbf{A}|, \tag{19$_2$}$$

ennek következtében

$$\mathbf{A}\cdot\text{adj. } \mathbf{A} = \text{adj. } \mathbf{A}\cdot\mathbf{A} = \begin{vmatrix} |\mathbf{A}| & 0 & \cdot & \cdot & 0 \\ 0 & |\mathbf{A}| & \cdots & & 0 \\ \cdot & \cdot & \cdots & & \cdot \\ 0 & 0 & \cdots & & |\mathbf{A}| \end{vmatrix} = |\mathbf{A}|\cdot\mathbf{E}. \tag{20}$$

Ebből az azonosságból kiolvasható a következő alapvető

II. Tétel: *Egy matrix a saját adjungáltjával mindig kommutálható és szorzatuk az egység-matrixtól csupán az* $|\mathbf{A}|$ *skalár szorzóban különbözik.*

A (20) identitás közvetlenül mutatja, hogy a (17) egyenletek mindegyike ugyanazon $\mathbf{X}$ matrix-szal kielégíthető, ha csak $|\mathbf{A}| \neq 0$. Ekkor ugyanis az $\mathbf{X} = |\mathbf{A}|^{-1}$ adj. $\mathbf{A}$ matrix mind a két egyenletet kielégíti, tehát a bal- és jobboldali reciprok matrixok megegyeznek és közös értékük $\mathbf{A}^{-1}$-nel jelölhető:

$$\mathbf{A}^{-1} = \frac{\text{adj.}\,\mathbf{A}}{|\mathbf{A}|}. \tag{21}$$

Valamint az $a \cdot x = 1$ egyenletnek nincs megoldása $a = 0$ esetén, épp úgy az $\mathbf{AX} = \mathbf{E}$ és $\mathbf{XA} = \mathbf{E}$ egyenleteknek sincs megoldásuk $|\mathbf{A}| = 0$ esetén. Ennek megfelelően a 0 determinánssal bíró matrixokat szingulárisoknak (vagy nem-invertálhatóknak) fogjuk nevezni.

Nyilván $|\mathbf{A}| \neq 0$ esetén a fentiek alapján $\mathbf{A}$-nak tetszőleges negatív egészkitevőjű hatványai is értelmezve vannak és azok úgy egymásközt, mint $\mathbf{A}$ természetes kitevőjű hatványaival felcserélhetők.

Ezekután tetszőleges matrixok osztását (nem-szinguláris osztó mellett) az $\mathbf{AX} = \mathbf{B}$ egyenletnél $\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$-vel, az $\mathbf{XA} = \mathbf{B}$ egyenletnél pedig $\mathbf{X} = \mathbf{BA}^{-1}$-nel értelmezzük. Az osztás eredménye tehát függ az osztandó és osztó sorrendjétől.

Matrix-szorzat reciprokjának a képzése a

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} \tag{22}$$

azonosság szerint történik, melynek helyességét a

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

egyenlőségek igazolják.

Ha az a és b skalárok kielégítik az $a \neq 0, ab = 0$ relációkat, akkor ezekből természetesen $b = 0$ következik. Matrixok esetén $\mathbf{A} \neq 0$, $\mathbf{AB} = 0$-ból általában nem következik $\mathbf{B} = 0$. Ha azonban $|\mathbf{A}| \neq 0$, akkor $\mathbf{A}^{-1}$ létezik és $\mathbf{AB} = 0$-ból $\mathbf{A}^{-1}$-nel való szorzással $\mathbf{B} = 0$ következik.

Ha az a, b és c skalárok kielégítik az $ac = bc = cb \neq 0$ relációkat, akkor ezekből nyilván $a = b$ következik. Viszont az előbbivel analóg meggondolás azt mutatja, hogy $\mathbf{AC} = \mathbf{BC} \neq 0$-ból csak akkor következik $\mathbf{A} = \mathbf{B}$, ha $|\mathbf{C}| \neq 0$.

Az $\mathbf{AC} = \mathbf{CB} \neq 0$ relációkból pedig $|\mathbf{C}| \neq 0$ esetén sem következik általában $\mathbf{A} = \mathbf{B}$. Mindamellett $|\mathbf{C}| \neq 0$ esetén a

$$\mathbf{AC} = \mathbf{CB}, \text{ ill. } \mathbf{A} = \mathbf{CBC}^{-1} \tag{23}$$

egyenletekkel összekapcsolt $\mathbf{A}$ és $\mathbf{B}$ matrixok közt szoros kapcsolat áll fenn. Először is (23)-ból következik

$$|\mathbf{A}| = |\mathbf{C}|\,|\mathbf{B}|\,|\mathbf{C}|^{-1} = |\mathbf{B}|.$$

Továbbá, ha $\mathbf{A}_1 = \mathbf{CB}_1\mathbf{C}^{-1}$ és $\mathbf{A}_2 = \mathbf{CB}_2\mathbf{C}^{-1}$, akkor

$$\mathbf{A}_1 + \mathbf{A}_2 = \mathbf{C}(\mathbf{B}_1 + \mathbf{B}_2)\mathbf{C}^{-1};\ \mathbf{A}_1\mathbf{A}_2 = \mathbf{CB}_1\mathbf{C}^{-1}\mathbf{CB}_2\mathbf{C}^{-1} = \mathbf{CB}_1\mathbf{B}_2\mathbf{C}^{-1}.$$

Ha tehát a $\mathbf{A} = \mathbf{CBC}^{-1}$ kapcsolatban álló matrixokat *hasonlók*nak nevezzük és a vonatkozást $\mathbf{A} \sim \mathbf{B}$-vel jelöljük, akkor $\mathbf{A}_1 \sim \mathbf{B}_1$, $\mathbf{A}_2 \sim \mathbf{B}_2$-ből $\mathbf{A}_1 + \mathbf{A}_2 \sim \mathbf{B}_1 + \mathbf{B}_2$, $\mathbf{A}_1\mathbf{A}_2 \sim \mathbf{B}_1\mathbf{B}_2$ következik. Minthogy továbbá valamely $\mathbf{A}$

matrixnak bármely skalár együtthatós $f(\mathbf{A})=\sum_{0}^{n} c_\nu \mathbf{A}^\nu$ polinomja **A**-ból összeadásokkal és szorzásokkal épül fel, tehát az előbbiek alapján érvényes a

III. Tétel: *Ha* $\mathbf{A} \sim \mathbf{B}$ *és* $f(x)$ *tetszőleges skalár polinom, akkor*

$$f(\mathbf{A}) \sim f(\mathbf{B}) \quad \textit{és} \quad |f(\mathbf{A})| = |f(\mathbf{B})|.$$

Pl. $f(x)=\lambda-x$ esetén $\mathbf{A} \sim \mathbf{B}$-ből következik

$$\lambda\mathbf{E}-\mathbf{A} \sim \lambda\mathbf{E}-\mathbf{B} \quad \text{és} \quad |\lambda\mathbf{E}-\mathbf{A}| = |\lambda\mathbf{E}-\mathbf{B}|.$$

II. Fejezet

Matrix rangja és diadikus felbontása. Vektorok lineáris függése. Projektor-matrix felbontása biortogonális diádok összegére

1. §. A matrix elmélet egyik legfontosabb fogalomalkotása a matrix rangja. Valamely **A** matrix rangja: $\varrho(\mathbf{A})$ számára két definíciót adunk, melyek egymással ekvivalenseknek fognak bizonyulni.

I. Definíció. **A** *rangja:* $\varrho(\mathbf{A})$ *egyenlő az*

$$\mathbf{A}=\begin{bmatrix} a_{11} \ldots a_{1m} \\ \vdots \\ a_{n1} \quad a_{nm} \end{bmatrix} = \sum_{k=1}^{r} \begin{bmatrix} u_{1k} \\ u_{2k} \\ \vdots \\ u_{nk} \end{bmatrix} [v_{k1}, v_{k2}, \ldots, v_{km}] = [\mathbf{u}_1 \ldots \mathbf{u}_r] \begin{bmatrix} \mathbf{v}_1^* \\ \vdots \\ \mathbf{v}_r^* \end{bmatrix} = \sum \mathbf{u}_k \mathbf{v}_k^* \tag{1}$$

alakú előállításhoz szükséges diádok minimális számával. Rövidség kedvéért a továbbiakban minden olyan (1) alakú előállítást, mely ennek a minimumkövetelménynek eleget tesz, az **A** matrix egy minimális diadikus előállításának fogunk nevezni.

Célszerű már e helyen bevezetni a vektorok lineáris függetlenségének a fogalmát.

Az $\mathbf{u}_1, \mathbf{u}_2, \ldots \mathbf{u}_r$ *vektorokat akkor nevezzük lineárisan függetleneknek, ha az*

$$x_1\mathbf{u}_1+x_2\mathbf{u}_2+\ldots+x_r\mathbf{u}_r=0 \tag{2}$$

egyenlet fennállásából $x_1=x_2=\ldots=x_r=0$ *következik.* Ha viszont az $\mathbf{u}_1, \mathbf{u}_2, \ldots \mathbf{u}_r$ vektorok közt fennáll olyan (2) alakú egyenlet, melyben nem mindegyik $x_i=0$, akkor azt mondjuk, hogy a vektorok közt lineáris összefüggés van.

Ebből a definícióból rögtön következik, hogy **A** *bármely minimális diadikus előállításában úgy a baloldali* $\mathbf{u}_1, \mathbf{u}_2, \ldots \mathbf{u}_r$ *tényezők, valamint a jobboldali* $\mathbf{v}_1^*, \mathbf{v}_2^*, \ldots, \mathbf{v}_r^*$ *tényezők lineárisan függetlenek.* Ha ugyanis pl. a baloldali tényezők közt fennállna egy (2) alakú egyenlet, melyben egyik skalár együttható, pl. $x_r \neq 0$, akkor abból $\mathbf{u}_r$ kifejezhető volna

$$\mathbf{u}_r = -\frac{x_1}{x_r}\mathbf{u}_1 - \frac{x_2}{x_r}\mathbf{u}_2 - \ldots - \frac{x_{r-1}}{x_r}\mathbf{u}_{r-1}$$

alakban és ennek folytán létezne egy

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

$$=\mathbf{u}_1\left(\mathbf{v}_1^*-\frac{x_1}{x_r}\mathbf{v}_r^*\right)+\mathbf{u}_2\left(\mathbf{v}^*-\frac{x_2}{x_r}\mathbf{v}_r^*\right)+\ldots+\mathbf{u}_{r-1}\left(\mathbf{v}_{r-1}^*-\frac{x_{r-1}}{x_r}\mathbf{v}_r^*\right)$$

alakú előállítás, tehát $\sum_{k=1}^{r}\mathbf{u}_k\mathbf{v}_k^*$ nem lenne minimális, feltevésünkkel ellentétben. (Később, a **4.** §.-ban ki fogjuk mutatni, hogy ha $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$, valamint $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ lineárisan függetlenek, akkor a $\sum_{k=1}^{r}\mathbf{u}_k\mathbf{v}_k^*$ előállítás minimális.)

II. Definíció. **A** *rangja:* $\varrho(\mathbf{A})$ *egyenlő a belőle kiválasztható maximális rendű nem-szinguláris minormatrixok rendszámával.*

2. §. $\varrho(\mathbf{A})$ megállapítására először egy olyan módszert ismertetünk, mely az I.) definíción alapszik és egyszersmind számítástechnikailag is alkalmas eljárást ad a rang meghatározására. Utólag verifikálni fogjuk, hogy az így meghatározott rang a II.) definíció követelményeit is kielégíti.

$\mathbf{A}=0$ esetén nyilván $\varrho(\mathbf{A})=0$.

Ha $\mathbf{A}\neq 0$, akkor van legalább egy 0-tól különböző eleme: $a_{\beta_1\gamma_1}$ tehát képezhető a

$$\mathbf{A}-\frac{1}{a_{\beta_1\gamma_1}}\begin{bmatrix}a_{1\gamma_1}\\ a_{2\gamma_1}\\ \vdots\\ a_{n\gamma_1}\end{bmatrix}[a_{\beta_1 1}a_{\beta_1 2}\ldots a_{\beta_1 m}]=\mathbf{A}-\mathbf{u}_1\mathbf{v}^*=\mathbf{A}'=[a'_{ij}] \qquad (3\cdot1)$$

$$(u_{1\gamma_1}v_{\beta_1 1}=a_{\beta_1\gamma_1}\neq 0)$$

matrix, melynek γ_1-dik oszlopa és β_1-dik sora csupa 0 elemet tartalmaz. Ha $\mathbf{A}'\neq 0$, akkor van legalább egy 0-tól különböző $a'_{\beta_2\gamma_2}$ eleme, tehát képezhető a

$$\mathbf{A}'-\frac{1}{a'_{\beta_2\gamma_2}}\begin{bmatrix}a'_{1\gamma_2}\\ a'_{2\gamma_2}\\ \vdots\\ a'_{n\gamma_2}\end{bmatrix}[a'_{\beta_2 1}a'_{\beta_2 2}\ldots a'_{\beta_2 m}]=\mathbf{A}'-\mathbf{u}_2\mathbf{v}_2^*=\mathbf{A}''=[a''_{ij}] \qquad (3\cdot2)$$

$$(u_{2\gamma_2}v_{\beta_2 2}=a'_{\beta_2\gamma_2}\neq 0)$$

matrix, melynek γ_1 és γ_2-dik oszlopai, valamint β_1 és β_2-dik sorai csupa 0 elemet tartalmaznak. (3·1) és (3·2) szerint

$$\mathbf{A}-\mathbf{u}_1\mathbf{v}_1^*-\mathbf{u}_2\mathbf{v}_2^*=\mathbf{A}''.$$

Nyilvánvaló, hogy a redukciós eljárást folytatva, eljutunk egy olyan $r\leqq n, m$ értékhez, melyre $\mathbf{A}^{(r)}=0$ és ezzel előállítottuk az adott matrixot a

$$\mathbf{A}=\mathbf{u}_1\mathbf{v}_1^*+\mathbf{u}_2\mathbf{v}^*+\ldots+\mathbf{u}_r\mathbf{v}_r^* \qquad (4)$$

alakban (a diádok előtt álló $a_{\beta_1\gamma_1}^{-1}$ stb. skalártényezőket tetszőleges arányban oszthatjuk szét az $\mathbf{u}_1, \mathbf{v}^*$ stb. vektortényezők közt).

Az I. definíció szerint **A** rangja nem változik, ha a sorok, ill. oszlopok sorrendjét felcseréljük, mert a sorok, ill. oszlopok cseréje a baloldali, ill.

jobboldali vektortényezők koordinátáinak felcserélésével ekvivalens. Ennek alapján — a könnyebb áttekintés kedvéért — vigyük át a $\beta_1, \beta_2, \ldots, \beta_r$-dik sort, úgyszintén a $\gamma_1, \gamma_2, \ldots, \gamma_r$-dik oszlopot az első, második, ... r-dik helyre; a többi sor és oszlop elrendezése tetszőleges. Ekkor a felcseréléssel nyert $\tilde{\mathbf{A}}$ matrix a következő alakú:[8]

$$\tilde{\mathbf{A}} = \sum_{k=1}^{r} \tilde{\mathbf{u}}_k \tilde{\mathbf{v}}_k^* = \sum_{k=1}^{r} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ u_{k,k} \\ u_{k+1,k} \\ \vdots \\ u_{nk} \end{bmatrix} [0 \ldots 0 \; v_{kk} \ldots v_{km}] = \tag{5}$$

$$= \begin{bmatrix} u_{11} & 0 & \ldots & 0 \\ u_{21} & u_{22} & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ u_{r1} & u_{r2} & \ldots & u_{rr} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ u_{n1} & u_{n2} & \ldots & u_{nr} \end{bmatrix} \begin{bmatrix} v_{11} & v_{12} & \ldots & v_{1m} \\ 0 & v_{22} & \ldots & v_{2m} \\ \cdot & \cdot & & \cdot \\ 0 & 0 & v_{rr} \ldots & v_{rm} \end{bmatrix} = \mathbf{UV}^*,$$

melyből nyilvánvaló, hogy

$$\pm \left| \mathbf{A}^{\beta_1 \beta_2 \ldots \beta_r}_{\gamma_2 \gamma_2 \ldots \gamma_r} \right| = \left| \tilde{\mathbf{A}}^{12 \ldots r}_{12 \ldots r} \right| = \prod_{k=1}^{r} u_{kk} v_{kk} \neq 0. \tag{6}$$

Ezzel kimutattuk, hogy az (5) előállítás minimális, mert r-nél kevesebb diád összegének minden r-edrendű minormatrixa szinguláris, (6)-tal ellentétben.

(5)-ből nyilvánvaló továbbá, hogy $\mathbf{A}$-nak minden $r+1$-edrendű minormatrixa szinguláris és ezzel (6) figyelembe vétele mellett az is verifikálva van, hogy az (1) előállítással meghatározott rangszám a klasszikus II.) definíció követelményeit is kielégíti.

A fenti redukciós eljárás effektív kivitele csupán másodrendű determinánsok szukcesszív számítását teszi szükségessé, miként azt az alábbi példákból látható.

$$\begin{bmatrix} -26 & -18 & -27 \\ 21 & 15 & 21 \\ 12 & 8 & 13 \end{bmatrix} = \begin{bmatrix} -9 \\ {}^{15}/_2 \\ 4 \end{bmatrix} [3 \; 2 \; {}^{13}/_4] + \begin{bmatrix} 1 \\ -{}^{3}/_2 \\ 0 \end{bmatrix} [1 \; 0 \; {}^{9}/_4] = \begin{bmatrix} -9 & 1 \\ {}^{15}/_2 & -{}^{3}/_2 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} 3 & 2 & {}^{13}/_4 \\ 1 & 0 & {}^{9}/_4 \end{bmatrix}.$$

[8] A β_i és γ_j indexek alkalmas választásával előzetes sor- és oszlopcsere *nélkül* is elérhető, hogy az (5)-ben szereplő tényezők *egyike* trapéz- ill. háromszög-matrix legyen.

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

$$=\begin{vmatrix} 1 \\ 0 \\ -1 \\ -2 \end{vmatrix}[1\ 2\ 3\ 4\ 5]+\begin{vmatrix} 1 \\ 2 \\ 0 \\ -1 \end{vmatrix}[2\ 1\ 3\ 1\ 0]+\begin{vmatrix} 0 & 0 & 0 & 0 & 0 \\ 3 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 3 & 2 & 1 & 0 & 0 \end{vmatrix}=$$

$$=\begin{vmatrix} 1 \\ 0 \\ -1 \\ -2 \end{vmatrix}[1\ 2\ 3\ 4\ 5]+\begin{vmatrix} 1 \\ 2 \\ 0 \\ -1 \end{vmatrix}[2\ 1\ 3\ 1\ 0]+\begin{vmatrix} 0 \\ 1 \\ 0 \\ 1 \end{vmatrix}[3\ 2\ 1\ 0\ 0]=$$

$$=\begin{vmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ -1 & 0 & 0 \\ -2 & -1 & 1 \end{vmatrix}\begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 1 & 0 \\ 3 & 2 & 1 & 0 & 0 \end{vmatrix};$$

megfelelő átrendezéssel

$$\tilde{\mathbf{A}}=\begin{vmatrix} -1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 2 & 1 \\ -2 & 1 & 1 \end{vmatrix}\begin{vmatrix} 5 & 4 & 1 & 2 & 3 \\ 0 & 1 & 2 & 1 & 3 \\ 0 & 0 & 3 & 2 & 1 \end{vmatrix}.$$

3. §. A fenti redukciós eljárás a diádok $\mathbf{u}_k, \mathbf{v}_k$ tényezőit rekurzív módon szolgáltatja. Egy közismert determinánstétel[9] felhasználásával azonban (3·1), (3·2), ... alapján a diadikus előállítás a következő explicit alakban is felírható:

$$\mathbf{A}=\sum_{k=1}^{r}\frac{1}{D_{k-1}D_k}\begin{vmatrix} a_{\beta_1\gamma_1} & a_{\beta_1\gamma_2} & \dots a_{\beta_1\gamma_k} \\ a_{\beta_2\gamma_1} & a_{\beta_2\gamma_2} & \dots a_{\beta_2\gamma_k} \\ \vdots & \vdots & \vdots \\ a_{\beta_{k-1}\gamma_1} & a_{\beta_{k-1}\gamma_2} & \dots a_{\beta_{k-1}\gamma_k} \\ \mathbf{a}_{\gamma_1} & \mathbf{a}_{\gamma_2} & \dots \mathbf{a}_{\gamma_k} \end{vmatrix}\begin{vmatrix} a_{\beta_1\gamma_1}a_{\beta_1\gamma_2}\dots a_{\beta_1\gamma_{k-1}}\mathbf{a}^{\beta_1} \\ a_{\beta_2\gamma_1}a_{\beta_2\gamma_2}\dots a_{\beta_2\gamma_{k-1}}\mathbf{a}^{\beta_2} \\ \\ \\ a_{\beta_k\gamma_1}a_{\beta_k\gamma_2}\dots a_{\beta_k\gamma_{k-1}}\mathbf{a}^{\beta_k} \end{vmatrix},$$

ahol

$$D_0=1,\ D_k=\begin{vmatrix} a_{\beta_1\gamma_1}a_{\beta_1\gamma_2}\dots a_{\beta_1\gamma_h} \\ a_{\beta_2\gamma_1}a_{\beta_2\gamma_2}\dots a_{\beta_2\gamma_k} \\ \dots \\ a_{\beta_k\gamma_1}a_{\beta_k\gamma_2}\dots a_{\beta_k\gamma_k} \end{vmatrix},\quad \mathbf{a}_i=\begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix},\quad \mathbf{a}^j=[a_{j1}, a_{j2}, \dots a_{jn}].$$

Adott matrixnak minimális számú diád összegeként való előállítása nyilván nem egyértelmű. Kimutatható azonban, hogy ha egy $\mathbf{A}=\mathbf{U}\mathbf{V}^*$ minimális

[9] L. pl. E. Pascal, Repertorium der höheren Mathematik. I. Erste Hälfte. Leipzig, 1910. pp. 120—121.

előállítás ismert, akkor abból az összes minimális előállítások

$$\mathbf{A}=(\mathbf{U}\mathbf{M})(\mathbf{M}^{-1}\mathbf{V}^*)$$

alakban adódnak, ahol $\mathbf{M}$ tetszőleges r-edrendű nem-szinguláris matrix.

A definíciók közvetlen következménye, hogy egy matrix rangja nem lehet nagyobb, mint a sorok, ill. az oszlopok száma. Quadratikus matrix rangja tehát nem nagyobb a rendszámnál és azzal csak akkor egyenlő, ha a matrix nem-szinguláris, azaz determinánsa 0-tól különböző. Ez esetben a matrix (5) előállításában trianguláris matrix-tényezők szerepelnek.

Az alábbi tételek, melyek összeg és szorzat rangjára vonatkoznak, szintén a definíciók egyszerű következményei.

IV. Tétel: *Összeg rangja nem nagyobb a tagok rangjainak összegénél:*

$$\varrho(\mathbf{A}+\mathbf{B})\leqq\varrho(\mathbf{A})+\varrho(\mathbf{B}). \tag{7}$$

Ha ugyanis $\varrho(\mathbf{A})=r$, $\varrho(\mathbf{B})=s$, akkor $\mathbf{A}+\mathbf{B}$ diadikus előállításához *legfeljebb* $r+s$ diád szükséges. Tehát $\varrho(\mathbf{A}+\mathbf{B})\leqq r+s$, q. e. d.

IV'. Tétel: 1. *Matrix rangja szorzás által nem növekszik.*

$$\varrho(\mathbf{A}\mathbf{B})\leqq\varrho(\mathbf{A}). \tag{8}$$

Legyen $\varrho(\mathbf{A})=r$. Ekkor $\mathbf{A}=\sum_1^r\mathbf{u}_k\mathbf{v}_k^*$, tehát $\mathbf{A}\mathbf{B}=\sum_1^r\mathbf{u}_k(\mathbf{v}_k^*\mathbf{B})$, következésképpen $\mathbf{A}\mathbf{B}$ diadikus előállításához *legfeljebb* r diád szükséges, azaz $\varrho(\mathbf{A}\mathbf{B})\leqq r$. Q. e. d.

Corollárium. Matrix rangja nem-szinguláris matrixszal való szorzás esetén nem változik.

Tegyük fel, hogy $\mathbf{M}$ nem szinguláris, azaz $\mathbf{M}^{-1}$ létezik. Ekkor (8) szerint

$$\varrho(\mathbf{A}\mathbf{M})\leqq\varrho(\mathbf{A})$$

és

$$\varrho(\mathbf{A})=\varrho\{(\mathbf{A}\mathbf{M})\mathbf{M}^{-1}\}\leqq\varrho(\mathbf{A}\mathbf{M}),$$

vagyis

$$\varrho(\mathbf{A}\mathbf{M})=\varrho(\mathbf{A}). \qquad \text{q. e. d.}$$

4. §. A matrix rangját evidenciába helyező (1) diadikus előállítás felhasználásával igen egyszerűen tárgyalható a lineáris algebrának következő két ekvivalens feladata:

I. *adott vektorok közti lineáris összefüggések megállapítása,*

II. *homogén, lineáris egyenletrendszer megoldása.*

Adott $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ n-dimenziós vektorok közti lineáris összefüggések megállapítása ugyanis a fentebb adott definíció szerint az $x_1, x_2, \ldots, x_m$ skalár ismeretleneket tartalmazó

$$\mathbf{a}_1x_1+\mathbf{a}_2x_2+\cdots+\mathbf{a}_mx_m=0 \tag{2}$$

egyenlet, vagy részletesen kiírva, a

$$\begin{aligned} a_{11}x_1+a_{12}x_2+\cdots+a_{1m}x_m&=0\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2m}x_m&=0\\ \cdots\cdots\cdots\cdots&\\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nm}x_m&=0 \end{aligned} \tag{2·1}$$

homogén, lineáris egyenletrendszer, illetve az ezzel ekvivalens

$$\begin{bmatrix} a_{11}\,a_{12}\,\ldots\,a_{1m}\\ a_{21}\,a_{22}\,\ldots\,a_{2m}\\ \cdots\cdots\cdots\\ a_{n1}\,a_{n2}\,\ldots\,a_{nm}\end{bmatrix}\begin{bmatrix}x_1\\x_2\\ \vdots\\ x_m\end{bmatrix}=\mathbf{A}\mathbf{x}=0 \tag{2·2}$$

matrix-egyenlet megoldásainak diszkusszióját teszi szükségessé.

Eszerint m számú n-dimenziós vektor közti lineáris összefüggések megállapítása m ismeretlent tartalmazó és n homogén lineár egyenletből álló rendszer megoldásával ekvivalens feladat.

Legyen az $\mathbf{A}$ koordináta-matrix (ill. együttható-matrix) rangja: $r \leqq m, n$, továbbá tegyük fel, hogy az egyenletek, valamint az ismeretlenek (5)-nek megfelelően át vannak számozva és legyen annak minimális diadikus előállítása: $\mathbf{A}=\mathbf{u}_1\mathbf{v}_1^*+\mathbf{u}_2\mathbf{v}_2^*+\cdots+\mathbf{u}_r\mathbf{v}_r^*$. Ekkor a megoldandó egyenletrendszerrel ekvivalens matrix-egyenlet:

$$\mathbf{A}\mathbf{x}=\mathbf{u}_1(\mathbf{v}_1^*\mathbf{x})+\mathbf{u}_2(\mathbf{v}_2^*\mathbf{x})+\cdots+\mathbf{u}_r(\mathbf{v}_r^*\mathbf{x})=0.$$

Ebből azonban az $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$ tényezők lineáris függetlensége miatt következik, hogy

$$\mathbf{v}_1^*\mathbf{x}=0,\ \mathbf{v}_2^*\mathbf{x}=0, \ldots, \mathbf{v}_r^*\mathbf{x}=0. \tag{9}$$

Az ily módon keletkező és (2)-vel ekvivalens egyenletrendszer azonban a fokozatos redukcióval nyert $\mathbf{v}_k^*$ vektoroknak (5)-ből látható struktúráját figyelembe véve a következő alakú:

$$\begin{aligned} v_{11}x_1+v_{12}x_2+\cdots+v_{1m}x_m&=0 \quad v_{11}\neq 0\\ v_{22}x_2+\cdots+v_{2m}x_m&=0 \quad v_{22}\neq 0\\ \cdots\cdots\cdots\cdots&\\ v_{rr}x_r+\cdots+v_{rm}x_m&=0 \quad v_{rr}\neq 0. \end{aligned} \tag{9·1}$$

Vizsgáljuk először az $r=m$ esetet. Ekkor a (9·1) rendszernek nyilván csupán a triviális megoldása: $x_1=x_2=\ldots=x_m=0$ létezik, tehát érvényes a

V. Tétel: *Ha* $\varrho[\mathbf{a}_1\mathbf{a}_2\ldots\mathbf{a}_m]=m$, *azaz a vektorok koordináta-matrixának* (az együtthatók matrixának) *rangja egyenlő a vektorok számával, akkor a vektorok lineárisan függetlenek* (csupán a triviális megoldás létezik).

Vizsgáljuk ezek után az $r<m$ esetet. Ekkor a (9·1) rendszerben $x_{r+1}, x_{r+2}, \ldots, x_m$-nek tetszőleges értékeket adva, a rendszer az $x_1, x_2, \ldots, x_r$ ismeretlenekre nézve nyilván rekurzív módon egyértelműen megoldható:

$$\begin{aligned} x_r&=b_{r,r+1}x_{r+1}+b_{r,r+2}x_{r+2}+\cdots+b_{rm}x_m\\ x_{r-1}&=b_{r-1,r+1}x_{r+1}+\cdots+b_{r-1,m}x_m\\ \cdots&\cdots\cdots\cdots\\ x_1&=b_{1,r+1}x_{r+1}+\cdots+b_{1,m}x_m. \end{aligned} \tag{9·2}$$

Ezek az egyenletek mindenekelőtt a (2·1) homogén lineár egyenletrendszer megoldásait, mint egy $m-r$ parametertől függő sokaságot szolgáltatják. Másszóval, az $\mathbf{A}\mathbf{x}=0$ egyenletnek a következő $m-r$ lineárisan független vektor tesz eleget:

$$\mathbf{x}=\mathbf{b}_{r+1}=\begin{bmatrix} b_{1,r+1} \\ \vdots \\ b_{r,r+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{x}=\mathbf{b}_{r+2}\begin{bmatrix} b_{1,r+2} \\ \vdots \\ b_{r,r+2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \mathbf{x}=\mathbf{b}_m=\begin{bmatrix} b_{1m} \\ \vdots \\ b_{rm} \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Ha továbbá a paramétereknek rendre az

$$x_{r+1}=0,\ldots,x_{k-1}=0, x_k=-1, x_{k+1}=0,\ldots,x_m=0$$

értékeket adjuk, akkor (2) és (9.2)-ből a

$$\mathbf{a}_1 b_{1k}+\mathbf{a}_2 b_{2k}+\cdots+\mathbf{a}_r b_{rk}=\mathbf{a}_k \qquad (k=r+1, r+2,\ldots,m)$$

egyenleteket nyerjük, melyek alapján kimondható a

VI. Tétel: *Ha* $\varrho[\mathbf{a}_1\mathbf{a}_2\ldots\mathbf{a}_m]=r<m$, *azaz a vektorok koordinátamatrixának rangja: r kisebb, mint a vektorok száma: m, akkor a vektorok közt van r számú lineárisan független:* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$, *míg a többiek:* $\mathbf{a}_{r+1}, \mathbf{a}_{r+2}, \ldots \mathbf{a}_m$ *ezeknek homogén lineáris függvényeiként kifejezhetők.*

1. Corollárium. Ha a vektorok száma nagyobb a dimenziónál, azaz $m>n$ és így szükségképpen $m>r$, akkor a vektorok nem lehetnek lineárisan függetlenek.

2. Corollárium. A most bizonyított tételekből következik, hogy ha $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$, valamint $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ lineárisan függetlenek, akkor $\mathbf{U}\mathbf{V}^*=\sum_1^r \mathbf{u}_k\mathbf{v}_k^*$ rangja r-rel egyenlő. Ugyanis egyrészt, miután $\mathbf{U}\mathbf{V}^*$ r diád összege, azért $\varrho(\mathbf{U}\mathbf{V}^*)\leqq r$. Másrészt a vektorok lineáris függetlensége folytán $\mathbf{U}$, ill. $\mathbf{V}^*$ tartalmaz egy nem-szinguláris $\mathbf{U}^{1\,2\ldots r}_{\gamma_1\gamma_2\ldots\gamma_r}$ ill. $\mathbf{V}^{*\beta_1\beta_2\ldots\beta_r}_{\;1\;2\ldots r}$ minor-matrixot. Eszerint $|(\mathbf{U}\mathbf{V}^*)^{\beta_1\ldots\beta_r}_{\gamma_1\ldots\gamma_r}|\neq 0$, tehát $\varrho(\mathbf{U}\mathbf{V}^*)=r$, q. e. d.

5. §. Skalárok szorzata tudvalevőleg csak akkor 0, ha legalább egyik skalártényező 0-sal egyenlő. Ennek közvetlen analogonja matrixokra a következő

VII. Tétel: *Ha* $\mathbf{A}^n_m\mathbf{B}^m_p\mathbf{C}^p_q=0$, *és* $\varrho(\mathbf{A}^n_m)=m, \varrho(\mathbf{C}^p_q)=p$, *akkor*

$$\mathbf{B}^m_p=0.$$

Bizonyítás: A tétel premisszáiból következik, hogy $\mathbf{A}$, ill. $\mathbf{C}$ tartalmaz egy nem-szinguláris $\mathbf{A}^{\alpha_1\alpha_2\ldots\alpha_m}_{1\;2\ldots m}$, ill. $\mathbf{C}^{1\,2\ldots p}_{\gamma_1\gamma_2\ldots\gamma_p}$ minor-matrixot. Továbbá $\mathbf{ABC}=0$-ból nyilván következik:

$$(\mathbf{ABC})^{\alpha_1\alpha_2\ldots\alpha_m}_{\gamma_1\gamma_2\ldots\gamma_m}=\mathbf{A}^{\alpha_1\alpha_2\ldots\alpha_m}_{1\;2\ldots m}\mathbf{B}\mathbf{C}^{1\,2\ldots p}_{\gamma_1\gamma_2\ldots\gamma_p}=0.$$

Ha itt balról $(\mathbf{A}_{1\ 2\ \ldots\ m}^{\alpha_1 \alpha_2 \ldots \alpha_m})^{-1}$-nel, jobbról $(\mathbf{C}_{\gamma_1 \gamma_2 \ldots \gamma_p}^{1\ 2\ \ldots\ p})^{-1}$-nel szorzunk, adódik $\mathbf{B} = 0$, q. e. d.

Matrix-szorzat azonban lehet 0 anélkül, hogy a tényezők bármelyike 0-matrix lenne, csupán a tényezők rangjainak kell bizonyos korlátozásnak eleget tenniök. Erre nézve érvényes a

VII'. Tétel: *Két quadratikus matrix szorzata csak akkor lehet* 0, *ha a tényezők rangjainak összege nem nagyobb a közös rendszámnál.* (Két skalár szorzatára vonatkozó analóg tétel ennek speciális esete, ha skalárt elsőrendű matrixnak, 0-tól különböző szám rangját 1-nek, 0 rangját 0-nak vesszük.)

Bizonyítás: Legyenek $\mathbf{A}$ és $\mathbf{B}$ n-edrendűek, $\varrho(\mathbf{A}) = r$, $\varrho(\mathbf{B}) = s$ és $\mathbf{AB} = 0$. A minimális diadikus $\mathbf{A} = \sum_1^r \mathbf{u}_k \mathbf{v}_k^*$, $\mathbf{B} = \sum_1^s \mathbf{w}_k \mathbf{z}_k^*$ előállítások alkalmazásával

$$\mathbf{AB} = [\mathbf{u}_1 \ldots \mathbf{u}_r] \begin{bmatrix} \mathbf{v}_1^* \\ \vdots \\ \mathbf{v}_r^* \end{bmatrix} [\mathbf{w}_1 \ldots \mathbf{w}_s] \begin{bmatrix} \mathbf{z}_1^* \\ \vdots \\ \mathbf{z}_s^* \end{bmatrix} = [\mathbf{u}_1 \ldots \mathbf{u}_r] \begin{bmatrix} \mathbf{v}_1^* \mathbf{w}_1 & \ldots & \mathbf{v}_1^* \mathbf{w}_s \\ \vdots & & \vdots \\ \mathbf{v}_r^* \mathbf{w}_1 & \ldots & \mathbf{v}_r^* \mathbf{w}_s \end{bmatrix} \begin{bmatrix} \mathbf{z}_1^* \\ \vdots \\ \mathbf{z}_s^* \end{bmatrix} = 0.$$

Minthogy az első és utolsó tényező maximális rangúak, azért az előbbi VII. tétel szerint kell, hogy a középső tényező 0 legyen. Eszerint az x_i és y_j változókra nézve identikusan

$$(\mathbf{v}_1^* x_1 + \cdots + \mathbf{v}_r^* x_r)(\mathbf{w}_1 y_1 + \cdots + \mathbf{w}_s y_s) \equiv 0.$$

Ha most feltesszük, hogy $r + s > n$, akkor a VI. tétel szerint a

$$\overline{\mathbf{v}}_1 \overline{x}_1 + \cdots + \overline{\mathbf{v}}_r \overline{x}_r = \mathbf{w}_1 y_1 + \cdots + \mathbf{w}_s y_s$$

homogén, lineáris egyenletnek van nem-triviális megoldása, melyben legalább egyik $\overline{x}_i \neq 0$ és egyik $y_j \neq 0$ (ellenkező esetben a $\mathbf{v}_i$ ill. $\mathbf{w}_j$ vektorok nem lennének lineárisan függetlenek). Ez a megoldás egy olyan vektort határoz meg, mely nem 0 és kielégíti a

$$(\mathbf{v}_1 x_1 + \cdots + \mathbf{v}_r x_r)^* (\overline{\mathbf{v}}_1 \overline{x}_1 + \cdots + \overline{\mathbf{v}}_r \overline{x}_r) = 0$$

egyenletet, ami ellenmondás. Ezzel tételünk be van bizonyítva.

6. §. Gyakran célszerű a fenti előállításban szereplő diádok tényezőit bizonyos módon normirozni. Ilyen esetekben a diadikus előállítás alakja következőképpen módosul:

VIII. Tétel: *r-edrangú matrix előállítható, mint r számú normirozott diádnak lineáris formája:*

$$\mathbf{A} = \lambda_1 \mathbf{u}_1 \mathbf{v}_1^* + \lambda_2 \mathbf{u}_2 \mathbf{v}_2^* + \cdots + \lambda_r \mathbf{u}_r \mathbf{v}_r^*, \tag{10}$$

ahol a $\lambda_1, \lambda_2, \ldots, \lambda_r$ skalár együtthatók a normirozás módjától függnek.

Valamely matrix diádikus előállításával kapcsolatban már most a matrixelmélet egyik alapvető problémája a következő alakban vethető fel: Létezik-e adott matrixnak olyan (10) alakú előállítása, melynél az $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$ és

$\mathbf{v}_1^*, \mathbf{v}_2^*, \ldots, \mathbf{v}_r^*$ vektorok biortogonális rendszert alkotnak, melynél tehát teljesülnek a

$$\mathbf{u}_k^* \mathbf{v}_l = \delta_{kl} \qquad (k, l = 1, 2, \ldots, r) \quad (11)$$

relációk.

A feltett kérdésre fokozatosan, az elmélet további kiépítése során fogunk végleges választ nyerni, azonban már e helyen megállapíthatjuk, hogy a (11) feltevésnek elegettevő (10) alakú, tehát biortogonális diádokkal való előállítás létezése esetén milyen jelentőséggel bírnak az abban szereplő λ_k skalár-tényezők és az $\mathbf{u}_k$, $\mathbf{v}_k^*$ vektortényezők.

Az

$$\mathbf{A} = \sum_{k=1}^{r} \lambda_k \mathbf{u}_k \mathbf{v}_k^* \quad , \quad \mathbf{u}_k^* \mathbf{v}_l = \delta_{kl}$$

relációkból ugyanis azonnal következik a

$$\mathbf{A}\mathbf{u}_k = \lambda_k \mathbf{u}_k \quad \text{és} \quad \mathbf{v}_k^* \mathbf{A} = \lambda_k \mathbf{v}_k^* \qquad (k = 1, 2, \ldots, r) \qquad (12)$$

egyenletek fennállása. E szerint az $\mathbf{u}_k$, ill. $\mathbf{v}_k^*$ vektorok az $\mathbf{A}$ matrixnak jobb- ill. baloldali *sajátvektorai* és a λ_k skalár a megfelelő *sajátérték*. Ezzel a megállapítással tehát a következő tételt nyertük:

IX. Tétel: *Ha egy quadratikus matrix elő van állítva, mint biortogonális diádok lineáris formája, akkor a diádok vektortényezői a matrixnak jobb- és baloldali sajátvektorai, a diádok skalár együtthatói pedig a matrix sajátértékei.*

7. §. A (10) alakú és a (11) feltételeket kielégítő előállítás lehetőségét először bizonyos speciális struktúrájú matrixok, az ú. n. projektor-matrixok esetében fogjuk megvizsgálni.

Projektor-matrixnak (v. idempotens matrixnak) nevezünk minden olyan quadratikus $\mathbf{P}$ matrixot, mely a

$$\mathbf{P}^2 = \mathbf{P} \qquad (13)$$

egyenletnek eleget tesz. Ezen projektor-matrixokra nézve ki fogjuk mutatni, hogy a (10) (11) alakban való előállítás nemcsak lehetséges, hanem egyedüli lehetőség $\lambda_1 = \lambda_2 = \ldots = \lambda_r = 1$ skalár együtthatókkal.

X. Tétel: *Ha egy r-edrangú* $\mathbf{P}$ *projektor-matrix elő van állítva lineárisan független diádok összegeként:*

$$\mathbf{P} = \sum_{k=1}^{r} \mathbf{u}_k \mathbf{v}_k^* \qquad (10_1)$$

alakban, akkor az itt szereplő $\mathbf{u}_k$, $\mathbf{v}_k^*$ *vektorok a projektor-matrixnak jobb- és baloldali sajátvektorai, melyek automatikusan biortogonalizálva vannak, azaz kielégítik a*

$$\mathbf{u}_k^* \mathbf{v}_l = \delta_{kl} \qquad (11)$$

relációkat.

Ez a tétel nyilván azt az állítást is magában foglalja, hogy egy r-edrangú projektor-matrixnak r számú saját értéke 1-gyel egyenlő[10], mert (10_1) és (11)-ből rögtön következik

$$\mathbf{P}\mathbf{u}_k = \mathbf{u}_k \quad \text{és} \quad \mathbf{v}_k^* \mathbf{P} = \mathbf{v}_k^* \quad (k = 1, 2, \ldots r).$$

A tétel bizonyításánál célszerű a diádikus előállításnak (5) alakját felhasználni. Ekkor a $\mathbf{P}^2 - \mathbf{P} = 0$ egyenlet a következő alakban írható:

$$[\mathbf{u}_1\, \mathbf{u}_2 \ldots \mathbf{u}_r] \begin{bmatrix} \mathbf{v}_1^* \\ \mathbf{v}_2^* \\ \vdots \\ \mathbf{v}_r^* \end{bmatrix} [\mathbf{u}_1\, \mathbf{u}_2 \ldots \mathbf{u}_r] \begin{bmatrix} \mathbf{v}_1^* \\ \mathbf{v}_2^* \\ \vdots \\ \mathbf{v}_r^* \end{bmatrix} - [\mathbf{u}_1\, \mathbf{u}_2 \ldots \mathbf{u}_r] \begin{bmatrix} \mathbf{v}_1^* \\ \mathbf{v}_2^* \\ \vdots \\ \mathbf{v}_r^* \end{bmatrix} = 0,$$

azaz

$$\begin{bmatrix} u_{11}\, u_{12} \ldots u_{1r} \\ u_{21}\, u_{22} \ldots u_{2r} \\ \cdot\ \cdot\ \cdot\ \cdot \\ u_{n1}\, u_{n2} \ldots u_{nr} \end{bmatrix} \left\{ \begin{bmatrix} \mathbf{v}_1^* \\ \mathbf{v}_2^* \\ \vdots \\ \mathbf{v}_r^* \end{bmatrix} [\mathbf{u}_1\, \mathbf{u}_2 \ldots \mathbf{u}_r] - \mathbf{E}_r^r \right\} \begin{bmatrix} v_{11}\, v_{12} \ldots v_{1n} \\ v_{21}\, v_{22} \ldots v_{2n} \\ \cdot\ \cdot\ \cdot\ \cdot \\ v_{r1}\, v_{r2} \ldots v_{rn} \end{bmatrix} = \mathbf{U}[\mathbf{v}_k^* \mathbf{u}_l - \delta_{kl}]\mathbf{V} = 0.$$

Az $\mathbf{u}_k$ és a $\mathbf{v}_k^*$ vektorok lineáris függetlenségéből következik, hogy az $\mathbf{U}$ és $\mathbf{V}$ matrixok r-edrangúak, tehát az **5. §.** VII. tétel alapján a fenti három-tényezős szorzat középső matrixtényezője eltűnik, azaz

$$[\mathbf{v}_k^* \mathbf{u}_l - \delta_{kl}] = 0,$$

q. e. d.

A most kimutatott tételnek elméleti és gyakorlati jelentőséget ad az a felismerés, hogy a projektor-matrixot nem is lehet másképpen diádok összegére bontani, mint úgy, hogy a diádok vektor-tényezői biortogonálisak.

Pl. a 429. oldalon bemutatott felbontásban szereplő

$$\mathbf{P} = \begin{vmatrix} -26 & -18 & -27 \\ 21 & 15 & 21 \\ 12 & 8 & 13 \end{vmatrix}$$

matrix projektor, mert kielégíti a $\mathbf{P}^2 = \mathbf{P}$ egyenletet. Ennek megfelelően az ottani felbontásnál nyert

$$\mathbf{u}_1 = \begin{bmatrix} -9 \\ 15/2 \\ 4 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 1 \\ -3/2 \\ 0 \end{bmatrix}, \quad \mathbf{v}_1^* = [3\ 2\ 13/4], \quad \mathbf{v}_2^* = [1\ 0\ 9/4]$$

vektorok valóban biortogonális rendszert alkotnak.

A projektor-matrixok definíciójából következik továbbá, hogy annak minden oszlop- ill. sor-vektora jobb- ill. baloldali saját-vektor $\lambda = 1$ saját-érték mellett. Ha ugyanis a projektor oszlopvektorokkal kifejezett alakja:

$$\mathbf{P} = [\mathbf{a}_1\, \mathbf{a}_2 \ldots \mathbf{a}_n],$$

[10] A többi $n-r$ számú sajátérték 0-val egyenlő és ezekhez az $\mathbf{E}-\mathbf{P}$ diadikus felbontásában szereplő vektortényezők tartoznak, mint bal- és jobboldali sajátvektorok.

akkor $\mathbf{P}^2 = \mathbf{P}$-ből következik

$$\mathbf{P}.[\mathbf{a}_1\,\mathbf{a}_2\ldots\mathbf{a}_n] = [\mathbf{P}\,\mathbf{a}_1, \mathbf{P}\mathbf{a}_2, \ldots \mathbf{P}\mathbf{a}_n] = [\mathbf{a}_1\,\mathbf{a}_2\ldots\mathbf{a}_n],$$

tehát

$$\mathbf{P}\,\mathbf{a}_i = \mathbf{a}_i. \tag{14. 1}$$

Hasonlóképpen bizonyítható

$$\mathbf{a}^j.\,\mathbf{P} = \mathbf{a}^j. \tag{14. 2}$$

8. §. A későbbi alkalmazásokra való tekintettel foglalkoznunk kell még hermitikus projektoroknak hermitikus diádok összegére való felbontásával. Egy $\mathbf{P}$ projektor hermitikus, ha egyidejűleg kielégíti a $\mathbf{P} = \overline{\mathbf{P}}^*$ és a $\mathbf{P}^2 = \mathbf{P}$ egyenleteket. Egy $\mathbf{u}\mathbf{v}^*$ diád pedig hermitikus, ha $\mathbf{u} = \overline{\mathbf{v}}$.

Legyen $\mathbf{P} = [a_{kl}]$, ahol $a_{lk} = \bar{a}_{kl}$, ekkor $\mathbf{P}^2 = \left[\sum_{\nu=1}^{n} a_{k\nu}\,a_{\nu l}\right]$, tehát a fődiagonális elemek összege $\mathbf{P}$-ben $\sum_{k=1}^{n} a_{kk}$ és $\mathbf{P}^2$-ben $\sum_{k=1}^{n}\left(\sum_{\nu=1}^{n} a_{k\nu}\,a_{\nu k}\right)$. Ebből a $\mathbf{P} = \mathbf{P}^2$ reláció folytán a

$$\sum_{k=1}^{n} a_{kk} = \sum_{k=1}^{n}\left(\sum_{\nu=1}^{n} a_{k\nu}\,a_{\nu k}\right) = \sum_{k}\sum_{\nu} |a_{k\nu}|^2$$

egyenlet adódik, melyből $\mathbf{P} \neq 0$ esetén következik, hogy az a_{kk} fődiagonális elemek közt van legalább egy pozitív. Legyen ez pl. a_{11}, akkor a

$$\mathbf{u}_1 = \frac{1}{\sqrt{a_{11}}}\begin{vmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{vmatrix}, \quad \overline{\mathbf{u}}_1^* = \frac{1}{\sqrt{a_{11}}}[a_{11}\,a_{12}\ldots a_{1n}]$$

jelölésekkel az $\mathbf{u}_1\overline{\mathbf{u}}_1^*$ diád hermitikus, továbbá $\mathbf{P} - \mathbf{u}_1\overline{\mathbf{u}}_1^*$ szintén hermitikus projektor, mert $\overline{\mathbf{u}}_1^*\mathbf{u}_1 = \dfrac{\Sigma a_{1\nu}\,a_{\nu 1}}{a_{11}} = 1$ és így a fenti (14) relációk folytán

$$(\mathbf{P} - \mathbf{u}_1\overline{\mathbf{u}}_1^*)^2 = \mathbf{P}^2 - \mathbf{P}\mathbf{u}_1\mathbf{u}_1^* - \mathbf{u}_1\overline{\mathbf{u}}_1^*\mathbf{P} + \mathbf{u}_1\overline{\mathbf{u}}_1^*\mathbf{u}_1\overline{\mathbf{u}}_1^* = \mathbf{P} - \mathbf{u}_1\overline{\mathbf{u}}_1^*.$$

E szerint $\varrho(\mathbf{P}) > 1$ esetén az $\mathbf{A} - \mathbf{u}_1\mathbf{u}_1^*$ projektorból a fenti eljárással egy további $\mathbf{u}_2\overline{\mathbf{u}}_2^*$ hermitikus diád választható le. Ha $\varrho(\mathbf{P}) = r$, akkor r számú lépés után ilymódon a 0 projektorhoz jutunk, tehát

XI. Tétel: Bármely hermitikus projektor előállítható

$$\mathbf{P} = \mathbf{u}_1\overline{\mathbf{u}}_1^* + \mathbf{u}_2\overline{\mathbf{u}}_2^* + \ldots + \mathbf{u}_r\overline{\mathbf{u}}^* \tag{15}$$

alakban és az itt szereplő vektorok kielégítik az $\mathbf{u}_k^*\overline{\mathbf{u}}_l = \delta_{kl}$ relációkat, azaz uniter-ortogonális rendszert alkotnak.

Valós hermitikus, azaz valós szimmetrikus projektor esetén nyilván $\overline{\mathbf{u}}_k = \mathbf{u}_k$, tehát egy r-edrangú valós szimmetrikus projektor r számú ortogonális, normált diád összegére bontható a következő alakban:

$$\mathbf{P} = \mathbf{u}_1\mathbf{u}_1^* + \mathbf{u}_2\mathbf{u}_2^* + \ldots + \mathbf{u}_r\mathbf{u}_r^*,$$

ahol

$$\mathbf{u}_k^*\,\mathbf{u}_l = \delta_{kl}.$$

III. Fejezet

Karakterisztikus polinom. Minimál polinom. Karakterisztikus-matrix rangcsökkenése

1. §. A projektor-matrixok tulajdonságainak levezetése (II. fej. 7. §) kizárólag azon a tényen alapul, hogy bármely $\mathbf{P}$ projektor kielégíti a $\mathbf{P}^2-\mathbf{P}=0$ egyenletet. Közelfekvő ezekután a következő kérdés: Kielégít-e egy tetszőleges $\mathbf{A}$ matrix valamely

$$c_0\mathbf{A}^m+c_1\mathbf{A}^{m-1}+\ldots c_{m-1}\mathbf{A}+c_m\mathbf{E}=0$$

alakú algebrai egyenletet? Erre a kérdésre ad választ a Hamilton—Cayley-től származó

XII. Tétel: *Bármely n-edrendű quadratikus* $\mathbf{A}$ *matrix kielégíti a hozzátartozó*

$$D(\lambda)\equiv|\mathbf{E}\lambda-\mathbf{A}|\equiv\lambda^n-\sum a_{ii}\lambda^{n-1}+\sum\sum\begin{vmatrix}a_{ii}&a_{ij}\\a_{ji}&a_{jj}\end{vmatrix}\lambda^{n-2}-\ldots+|\mathbf{A}|=0 \quad (1)$$

karakterisztikus egyenletet, azaz $D(\mathbf{A})=0$.

A matrix-algebra ezen középponti tételének a bizonyitása legegyszerűbben arra a tényre alapítható, hogy az $\mathbf{E}\lambda-\mathbf{A}=0$ matrix-egyenlet a $\lambda=\mathbf{A}$ helyettesítéssel triviális módon kielégíthető. Ugyanis [I. Fej. 20] szerint λ-ban identikusan

$$(\mathbf{E}\lambda-\mathbf{A})\,\mathrm{adj.}\,(\mathbf{E}\lambda-\mathbf{A})\equiv\mathrm{adj.}\,(\mathbf{E}\lambda-\mathbf{A})\cdot(\mathbf{E}\lambda-\mathbf{A})\equiv\mathbf{E}|\mathbf{E}\lambda-\mathbf{A}|\equiv\mathbf{E}D(\lambda).$$

Ez a λ skalárra vonatkozó azonosság általában nem marad érvényben, ha λ helyébe egy tetszőleges matrixot helyettesítünk, mert az azonosság levezetésében λ-nak bármely tényezővel való felcserélhetősége implicite fel volt tételezve. Azonban $\mathbf{E}\lambda-\mathbf{A}$ és adj. $(\mathbf{E}\lambda-\mathbf{A})$ felcserélhetősége következtében nyilván $\mathbf{A}$ az adj. $(\mathbf{E}\lambda-\mathbf{A})=\mathbf{A}_0\lambda^{n-1}+\mathbf{A}_1\lambda^{n-2}+\cdots+\mathbf{A}_{n-1}$ polinom minden egyes matrix-együtthatójával felcserélhető, ezért a (2) azonosság a $\lambda=\mathbf{A}$ helyettesítésnél érvényben marad és így adódik

$$\mathbf{E}.D(\mathbf{A})=(\mathbf{E}\mathbf{A}-\mathbf{A})\,\mathrm{adj.}\,(\mathbf{E}\mathbf{A}-\mathbf{A})=0,$$

tehát $\mathbf{E}\neq 0$ következtében $D(\mathbf{A})=0$, q. e. d.

2. §. Miként (1)-ből látható a $|\mathbf{E}\lambda-\mathbf{A}|=0$ karakterisztikus egyenlet fokszáma megegyezik az $\mathbf{A}$ matrix rendszámával, λ^k együtthatója pedig (I. I. (16)) az $|\mathbf{A}|$ determináns $n-k$-adrendű főminorainak összegével egyenlő.

Pl. ha a $\mathbf{P}=\mathbf{u}\mathbf{v}^*=[u_iv_j]$ diád tényezői kielégítik az $\mathbf{u}^*\mathbf{v}=\mathbf{v}^*\mathbf{u}=\sum u_kv_k=1$ relációt, akkor a hozzátartozó karakterisztikus egyenlet:

$$|\mathbf{E}\lambda-\mathbf{P}|\equiv\begin{vmatrix}\lambda-u_1v_1&-u_1v_2\ldots&-u_1v_n\\-u_2v_1&\lambda-u_2v_2\ldots&-u_2v_n\\\ldots&\ldots&\ldots\\-u_nv_1&-u_nv_2\ldots&\lambda-u_nv_n\end{vmatrix}\equiv\lambda^n-\lambda^{n-1}\sum_{k=1}^n u_kv_k\equiv\lambda^n-\lambda^{n-1}=0$$

($\varrho(\mathbf{P})=1$, következésképpen $\mathbf{P}$ másod- és magasabbrendű minorai eltűnnek).

Tehát a Cayley—Hamilton-tétel szerint **P** kielégíti a $\mathbf{P}^n - \mathbf{P}^{n-1} = 0$ egyenletet.

Másrészt könnyen felismerhető, hogy **P** projektor, ugyanis a $\mathbf{v}^*\mathbf{u} = 1$ feltétel figyelembevétele mellett $\mathbf{P}^2 = \mathbf{u}\mathbf{v}^*\mathbf{u}\mathbf{v}^* = \mathbf{u}\mathbf{v}^* = \mathbf{P}$, tehát **P** kielégíti a projektorokat jellemző $\mathbf{P}^2 - \mathbf{P} = 0$ egyenletet, melynek fokszáma $n > 2$ esetén kisebb, mint a fent levezetett karakterisztikus egyenleté.

Ez a példa mutatja, hogy vannak matrixok, melyek a rendszámuknál alacsonyabb fokú egyenletet is kielégítenek és ezzel a következő probléma felvetésére vezet.

Meghatározandó adott **A** matrixnak a „minimál-egyenlet"-e (redukált karakterisztikus egyenlete), vagyis az a legalacsonyabb fokú algebrai egyenlet, melyet az **A** matrix kielégít.

A megoldás kiinduló pontja ez esetben is a

$$(\mathbf{E}\lambda - \mathbf{A}).\mathrm{adj}.(\mathbf{E}\lambda - \mathbf{A}) \equiv \mathbf{E}.D(\lambda) \tag{2}$$

identitás.

Vizsgáljuk, lehet-e ezen identitás két oldalán álló matrixokat λ-nak valamely skalár polinomjával osztani. Miután egy matrixnak skalárral való osztása ekvivalens a matrix valamennyi elemének az osztásával, nyilvánvaló, hogy a baloldali első tényező: $\mathbf{E}\lambda - \mathbf{A} = [\delta_{kl}\lambda - a_{kl}]$ λ-nak semilyen (valódi) polinomjával nem osztható, tehát oszthatóság szempontjából a baloldalon csupán az $\mathrm{adj}.(\mathbf{E}\lambda - \mathbf{A}) = [D_{kl}(\lambda)]$ tényező jön tekintetbe. Ezen matrix $D_{kl}(\lambda)$ elemei λ-nak legfeljebb $n-1$-edfokú polinomjai, melyeknek esetleg van nem-konstans legnagyobb közös osztójuk, jelöljük ezt $\theta(\lambda)$-val. A jobboldalon álló $D(\lambda)$ determináns lineáris formája a saját $n-1$-edrendű minorainak, vagyis a $D_{kl}(\lambda)$ elemeknek, tehát nyilván osztható ezek legnagyobb közös osztójával: $\theta(\lambda)$-val. Minthogy továbbá $\theta(\lambda) \not\equiv 0$, tehát a (2) identitás mindkét oldalát oszthatjuk $\theta(\lambda)$-val:

$$(\mathbf{E}\lambda - \mathbf{A})\frac{\mathrm{adj}.(\mathbf{E}\lambda - \mathbf{A})}{\theta(\lambda)} = \mathbf{E}\frac{D(\lambda)}{\theta(\lambda)} \equiv \mathbf{E}\Delta(\lambda). \tag{3}$$

Ha most az így nyert identitásban a fenti meggondolások alapján λ helyébe az **A** matrixot helyettesítjük, azt nyerjük, hogy $\Delta(\mathbf{A}) = 0$. Kimutatható, hogy a $\Delta(\lambda) = 0$ egyenlet, melynek baloldala $D(\lambda)$-nak és a $D_{kl}(\lambda)$ minorok legnagyobb közös osztójának hányadosa, a legalacsonyabb fokú algebrai egyenlet, melyet az **A** matrix kielégít.[11]

3. §. A fokszám redukciója, vagyis $D(\lambda)$ és $\Delta(\lambda)$ fokszámainak különbségére nézve felső korlát állapítható meg a következő módon. Legyen $D(\lambda)$ prímtényezős alakja:

$$D(\lambda) = (\lambda - \lambda_1)^{\alpha_1}(\lambda - \lambda_2)^{\alpha_2}\dots(\lambda - \lambda_s)^{\alpha_s}; \; \alpha_1 + \alpha_2 + \dots + \alpha_s = n \tag{1$_1$}$$

Ekkor $\theta(\lambda)$, mint $D(\lambda)$ osztója szükségképpen a következő alakkal bír

$$\theta(\lambda) = (\lambda - \lambda_1)^{\beta_1}(\lambda - \lambda_2)^{\beta_2}\dots(\lambda - \lambda_s)^{\beta_s}; \; \beta_k \leq \alpha_k$$

[11] L. pl. C. C. Mac Duffee l. c.[6] pp. 20—21.

Kimutatjuk, hogy $\beta_n \leqq \alpha_k - 1$. Ugyanis $\theta(\lambda)$, mint a $D_{kl}(\lambda)$ minorok legnagyobb közös osztója, egyszersmind ezen minorok bármely homogén lineár formájának, így a $D'(\lambda) = \sum_k D_{kk}(\lambda)$ derivált polinomnak is osztója, következésképpen a $\lambda - \lambda_k$ prímtényezőt nem tartalmazhatja $\alpha_k - 1$-nél magasabb hatványon.

Ebből a megállapításból először is következik, hogy $\Delta(\lambda) = D(\lambda)$, ha a $D(\lambda) = 0$ karakterisztikus polinom csupa egyszeres gyökkel bír.

Továbbá, többszörös gyökök esetén a maximális fokszámredukció akkor áll be, ha (konstans faktortól eltekintve)

$$\theta(\lambda) = (\lambda - \lambda_1)^{\alpha_1 - 1}(\lambda - \lambda_2)^{\alpha_2 - 1} \dots (\lambda - \lambda_s)^{\alpha_s - 1} = (D(\lambda), D'(\lambda)).$$

Ebben az esetben

$$\Delta(\lambda) = \frac{D(\lambda)}{\theta(\lambda)} = (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_s),$$

tehát a $\Delta(\lambda) = 0$ minimál-egyenlet a $D(\lambda) = 0$ karakterisztikus egyenlet valamennyi gyökét egyszeres multiplicitással tartalmazza.

Mint látni fogjuk, azok a matrixok, melyeknek minimál-egyenlete csupa egyszeres gyökkel bír, a matrixoknak egy fontos alosztályát alkotják, mely az alkalmazásokban előforduló matrixoknak túlnyomó többségét felöleli és amelyeknek az elmélete elemi eszközökkel felépíthető.

4. §. Ezekután önként felmerül a kérdés, vannak-e olyan matrixok, melyeknek külső struktúrájából (vagyis a $\theta(\lambda)$ legnagyobb közös osztó direkt kiszámítása nélkül) megállapítható, hogy minimál-egyenletük csupa egyszeres gyökkel bír.

Ki fogjuk mutatni, hogy egy hermitikus (és ennek speciális eseteként egy valós szimmetrikus) matrix mindig ilyen tulajdonságú.

A bizonyításhoz néhány segédtételt bocsátunk előre.

Valamely hermitikus matrix karakterisztikus egyenletének nincs tisztán képzetes gyöke.

Tegyük fel ezzel ellentétben, hogy egy hermitikus, tehát az $\mathbf{A} = \overline{\mathbf{A}}^*$ relációt kielégítő matrixnál $|\mathbf{E}i\mu - \mathbf{A}| = 0$, ahol μ 0-tól különböző valós szám. Akkor fennáll a

$$|\mathbf{E}i\mu - \mathbf{A}|\,|\mathbf{E}i\mu + \mathbf{A}| = -|\mathbf{E}\mu^2 + \mathbf{A}^2| = -|\mathbf{E}\mu^2 + \mathbf{A}\overline{\mathbf{A}}^*| = 0$$

egyenlet, illetve részletesen kiírva, a

$$0 = \mu^{2n} + \mu^{2n-2} \sum_k \sum_l a_{kl}\bar{a}_{kl} + \dots + \begin{vmatrix} a_{11} \dots a_{1n} \\ \dots\dots \\ a_{n1} \dots a_{nn} \end{vmatrix} \begin{vmatrix} \bar{a}_{11} \dots \bar{a}_{1n} \\ \dots\dots \\ \bar{a}_{n1} \dots \bar{a}_{nn} \end{vmatrix}$$

egyenlet is, ami 0-tól különböző valós μ esetén nyilván ellenmondás. Ezzel tételünk be van bizonyítva.

Valamely hermitikus matrix karakterisztikus egyenletének összes gyökei valósak.

Tegyük fel ezzel ellentétben, hogy az $\mathbf{A} = \overline{\mathbf{A}}^*$ hermitikus matrix karakterisztikus egyenletét kielégíti a $\lambda = \varkappa + i\mu$ $(\mu \neq 0)$ komplex szám, azaz fennáll a $|\mathbf{E}(\varkappa + i\mu) - \mathbf{A}| = 0$ egyenlet. Ekkor azonban nyilván $\mathbf{B} = \mathbf{A} - \varkappa\mathbf{E}$ matrix is hermitikus és erre nézve fennállna az $|\mathbf{E}i\mu - \mathbf{B}| = 0$ egyenlet, ami ellenmondásban van az előbb bizonyított tétellel, az $|\mathbf{E}\lambda - \mathbf{A}| = 0$ egyenletnek tehát nincsenek komplex gyökei. Q. e. d.

XIII. Tétel: *Ha* $\mathbf{A}$ *hermitikus, akkor* $|\mathbf{E}\lambda - \mathbf{A}| \equiv D(\lambda)$ *összes* $D_{kl}(\lambda)$ *minorai oszthatók* $(D(\lambda), D'(\lambda))$*-val.*

Ismeretes, hogy[12]

$$\begin{vmatrix} D_{kk}(\lambda), & D_{kl}(\lambda) \\ D_{lk}(\lambda), & D_{ll}(\lambda) \end{vmatrix} = D(\lambda) D_{\substack{kl \\ kl}}(\lambda),$$

eszerint

$$\sum_{k=1}^{n} \sum_{l=1}^{n} D_{kl}(\lambda) D_{lk}(\lambda) = \sum_{k=1}^{n} D_{kk}(\lambda) \sum_{l=1}^{n} D_{ll}(\lambda) - D(\lambda) \sum_{k=1}^{n} \sum_{l=1}^{n} D_{\substack{kl \\ kl}}(\lambda).$$

Azonban $\Sigma D_{kk}(\lambda) = D'(\lambda)$, $\Sigma\Sigma D_{\substack{kl \\ kl}}(\lambda) = D''(\lambda)$ és λ *valós értékei mellett* $\overline{D_{lk}(\lambda)} = D_{kl}(\lambda)$, tehát valós λ-ra fennáll a következő identitás:

$$\sum_{k=1}^{n} \sum_{l=1}^{n} |D_{kl}(\lambda)|^2 \equiv D'(\lambda)^2 - D(\lambda) D''(\lambda).$$

Az előzők szerint $D(\lambda) = \prod_{k=1}^{s} (\lambda - \lambda_k)^{\alpha_k}$ összes 0-helyei valósak, ennélfogva a jobboldal és vele együtt a baloldal is osztható a $\lambda - \lambda_k$ *valós* prímtényező $2(\alpha_k - 1)$-edik hatványával. Ez azonban csak úgy lehetséges, hogy a baloldali négyzetösszeg *minden tagja* osztható $(\lambda - \lambda_k)^{2(\alpha_k - 1)}$-nel. Minthogy pedig az oszthatóság k minden értékére fennáll, tehát mindegyik $D_{kl}(\lambda)$ minor osztható $\prod_{n=1}^{s} (\lambda - \lambda_k)^{\alpha_k - 1}$-gyel, azaz $(D(\lambda), D'(\lambda))$-val és így a fentiek figyelembevételével ki van mutatva, hogy a $D_{kl}(\lambda)$ minorok legnagyobb közös osztója: $\theta(\lambda) = (D(\lambda), D'(\lambda))$-val. Érvényes tehát a következő

XIV. Tétel: *Hermitikus (valós szimmetrikus) matrix minimál-egyenlete csupa egyszeres, valós gyökkel bír.*

[12] L. pl. E. Pascal l. c.[9] p. 61.

IV. Fejezet

Matrix analitikus függvénye és annak redukciója minimális fokszámú matrix-polinomra. Matrix-függvény kanonikus előállítása egyszeres gyökökkel bíró minimálegyenlet esetén

1. §. Legyen $f(z)=\sum_{\nu=0}^{\infty} c_\nu z^\nu$ a z skalárváltozónak közönséges hatványsora. Ha valamely n-edrendű $\mathbf{A}$ matrix behelyettesítésénél a

$$S_N(\mathbf{A})=\sum_{\nu=0}^{N} c_\nu \mathbf{A}^\nu$$

matrix minden eleme $N\rightarrow\infty$ mellett véges limeshez tart, akkor azt mondjuk, hogy a

$$f(\mathbf{A})=\sum_{\nu=0}^{\infty} c_\nu \mathbf{A}^\nu \tag{1}$$

matrix-hatványsor konvergens és az $\mathbf{A}$ matrixnak $f(\mathbf{A})$ analitikus függvényét definiálja.

A (1) matrix-hatványsor konvergenciájára nézve érvényes a

XV. Tétel: *Ha az $\mathbf{A}$ matrix karakterisztikus gyökei az $f(z)=\sum c_\nu z^\nu$ hatványsor konvergencia-körének belsejébe esnek, akkor a $\sum c_\nu \mathbf{A}^\nu$ matrix-hatványsor konvergens és a minimálpolinomnál alacsonyabb fokszámú matrix-polinomra redukálható.*

Legyen az $\mathbf{A}$ matrix minimálpolinomja

$$\Delta(z)=(z-\lambda_1)^{\gamma_1}(z-\lambda_2)^{\gamma_2}\cdots(z-\lambda_s)^{\gamma_s};\quad \gamma_1+\gamma_2+\ldots+\gamma_s=m\leqq n \tag{2}$$

és képezzük az $S_N(z)$ polinomnak a $\Delta(z)$ osztóra vonatkozó (legfeljebb $m-1$-ed fokú) $R_N(z)$ maradékát, melyet a

$$S_N(z)=\Delta(z)\cdot Q_N(z)+R_N(z) \tag{3}$$

azonosság definiál. Ebből $\Delta(\mathbf{A})=0$-ra való tekintettel következik

$$S_N(\mathbf{A})=R_N(\mathbf{A}). \tag{3. 1}$$

Az $R_N(z)$ maradékpolinom-sorozat effektív kiszámítása osztás útján kivihetetlennek minősítendő úgy formális, mint még inkább numerikus számítás szempontjából. Éppen ezért — nem csak e helyen, de az egész matrix-redukálási elmélet szempontjából — alapvető az a felismerés, hogy a $R_N(z)$ osztási maradék explicite kifejezhető a Hermite-f. interpolációs polinomok segélyével.

(3) szerint ugyanis $S_N(z)-R_N(z)$ osztható $\Delta(z)=\Pi(z-\lambda_k)^{\gamma_k}$-val, ennek következtében

$$R_N(\lambda_k)=S_N(\lambda_k),\ R'_N(\lambda_k)=S'_N(\lambda_k),\ \ldots,\ R_N^{(\gamma_k-1)}(\lambda_k)=S_N^{(\gamma_k-1)}(\lambda_k)\ (k=1,2,\ldots,s). \tag{4}$$

Minthogy pedig $R_N(z)$ fokszáma legfeljebb $m-1=\sum\gamma_k-1$, tehát $R_N(z)$ a

$\sum \gamma_k$ számú (4) feltétellel egyértelműen meg van határozva. Bevezetve a[13]

$$H_{k\varkappa}^{(\nu)}(\lambda_\mu) = \delta_{k\mu}\delta_{\varkappa\nu} \qquad \begin{matrix} \mu, k = 1, 2, \ldots, s \\ \varkappa, \nu = 0, 1, \ldots, \gamma_k - 1 \end{matrix} \tag{5}$$

relációkkal definiált, legfeljebb $m-1$-ed fokú Hermite-f. $H_{k\varkappa}(\lambda)$ alappolinomokat, az $R_N(z)$ osztási maradéknak következő előállításához jutunk:

$$R_N(z) = \sum_{k=1}^{s} \{S_N(\lambda_k)H_{k0}(z) + S_N'(\lambda_k)H_{k1}(z) + \ldots + S_N^{(\gamma_k-1)}(\lambda_k)H_{k,\gamma_k-1}(z)\}.$$

Ebből (3_1) figyelembe vételével

$$S_N(\mathbf{A}) = \sum_{k=1}^{s} \{S_N(\lambda_k)H_{k0}(\mathbf{A}) + S_N'(\lambda_k)H_{k1}(\mathbf{A}) + \ldots + S_N^{(\gamma_k-1)}(\lambda_k)H_{k,\gamma_k-1}(\mathbf{A})\}. \tag{6}$$

Minthogy a jobboldalon álló matrixok N-től függetlenek, az $N \to \infty$ határátmenet nehézség nélkül elvégezhető és pedig tekintettel arra, hogy feltevés szerint $\mathbf{A}$ karakterisztikus gyökei: $\lambda_1, \lambda_2, \ldots, \lambda_s$ $\sum c_\nu z^\nu$ konvergenciakörének belsejébe esnek, $\lim\limits_{N\to\infty} S_N(\lambda_k) = f(\lambda_k)$, $\lim\limits_{N\to\infty} S_N'(\lambda_k) = f'(\lambda_k), \ldots$, tehát

$$f(\mathbf{A}) = \lim_{N\to\infty} S_N(\mathbf{A}) = \sum_{k=1}^{s} \{f(\lambda_k)H_{k0}(\mathbf{A}) + \ldots + f^{(\gamma_k-1)}(\lambda_k)H_{k,\gamma_k-1}(\mathbf{A})\}. \tag{7}$$

Ezzel kimutattuk, hogy az előírt feltételek mellett az $f(\mathbf{A})$ matrix-függvény értelmezve van és kifejeztük azt a $H_{k\varkappa}(\mathbf{A})$ Hermite-f. matrix-polinomok lineáris formájaként.

2. §. Az $f(\mathbf{A})$ matrix-függvény fenti (7) kifejezése lényegesen egyszerűsödik és további redukcióra nyújt lehetőséget akkor, ha a $\Delta(z)$ minimálpolinom csupa egyszeres gyökkel bír, azaz

$$\Delta(z) = (z-\lambda_1)(z-\lambda_2)\ldots(z-\lambda_s) \qquad s \leqq n \tag{2. 1}$$

alakú. Ekkor ugyanis csupán a

$$L_k(\lambda_\varkappa) = \delta_{k\varkappa} \qquad (k, \varkappa = 1, 2, \ldots, s) \tag{5. 1}$$

relációkkal definiált és a

$$L_k(z) \equiv \frac{\Delta(z)}{\Delta'(\lambda_k)(z-\lambda_k)} \tag{5. 2}$$

explicit formulákkal előállított, legfeljebb $s-1$-ed fokú Lagrange-f. alappolinomok alkalmazása válik szükségessé és a (7) előállításnak jelen esetre való alkalmazásával a következő eredményre jutunk:

XVI. Tétel: *Ha az $\mathbf{A}$ matrix minimálegyenletének valamennyi $\lambda_1, \lambda_2, \ldots, \lambda_s$ gyöke egyszeres, ha továbbá az $f(z) = \sum c_\nu z^\nu$ hatványsor konvergenciaköre ezeket a gyököket a belsejében tartalmazza, akkor az $f(\mathbf{A})$ matrix-függvény előállítható Lagrange-f. matrix-polinómok lineáris formájaként a következő alakban:*

$$f(\mathbf{A}) = f(\lambda_1)L_1(\mathbf{A}) + f(\lambda_2)L_2(\mathbf{A}) + \ldots + f(\lambda_s)L_s(\mathbf{A}). \tag{8}$$

[13] A Hermite f. alappolinomokra nézve l. pl. Szász Pál, A differenciál és integrálszámítás elemei. Budapest, 1951, II. köt. 22—29. old.

3. §. Állapítsuk meg az ebben az előállításban szereplő $L_k(\mathbf{A})$ matrixok $\varrho(L_k)$ rangszámát. A Lagrange-polinómok közt fennálló $\sum L_k(z) \equiv 1$ identitásból következik

$$\sum_{k=1}^{s} L_k(\mathbf{A}) = \mathbf{E},$$

tehát II. *Fejezet* IV. Tétel szerint

$$n = \varrho(\mathbf{E}) = \varrho\left(\sum_{k=1}^{s} L_k(\mathbf{A})\right) \leqq \sum_{k=1}^{s} \varrho(L_k(\mathbf{A})). \tag{9}$$

Másrészt (5_2) szerint

$$(z-\lambda_k)L_k(z) = \Delta'(\lambda_k)^{-1}\cdot\Delta(z),$$

tehát

$$(\mathbf{A}-\lambda_k\mathbf{E})L_k(\mathbf{A}) = 0.$$

Ha továbbá (I. III. (1_1)) $D(\lambda) = \prod_{k=1}^{s}(\lambda-\lambda_k)^{\alpha_k}$ akkor $D^{(\alpha_k)}(\lambda_k) \neq 0$. De $D^{(\alpha_k)}(\lambda)$ egyenlő $|\lambda\mathbf{E}-\mathbf{A}|$ $n-\alpha_k$-ad rendű főminorainak összegével, ezek tehát nem tünhetnek el mind a $\lambda = \lambda_k$ helyen, vagyis $\varrho(\lambda_k\mathbf{E}-\mathbf{A}) \geqq n-\alpha_k$. Innen II. Fejezet VII. Tétel felhasználásával következik, hogy

$$\varrho(L_k(\mathbf{A})) \leqq \alpha_k\,; \quad \sum_{k=1}^{s} \varrho(L_k(\mathbf{A})) \leqq \sum_{k=1}^{s} \alpha_k = n \tag{10}$$

A (9) és (10) egyenlőtlenségek azonban csak akkor állhatnak fenn egyidejűleg, ha $\varrho(L_k(\mathbf{A})) = \alpha_k$. Ezzel kimutattuk, hogy a (8) előállításban szereplő $L_k(\mathbf{A})$ Lagrange-f. matrix-polinom rangja egyenlő a λ_k karakterisztikus gyök multiplicitásával:

$$\varrho(L_k(\mathbf{A})) = \alpha_k.$$

4. §. Az $L_k(z)$ Lagrange-f. alappolinomok bizonyos oszthatósági tulajdonságokkal bírnak, melyekből az $f(\mathbf{A})$ matrix-függvény további redukciója szempontjából alapvető következtetéseket fogunk levonni.

1. $L_k(z)$ (5) szerint eltűnik a $\lambda_1 \ldots, \lambda_{k-1}, \lambda_{k+1} \ldots \lambda_s$ helyeken, és $1-L_k(z)$ eltűnik a λ_k helyen. Eszerint $L_k(z)\,(L_k(z)-1)$ eltűnik $\Delta(z)$ valamennyi 0-helyén, tehát

$$L_k(z)(L_k(z)-1) \equiv L_k(z)^2 - L_k(z) \equiv \Delta(z)\cdot Q(z)$$

alakban írható. Innen $z = \mathbf{A}$ helyettesítéssel

$$L_k(\mathbf{A})^2 - L_k(\mathbf{A}) = 0 \tag{11}$$

Ezzel bebizonyítottuk a következő

XVII. Tételt: *Ha az* **A** *matrix karakterisztikus egyenletének különböző gyökei:* $\lambda_1, \lambda_2, \ldots, \lambda_s$, *akkor az ezeken a helyeken interpoláló Lagrange-f. matrix-polinomok mindegyike projektor.*

Fentebb kimutattuk, hogy $\varrho(L_k(\mathbf{A}))=\alpha_k$. Ezt figyelembe véve, a (8) tétel szerint $L_k(\mathbf{A})$ felbontható α_k számú diád összegére:

$$L_k(\mathbf{A})=\mathbf{u}_{k1}\mathbf{v}_{k1}^*+\mathbf{u}_{k2}\mathbf{v}_{k2}^*+\ldots+\mathbf{u}_{k\alpha_k}\mathbf{v}_{k\alpha_k}^*. \tag{12}$$

Minthogy azonban a most bizonyított tétel szerint $L_k(\mathbf{A})$ projektor, azért a II. Fejezet X. Tétel értelmében $L_k(\mathbf{A})$ bármely (12) alakú felbontásában a diádok automatikusan biortogonalizálva vannak, azaz

$$\mathbf{u}_{ki}^*\mathbf{v}_{kj}=\delta_{ij} \qquad (i,j=1,2,\ldots,\alpha_k) \tag{12'}$$

Ennek következtében az $\mathbf{u}_{ki}$ ill. $\mathbf{v}_{kj}^*$ tényezők egyszersmind az $L_k(\mathbf{A})$ projektor jobb- ill. baloldali saját vektorai.

2. (5_2)-ből nyilvánvaló, hogy $k\neq h$ esetén $L_k(z)L_h(z)$ eltűnik $\Delta(z)$ valamennyi 0-helyén, tehát

$$L_k(z)\cdot L_h(z)\equiv\Delta(z)\cdot Q(z) \qquad k\neq h$$

alakban írható. Innen $z=\mathbf{A}$ helyettesítéssel

$$L_k(\mathbf{A})L_h(\mathbf{A})=L_h(\mathbf{A})L_k(\mathbf{A})=0. \tag{13}$$

Eszerint, ha az α_k-ad rendű $L_k(\mathbf{A})$, ill. α_h-ad rangú $L_h(\mathbf{A})$ matrixok diadikus előállításai

$$L_k(\mathbf{A})=\mathbf{u}_{k1}\mathbf{v}_{k1}^*+\ldots+\mathbf{u}_{k\alpha_k}\mathbf{v}_{k\alpha_k}^*$$

és (14)

$$L_h(\mathbf{A})=\mathbf{u}_{h1}\mathbf{v}_{h1}^*+\ldots+\mathbf{u}_{h\alpha_h}\mathbf{v}_{h\alpha_h},$$

akkor $L_k(\mathbf{A})L_h(\mathbf{A})=[\mathbf{u}_{k1}\ldots\mathbf{u}_{k\alpha_k}]\left\{\begin{bmatrix}\mathbf{v}_{k1}^*\\ \vdots\\ \mathbf{v}_{k\alpha_k}^*\end{bmatrix}[\mathbf{u}_{h1}\ldots\mathbf{u}_{h\alpha_h}]\right\}\begin{bmatrix}\mathbf{v}_{h1}^*\\ \vdots\\ \mathbf{v}_{h\alpha_h}^*\end{bmatrix}=0,$

tehát a II. Fejezet VII. tétel értelmében

$$\mathbf{v}_{ki}^*\mathbf{u}_{hj}=0 \qquad k\neq h, \quad \begin{matrix} i=1,2,\ldots,\alpha_k\\ j=1,2,\ldots,\alpha_h\end{matrix}. \tag{14'}$$

Ezzel kimutattuk, hogy bármelyik $L_k(\mathbf{A})$ projektornak valamennyi bal (jobb) oldali saját vektora ortogonális az összes többi $L_h(\mathbf{A})$ projektoroknak valamennyi jobb (bal) oldali saját vektorára.

5. §. Az (12') és (14') relációk egyidejű figyelembevételéből közvetlenül következik a

(XVIII). Matrixfüggvények kanonikus előállításának alaptétele. *Ha az* $\mathbf{A}$ *matrix karakterisztikus polinomja:* $|\lambda\mathbf{E}-\mathbf{A}|=D(\lambda)=\prod_1^s(\lambda-\lambda_k)^{\alpha_k}$, *minimálpolinomja:* $\Delta(\lambda)=\prod_1^s(\lambda-\lambda_k)$, *ha továbbá az* $f(z)=\sum c_\nu z^\nu$ *hatványsor konvergenciaköre a* $\lambda_1,\lambda_2,\ldots,\lambda_s$ *0-helyeket a belsejében tartalmazza, akkor a* $f(\mathbf{A})$ *matrix-függvény a következő kanonikus alapban állítható elő:*

$$f(\mathbf{A})=\sum_{k=1}^s f(\lambda_k)L_k(\mathbf{A})=\sum_{k=1}^s f(\lambda_k)(\mathbf{u}_{k1}\mathbf{v}_{k1}^*+\cdots+\mathbf{u}_{k\alpha_k}\mathbf{v}_{k\alpha_k}^*), \tag{15}$$

ahol $L_k(z)=\dfrac{\Delta(z)}{\Delta'(\lambda_k)(z-\lambda_k)}$ *és a* $\mathbf{u}_{ki}$, $\mathbf{v}_{lj}$ *vektorok biortogonális rendszert alkotnak, azaz*

$$\mathbf{u}_{kl}^{*}\mathbf{v}_{lj}=\delta_{kl}\delta_{ij}\begin{pmatrix} k,l=1,2,\dots s \\ i=1,2,\dots \alpha_k \\ j=1,2,\dots \alpha_l \end{pmatrix}. \tag{15. 1}$$

Az $\mathbf{u}_{ki}$ *ill.* $\mathbf{v}_{kj}^{*}$ *vektoroknak* **A**-*nak* λ_k-*hoz tartozó jobb- ill. baloldali saját vektorai.*

Innen az $f(z)\equiv z$ esetben magának az **A** matrixnak

$$\mathbf{A}=\sum_{k=1}^{s}\lambda_k(\mathbf{u}_{k1}\mathbf{v}_{k1}^{*}+\dots+\mathbf{u}_{k\alpha_k}\mathbf{v}_{k\alpha_k}^{*}) \tag{16}$$

kanonikus felbontása' adódik.

Ha $\mathbf{A}=\overline{\mathbf{A}}^{*}$, azaz **A** hermitikus, akkor a minimálpolinomnak csupa egyszeres 0-helye van, tehát minden hermitikus matrix előállítható a

$$\mathbf{A}=\sum_{k=1}^{s}\lambda_k(\mathbf{u}_{k1}\overline{\mathbf{u}}_{k1}^{*}+\dots\ \mathbf{u}_{k\alpha_k}\overline{\mathbf{u}}_{k\alpha_k}^{*}) \tag{17}$$

kanonikus alakban.

Ha $\mathbf{A}=\overline{\mathbf{A}}$, és $\mathbf{A}=\mathbf{A}^{*}$, azaz **A** valós, szimmetrikus, akkor kanonikus előállítása:

$$\mathbf{A}=\sum_{k=1}^{s}\lambda_k(\mathbf{u}_{k1}\mathbf{u}_{k1}^{*}+\dots+\mathbf{u}_{k\alpha_k}\mathbf{u}_{k\alpha_k}^{*}). \tag{18}$$

Ha ezt az eredményt a $\Sigma\Sigma a_{ij}x_ix_j=\mathbf{x}^{*}\mathbf{A}\mathbf{x}$ quadratikus alakra alkalmazzuk, nyerjük annak ortogonális transzformációval négyzetösszegbe átvitt alakját:

$$\begin{aligned}\mathbf{x}^{*}\mathbf{A}\mathbf{x}&=\mathbf{x}^{*}\sum_{k=1}^{s}\lambda_k(\mathbf{u}_{k1}\mathbf{u}_{k1}^{*}+\dots+\mathbf{u}_{k\alpha_k}\mathbf{u}_{k\alpha_k}^{*})\mathbf{x}=\\&=\sum_{k=1}^{s}\lambda_k\{(\mathbf{u}_{k1}^{*}\mathbf{x})^2+\dots:(\mathbf{u}_{k\alpha_k}^{*}\mathbf{x})^2\}.\end{aligned} \tag{19}$$

6. §. Ha az adott **A** matrix rendszáma: n határozatlan, akkor az $L_k(\mathbf{A})$ Lagrange-f. matrix-polinomok effektív kiszámítására gyakran a direkt behelyettesítésnél előnyösebb a következő eljárás, mely az $L_k(\mathbf{A})$ és az adj. $(\lambda\mathbf{E}-\mathbf{A})$ matrixok közti összefüggésen alakul. Ezen összefüggés levezetésének kiinduló pontja a következő

Lemma. Ha $D(x)=\prod_{k=1}^{n}(x-\lambda_k)$ n-ed fokú polinom, melynek nincsenek többszörös 0-helyei, akkor érvényes a következő identitás

$$\frac{D(x)-D(y)}{x-y}=\sum_{k=1}^{n}D'(\lambda_k)L_k(x)L_k(y), \tag{20}$$

ahol

$$L_k(x)=\frac{D(x)}{D'(\lambda_k)(x-\lambda_k)}. \tag{5. 3}$$

Bizonyítás: $\dfrac{D(x)-D(y)}{x-y}$ az y változónak $n-1$-ed fokú polinomja, tehát y szerint a $\lambda_1, \lambda_2, \dots \lambda_n$ helyeken interpolálva

$$\frac{D(x)-D(y)}{x-y}=\sum_{k=1}^{n}\frac{D(x)}{x-\lambda_k}L_k(y).$$

Ha itt $\dfrac{D(x)}{x-\lambda_k}$ helyébe annak (5. 3)-ból adódó kifejezését helyettesítjük, éppen a bizonyítandó egyenlőséget nyerjük.

7. §. Legyenek $\lambda_1\lambda_2\dots\lambda_n$ az $\mathbf{A}$ matrix egyszeres sajátértékei és helyettesítjük az I. 4. §. befejező megjegyzés alapján a (20) identitásban x helyébe a $\lambda\mathbf{E}$ matrixot és y helyébe az $\mathbf{A}$ matrixot. Ekkor a $D(\mathbf{A})=0$ és a (III. (2)) egyenletek következtében

$$\text{adj.}(\lambda\mathbf{E}-\mathbf{A})=\frac{\mathbf{E}D(\lambda)}{\lambda\mathbf{E}-\mathbf{A}}\sum_{k=1}^{n}D'(\lambda_k)L_k(\lambda)L_k(\mathbf{A}).$$

Ebből $\lambda=\lambda_k$ helyettesítéssel $L_k(\mathbf{A})$ következő kifejezéséhez jutunk:

$$L_k(\mathbf{A})=\frac{1}{D'(\lambda_k)}\text{adj.}(\lambda_k\mathbf{E}-\mathbf{A}). \tag{21}$$

(Ha $\mathbf{A}$ karakterisztikus polinómja: $D(\lambda)=\Pi(\lambda-\lambda_k)^{\alpha_k}$ és minimálpolinomja: $\Delta(\lambda)=\Pi(\lambda-\lambda_k)$, akkor a fentiekhez hasonlóan kimutatható, hogy $L_k(\mathbf{A})$ és adj. $(\lambda\mathbf{E}-\mathbf{A})$ közt a következő összefüggés áll fenn:

$$L_k(\mathbf{A})=\frac{\alpha_k}{D^{(\alpha_k)}(\lambda_k)}\{\text{adj.}(\lambda\mathbf{E}-\mathbf{A})\}^{(\alpha_k-1)}_{\lambda=\lambda_k}. \tag{21'}$$

8. §. A matrixok kanonikus előállításáról szóló fenti alaptétel még a következőképpen is fogalmazható:

XIX. Tétel: *Minden olyan* $\mathbf{A}$ *matrix, melynek minimálpolinomja csupa egyszeres* 0*-hellyel bír, hasonló a sajátértékeiből alkotott diagonális matrixhoz:*

$$\mathbf{A}\sim\langle\lambda_{11}\dots\lambda_{1\alpha_1}\dots\lambda_{s1}\dots\lambda_{s\alpha_s}\rangle,\quad \lambda_{k1}=\lambda_{k2}=\dots=\lambda_{k\alpha_k}=\lambda_k;\quad k=1,2,\dots,s$$

azaz, létezik olyan $\mathbf{U}$ *matrix, melyre*

$$\mathbf{A}=\mathbf{U}\langle\lambda_{11}\dots\lambda_{1\alpha_1}\dots\lambda_{s1}\dots\lambda_{s\alpha_s}\rangle\mathbf{U}^{-1}.$$

Bizonyítás: Az egyszerűbb írásmód kedvéért jelöljük a fenti sorrendben vett (multiplicitásuknak megfelelően ismételt) saját értéket $\lambda_1, \lambda_2, \dots, \lambda_n$-nel és a hozzájuk tartozó jobb- és baloldali biortogonalizált saját vektorokat $\mathbf{u}_1\dots\mathbf{u}_n$, ill. $\mathbf{v}_1^*\dots\mathbf{v}_n^*$-gal. Ekkor a (16) kanonikus előállítás így írható:

$$\mathbf{A}=\sum_{k=1}^{n}\lambda_k\mathbf{u}_k\mathbf{v}_k^*=[\mathbf{u}_1\mathbf{u}_2\dots\mathbf{u}_n]\begin{bmatrix}\lambda_1 & 0 & \dots & 0\\ 0 & \lambda_2 & \dots & 0\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & & \lambda_n\end{bmatrix}\begin{bmatrix}\mathbf{v}_1^*\\ \mathbf{v}_2^*\\ \vdots\\ \mathbf{v}_n^*\end{bmatrix}. \tag{22}$$

Azonban az $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \ldots \mathbf{u}_n]$ és $\mathbf{V}^* = \begin{bmatrix} \mathbf{v}_1^* \\ \mathbf{v}_2^* \\ \vdots \\ \mathbf{v}_n^* \end{bmatrix}$ matrixok az $\mathbf{v}_i^* \mathbf{u}_j = \delta_{ij}$ biortogonálási relációk következtében egymásnak reciprokjai, azaz $\mathbf{V}^* = \mathbf{U}^{-1}$, tehát (22)-ből következik

$$\mathbf{A} = \mathbf{U}\langle\lambda_1 \ldots \lambda_n\rangle\mathbf{U}^{-1}.$$

Hermitikus, $\mathbf{A}$ matrix mindig hasonló egy diagonális matrixhoz és a (17) reláció ez esetben a

$$\mathbf{A} = \mathbf{U}\langle\lambda_1 \ldots \lambda_n\rangle\bar{\mathbf{U}}^*$$

alakot ölti. Valós, szimmetrikus matrix esetén (18) alapján hasonló meggondolással nyerjük, hogy

$$\mathbf{A} = \mathbf{U}\langle\lambda_1 \ldots \lambda_n\rangle\mathbf{U}^*.$$

V. Fejezet

Lineáris, állandó együtthatós differenciálegyenlet-rendszerek megoldása a matrix-számítás felhasználásával

1. §. Az $x = e^{At}$ függvény kielégíti az $\dot{x} = Ax$ differenciálegyenletet és a $t = 0$ helyen 1 értéket vesz fel. Következésképpen az

$$\dot{x} = Ax$$

elsőrendű homogén differenciálegyenletnek az

$$x(0) = x_0$$

kezdőfeltételt kielégítő $x(t; x_0)$ megoldását nyilván a

$$x(t; x_0) = e^{At}.x_0 \tag{1}$$

formula szolgáltatja.

Az $\dot{x} - Ax = e^{At}\frac{d}{dt}(e^{-At}.x)$ azonosságból pedig következik, hogy az

$$\dot{x} - Ax = g(t)$$

elsőrendű inhomogén differenciálegyenletnek az

$$x(0) = 0$$

kezdőfeltételt kielégítő megoldását a

$$x(t) = \int_0^t e^{A(t-\tau)}.g(\tau)\,d\tau \tag{2}$$

formula adja.

Az $x = \cos(\sqrt{A}.t)$, ill. a $x = \frac{\sin(\sqrt{A}t)}{\sqrt{A}}$ függvények kielégítik a $\ddot{x} + Ax = 0$ differenciálegyenletet, továbbá a $t = 0$ helyen a $x(0) = 1$, $\dot{x}(0) = 0$, ill. a $x(0)=0; \dot{x}(0)=1$ kezdőfeltételeket. Következésképpen az

$$\ddot{x}+Ax=0$$

másodrendű homogén differenciálegyenletnek a

$$x(0)=x_0; \quad \dot{x}(0)=\dot{x}_0$$

kezdőfeltételeket kielégítő $x(t; x_0, \dot{x}_0)$ megoldását nyilván a

$$x(t; x_0\, \dot{x}_0)=\cos(\sqrt{A}t).x_0+\frac{\sin(\sqrt{A}t)}{\sqrt{A}}\dot{x}_0 \tag{3}$$

formula szolgáltatja.

Direkt számítással verifikálható továbbá, hogy a

$$\ddot{x}+Ax=g(t)$$

másodrendű inhomogén differenciálegyenletnek az

$$x(0)=0; \quad \dot{x}(0)=0$$

kezdőfeltételeket kielégítő megoldását az

$$x(t)=\int_0^t \frac{\sin\sqrt{A}(t-\tau)}{\sqrt{A}} g(\tau)\, d\tau \tag{4}$$

formula adja.

(Ha az inhomogén differenciálegyenleteknél más kezdőfeltételek vannak előírva, akkor azokat a homogén differenciálegyenlet megfelelő partikuláris megoldásának szuperponálásával elégítjük ki.)

2. §. A matrix-számításnak az analóg differenciálegyenletrendszerek megoldására való alkalmazása mármost lényegileg úgy jellemezhető, hogy a fentebb idézett egyszerű megoldási formulák rendszerekre is érvényesek maradnak, ha az azokban szereplő skalárokat megfelelő matrixokkal helyettesítjük.

Tekintsük először az $\dot{x}=Ax$ egyenletnek és az $x(0)=x_0$ kezdeti feltételeknek megfelelő

$$\begin{aligned}
\dot{x}_1 &= a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n & x_1(0)&=x_{10}\\
\dot{x}_2 &= a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n; & x_2(0)&=x_{20}\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & &\\
\dot{x}_n &= a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n & x_n(0)&=x_{n0}
\end{aligned}$$

homogén differenciálegyenletrendszert és kezdőfeltételt. Ezek a

$$\mathbf{x}(t)=\begin{bmatrix} x_1(t)\\ x_2(t)\\ \vdots\\ x_n(t)\end{bmatrix}, \quad \mathbf{x}_0=\begin{bmatrix} x_{10}\\ x_{20}\\ \vdots\\ x_{n0}\end{bmatrix}$$

vektorok és az $[a_{ij}]=\mathbf{A}$ matrix bevezetésével nyilván az

$$\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}; \quad \mathbf{x}(0)=\mathbf{x}_0 \tag{1_1)(1_2}$$

matrix-egyenletekké foglalhatók össze.

Kimutatjuk, hogy az (1) skalár-formulából a megfelelő helyettesítésekkel keletkező

$$\mathbf{x}(t;\mathbf{x}_0)=e^{\mathbf{A}t}\cdot\mathbf{x}_0$$

vektor kielégíti a (1_1) differenciálegyenletet és az (1_2) kezdőfeltételt. Az $\mathbf{A}$ matrixnak itt szereplő

$$e^{\mathbf{A}t}=\mathbf{E}+\mathbf{A}t+\frac{1}{2!}\mathbf{A}^2t^2+\cdots$$

függvénye ugyanis minden $\mathbf{A}$ matrixra és minden t-re konvergens hatványsorral van definiálva. A t szerint tagonkénti differenciálással nyert sor:

$$\frac{d}{dt}e^{\mathbf{A}t}=\mathbf{A}+\mathbf{A}^2t+\frac{1}{2!}\mathbf{A}^3t^2+\cdots=\mathbf{A}e^{\mathbf{A}t},$$

tehát az $e^{\mathbf{A}t}$ matrix kielégíti az (5) differenciálegyenletet (abban az értelemben, hogy $e^{\mathbf{A}t}$ oszlop-vektorai kielégítik az (5) rendszert és annak egy alaprendszerét alkotják). Továbbá a $t=0$ helyen $e^{\mathbf{A}t}$ nyilván az egységmatrixba megy át.

Innen azonban rögtön következik, hogy az $e^{\mathbf{A}t}\cdot\mathbf{x}_0$ vektor mind az (1_1) differenciálegyenletet, mind az (1_2) kezdőfeltételt kielégíti. $e^{\mathbf{A}t}\cdot\mathbf{x}_0$ ugyanis $e^{\mathbf{A}t}$ oszlop-vektorainak, vagyis partikuláris megoldásoknak homogén lineár kombinációja, tehát maga is megoldás. Továbbá $\lim\limits_{t\to 0}(e^{\mathbf{A}t}\cdot\mathbf{x}_0)=\mathbf{E}\mathbf{x}_0=\mathbf{x}_0$, tehát a kezdőfeltétel is ki van elégítve.

Az analóg meggondolásokat a fentidézett többi egyenlettípusra is elvégezve, a következő eredményekhez jutunk:

Az $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}$ egyenletrendszernek az $\mathbf{x}(0)=\mathbf{x}_0$ kezdőfeltételt kielégítő megoldása:

$$\mathbf{x}(t;\mathbf{x}_0)=e^{\mathbf{A}t}\cdot\mathbf{x}_0. \tag{1'}$$

Az $\dot{\mathbf{x}}-\mathbf{A}\mathbf{x}=\mathbf{g}(t)$ egyenletrendszernek az $\mathbf{x}(0)=0$ kezdőfeltételt kielégítő megoldása:

$$\mathbf{x}(t)=\int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{g}(\tau)\,d\tau. \tag{2'}$$

Az $\ddot{\mathbf{x}}+\mathbf{A}\mathbf{x}=0$ egyenletrendszernek az $\mathbf{x}(0)=\mathbf{x}_0$, $\dot{\mathbf{x}}(0)=\dot{x}_0$ kezdőfeltételeket kielégítő megoldása:

$$\mathbf{x}(t;\mathbf{x}_0,\dot{\mathbf{x}}_0)=\cos(\sqrt{\mathbf{A}}t)\cdot\mathbf{x}_0+\frac{\sin(\sqrt{\mathbf{A}}t)}{\sqrt{\mathbf{A}}}\dot{\mathbf{x}}_0. \tag{3'}$$

Az $\ddot{\mathbf{x}}+\mathbf{A}\mathbf{x}=\mathbf{g}(t)$ rendszernek az $\mathbf{x}(0)=0$, $\dot{\mathbf{x}}(0)=0$ kezdőfeltételeket kielégítő megoldása:

$$\mathbf{x}(t)=\int_0^t\frac{\sin\sqrt{\mathbf{A}}(t-\tau)}{\sqrt{\mathbf{A}}}\mathbf{g}(\tau)\,d\tau. \tag{4'}$$

(Az $\mathbf{A}$ matrix négyzetgyökeinek itteni szereplése látszólagos, tekintettel a cos és sin függvények hatványsorainak paritására.)

A gyakorlati alkalmazásoknál e megoldási formulákban szereplő matrixfüggvényeket a IV. (7) alatti redukált, ill. a IV. (15) alatti kanonikus alakra hozzuk. Redukció után a (2') (4') megoldási formulák csupán skalár-függvények integráljait tartalmazzák.

Ciklikus matrix kanonikus előállítása és annak alkalmazása egy valószínűség-számítási problémára

3. §. Egy n-edrendű $\mathbf{A}=[a_{ij}]$ matrixot akkor nevezünk ciklikusnak, ha annak elemei az indexek különbségének modulo n periodikus függvényei, azaz

$$a_{ij}=c_{j-i};\quad c_{-k}=c_{n-k}\quad (i,j,k=0,1,2,\dots,n-1)$$

alakúak. Ezen definícióból következik, hogy egy ciklikus matrix teljesen meg van határozva első sorának elemei által. Erre való tekintettel egy ciklikus matrixot a következőképpen fogunk jelölni:

$$\begin{bmatrix} c_0 & c_1 c_2 \dots c_{n-1} \\ c_{n-1} c_0 c_1 \dots c_{n-2} \\ \dots\dots\dots \\ c_1 & c_2 c_3 \dots \quad c_0 \end{bmatrix} = \mathbf{C}(c_0, c_1, c_2, \dots, c_{n-1}). \tag{5}$$

A ciklikus matrixokra jellemző a következő tulajdonság:

Bármely n-edrendű $\mathbf{C}(c_0, c_1, \dots, c_{n-1})$ ciklikus matrix kifejezhető, mint az $\Omega=\mathbf{C}(0,1,0,\dots,0)$ primitív ciklikus matrixnak legfeljebb $n-1$-edfokú polinomja. Ugyanis nyilván

$$\mathbf{C}(c_0, c_1, \dots, c_{n-0})=c_0\mathbf{E}+c_1\mathbf{C}(0,1,0,\dots,0)+\cdots,c_{n-1}\mathbf{C}(0,0,\dots,1). \tag{6}$$

Továbbá Ω definíciójából következik, hogy

$$\begin{aligned} \Omega^2 &= \mathbf{C}(0,0,1,0,\dots,0) \\ &\dots\dots\dots\dots \\ \Omega^{n-1} &= \mathbf{C}(0,0,\dots\dots 1) \\ \Omega^n &= \mathbf{E}, \end{aligned}$$

tehát (6)-ból

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

Ezzel feladatunk az $\Omega=\mathbf{C}(0,1,0,\dots,0)$ primitív ciklikus matrix polinomjának kanonikus előállítására van visszavezetve.

Ω karakterisztikus polinomja:

$$|\lambda\mathbf{E}-\Omega|=D(\lambda)=\begin{vmatrix} \lambda & -1 & 0 \dots 0 \\ 0 & \lambda & -1 \dots 0 \\ \dots & \dots & \dots \\ -1 & 0 & 0 \dots \lambda \end{vmatrix}=\lambda^n-1=\prod_{k=0}^{n-1}(\lambda-\omega_k);\ \omega_k=e^{\frac{2k\pi i}{n}}. \tag{7}$$

E szerint Ω sajátértékei az n-edik egységgyökök, ezek mind különbözők és így $D(\lambda)$ Ω-nak egyszersmind minimal polinomja. IV. (8) szerint tehát Ω tetszőleges $\varphi(\Omega)$ polinomjának kanonikus előállítása a következő:

$$\varphi(\Omega)=\sum_{k=1}^{n-1}\varphi(\omega_k)L_k(\Omega), \qquad \omega_k=e^{\frac{2k\pi i}{n}};$$

itt

$$L_k(x)=\frac{D(x)}{D'(\omega_k)(x-\omega_k)}=\frac{x^n-1}{n\omega_k^{n-1}(x-\omega_k)}=\frac{1}{n}\{1+\bar{\omega}_k x+\bar{\omega}_k^2x^2+\cdots+\bar{\omega}_k^{n-1}x^{n-1}\},$$

tehát

$$L_k(\Omega)=\frac{1}{n}(\mathbf{E}+\bar{\omega}_k\Omega+\bar{\omega}_k^2\Omega^2+\cdots+\bar{\omega}_k^{n-1}\Omega^{n-1}),$$

vagyis (7) figyelembevételével

$$L_k(\Omega)=\frac{1}{n}\{\mathbf{C}(1,0,\ldots,0)+\bar{\omega}_k\mathbf{C}(0,1,0,\ldots,0)+\cdots+\bar{\omega}_k^{n-1}\mathbf{C}(0,0,\ldots,1)\}=$$

$$=\frac{1}{n}\mathbf{C}(1,\bar{\omega}_k,\ldots,\bar{\omega}_k^{n-1})=\frac{1}{n}\begin{bmatrix}1 & \bar{\omega}_k & \ldots & \bar{\omega}_k^{n-1}\\ \omega_k & 1 & \ldots & \bar{\omega}_k^{n-2}\\ \cdot & \cdot & \cdot & \cdot\\ \omega_k^{n-1} & \omega_k^{n-2} & \ldots & 1\end{bmatrix}=\frac{1}{n}\begin{bmatrix}1\\ \omega_k\\ \vdots\\ \omega_k^{n-1}\end{bmatrix}[1\,\bar{\omega}_k\ldots\bar{\omega}_k^{n-1}]. \quad (8)$$

$L_k(\Omega)$ tehát hermitikus diád és ennek felhasználásával

$$\varphi(\Omega)=\frac{1}{n}\sum_{k=0}^{n-1}\varphi(\omega_k)\begin{bmatrix}1\\ \omega_k\\ \vdots\\ \omega_k^{n-1}\end{bmatrix}[1\,\bar{\omega}_k\ldots\bar{\omega}_k^{n-1}]. \quad (9)$$

Ha most φ az adott ciklikus matrixnak

$$\mathbf{C}(c_0,c_1,\ldots,c_{n-1})=c_0\mathbf{E}+c_1\Omega+\cdots+c_{n-1}\Omega^{n-1}=\varphi(\Omega)$$

alakú előállításában szereplő polinom, akkor (8)-ból következik, hogy

Valamely n-edrendű $\mathbf{C}(c_0,c_1,\ldots,c_{n-1})$ ciklikus matrix sajátértékei:

$$\lambda_k=c_0+c_1\omega_k+\cdots+c_{n-1}\omega_k^{n-1},\ \omega_k=e^{\frac{2k\pi i}{n}},\ (k=0,1,\ldots,n-1). \quad (10)$$

Sajátvektorai:

$$\mathbf{u}_k=\frac{1}{\sqrt{n}}\begin{bmatrix}1\\ \omega_k\\ \vdots\\ \omega_k^{n-1}\end{bmatrix};\ \mathbf{u}_k^*=\frac{1}{\sqrt{n}}[1,\bar{\omega}_k,\ldots,\bar{\omega}_k^{n-1}] \quad (11)$$

uniter-ortogonális rendszert alkotnak. azaz $\bar{\mathbf{u}}_k^*\mathbf{u}_h=\delta_{kh}$.

Kanonikus előállítása:

$$\mathbf{C}(c_0,c_1,\ldots,c_{n-1})=\sum_{k=0}^{n-1}(c_0+c_1\omega_k+\cdots c_{n-1}\omega_k^{n-1})\,\mathbf{u}_k\bar{\mathbf{u}}_k^*. \quad (12)$$

4. §. Az alábbi alkalmazásra való tekintettel foglalkozzunk a következő elsőrendű differenciálegyenletrendszerrel:

$$\frac{dx_0}{dt} = -\left(\sum_1^{n-1} c_\nu\right) x_0 + c_1 x_1 + \dots + c_{n-1} x_{n-1}$$

$$\frac{dx_1}{dt} = c_{n-1} x_0 - \left(\sum_1^{n-1} c_\nu\right) x_1 + \dots + c_{n-2} x_{n-1}$$

$$\frac{dx_{n-1}}{dt} = c_1 x_0 + c_2 x_2 + \dots - \left(\sum_1^{n-1} c_\nu\right) x_{n-1}.$$

Itt $c_0, c_1, \dots c_{n-1}$ nem-negatív számok, melyek nem mind tűnnek el. A $t=0$-ra előírt kezdőértékek tetszőleges $\xi_0, \xi_1, \dots \xi_{n-1}$ nem-negatív számok, melyek a $\sum_0^{n-1} \xi_\nu = 1$ normirozást kielégítik.

Matrix-formában

$$\frac{d}{dt}\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{bmatrix} = \begin{bmatrix} -\Sigma c_\nu & c_1 & \dots & c_{n-1} \\ c_{n-1} & -\Sigma c_\nu & \dots & c_{n-2} \\ \cdot & \cdot & \dots & \cdot \\ c_1 & c_2 & & -\Sigma c_\nu \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{bmatrix}; \quad \begin{bmatrix} x_0(0) \\ x_1(0) \\ \vdots \\ x_{n-1}(0) \end{bmatrix} = \begin{bmatrix} \xi_0 \\ \xi_1 \\ \vdots \\ \xi_{n-1} \end{bmatrix}; \quad \sum_0^{n-1} \xi_\nu = 1,$$

azaz

$$\dot{\mathbf{x}} = \mathbf{C}\mathbf{x}; \quad \mathbf{x}(0) = \xi; \quad \sum_0^{n-1} \xi_\nu = 1. \qquad (11)\ (12)$$

A $\mathbf{C}$ ciklikus matrix sajátértékei (10) szerint

$$\lambda_k = \sum_1^{n-1} c_\nu \left(e^{\frac{2k\nu\pi i}{n}} - 1\right) = -2\sum_1^{n-1} c_\nu \sin^2 \frac{k\nu\pi}{n} + i \sum_1^{n-1} c_\nu \sin \frac{2k\nu\pi}{n}; \quad \lambda_k = \overline{\lambda_{n-k}}. \qquad (13)$$

$$(k = 0, 1, \dots n-1).$$

E szerint ahhoz, hogy $\lambda_0 = 0$ kivételével valamennyi sajátérték negatív valós résszel bírjon, elegendő, hogy az alábbi feltételek valamelyike teljesüljön:

Valamennyi c_ν pozitív, vagy $\sum_1^{n-1} c_\nu > 0$ és n primszám. (Vagy általánosabban: azon ν indexek legnagyobb közös osztója, melyekre $c_\nu > 0$, relatív prim n-hez.)

Ekkor a (11) differenciálegyenletnek a (12) kezdeti feltételt kielégítő megoldása

$$\mathbf{x}(t;\xi) = e^{\mathbf{C}t}\xi = \frac{1}{n}\left\{ \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} [1, 1, \dots 1] + \sum_1^{n-1} e^{\lambda_k t} \begin{bmatrix} 1 \\ \omega_k \\ \vdots \\ \omega_k^{n-1} \end{bmatrix} [\bar{1}\ \bar{\omega}_k \dots \bar{\omega}_k^{n-1}] \right\} \xi. \qquad (14)$$

Ha az előbbi feltételek valamelyike teljesül, akkor $\Re(\lambda) < 0$, $k = 1, 2, \dots n-1$, továbbá (12) szerint $[1, 1, \dots 1]\,\xi = 1$, tehát

$$\lim_{t\to\infty}\mathbf{x}(t;\xi)=\frac{1}{n}\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}[1,1,\dots 1]\xi=\frac{1}{n}\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix},$$

vagyis az $\mathbf{x}(t;\xi)$ vektor $t\to\infty$ esetén a ξ kezdeti vektortól független $\left[\frac{1}{n},\frac{1}{n},\dots\frac{1}{n}\right]^*$ vektorhoz konvergál.

5. §. A ciklikus matrixok kanonikus előállítására vonatkozó fenti eredmények érdekes alkalmazást nyerhetnek a valószínűségszámításban, mégpedig mod n additív és homogén Markov-folyamatok ergodicitásának vizsgálatánál. Ezek a folyamatok a következő egyszerű modellel realizálhatók: Egy pont végezzen „bolyongás"-t az egységkör kerületén, lehetséges helyzetei legyenek egy az egységkörbe írt szabályos n-szög $A_1, A_2,\dots A_n$ csúcspontjai. Tegyük fel, hogy a $t>0$ időpontban a pont az A_j csúcsban van, és ezen feltétel mellett vizsgáljuk annak a valószínűségét, hogy a $t+\tau$ időpontban ($\tau\geqq 0$) a pont az A_k csúcsban legyen. Tegyük fel, hogy ez a valószínűség nem függ attól, hogy egy tetszőleges $\tau_0<t$ időpontban hol volt a pont (vagyis, hogy a folyamat Markov-típusú), továbbá, hogy ez a valószínűség nem függ t-től (vagyis, hogy a folyamat időben homogén) végül, hogy ez a valószínűség t-n kívül csak a $k-j$ különbség mod n vett maradékától függ. (Vagyis, hogy a folyamat mod n additív.)

Ezen feltevések mellett a szóbanforgó valószínűséget jelölhetjük $p_{k-j}(t)$-vel, ahol definíció szerint $p_r(t)\equiv p_s(t)$, ha $r\equiv s$ (mod n).

A $p_\nu(t)$ függvények értelmezéséből következik, hogy

$$p_\nu(t)\geqq 0(\nu=0,1,\dots n-1) \text{ és } \sum_{\nu=0}^{n-1}p_\nu(t)=1, \tag{15}$$

továbbá, hogy $\mathbf{P}(t)$-vel jelölve a $[p_{k-j}(t)]$ ciklikus matrixot,

$$\mathbf{P}(t+\tau)=\mathbf{P}(t)\mathbf{P}(\tau) \text{ és } \mathbf{P}(0)=\mathbf{E}. \qquad (16)\ (17)$$

Ha még feltesszük, hogy léteznek a $p'_\nu(0)=c_\nu$ deriváltak [14] és a $[c_{k-j}]$ ciklikus matrixot $\mathbf{C}$-vel jelöljük, úgy (16)-ből:

$$\mathbf{P}'(t)=\mathbf{P}(t)\cdot\mathbf{C}=\mathbf{C}\cdot\mathbf{P}(t) \tag{18}$$

($\mathbf{P}'(t)$ létezését $t>0$-ra nem kell feltenni, ez következik $\mathbf{P}'(0)$ létezéséből és (16)-ből.)

A (15) feltevésekből következik, hogy $c_\nu>0$, ha $\nu=1,2,\dots n-1$ és $c_0=-\sum_{1}^{n-1}c_\nu$. Eszerint a probléma a fentebb tárgyalt differenciálegyenletrendszerre van visszavezetve és így az előbb nyert eredményünk értelmében (ha azon ν számok legnagyobb közös osztója, melyekre $c_\nu>0$, relatív prím n-hez), akkor

[14] Elegendő a $p_\nu(t)$ függvények folytonosságát a $t=0$ helyen feltételezni. L. pl. J. L. Doob, Stochastic Processes. Wiley, 1953. Chap. VI.

$$\lim_{t\to\infty} p_\nu(t) = \frac{1}{n} \qquad (\nu = 0, 1, \ldots n-1). \quad (19)$$

Tegyük most fel, hogy a $t=0$ időpontban a pont w_j valószínűséggel tartózkodott az A_j csúcsban $\left(\sum_1^n w_j = 1\right)$, úgy annak valószínűségét, hogy a $t>0$ időpontban a pont az A_k csúcsban legyen, $W_k(t)$-vel jelölve

$$W_k(t) = \sum_{j=1}^{n} w_j p_{k-j}(t)$$

és így (19) szerint

$$\lim_{t\to\infty} W_k(t) = \frac{1}{n} \qquad (k = 1, 2, \ldots n).$$

Tehát tetszőleges kezdeti valószínűség eloszlás mellett a pont $t\to\infty$-re határértékben egyforma valószínűséggel lesz található a szabályos n-szög bármely csúcsában.

Ilyen módon a ciklikus matrix kanonikus előállítása segélyével kimutattuk a vizsgált Markov-folyamat ergodicitását.[15]

Szimmetrikus matrix kanonikus előállítása és annak alkalmazása egy mechanikai problémára

6. §. Végesszámú szabadsági fokkal bíró, konzervatív mechanikai rendszernek stabilis egyensúlyi konfiguráció kis környezetében végbemenő rezgései tudvalevőleg másodrendű lineáris állandó együtthatós differenciálegyenletrendszer által vannak meghatározva, melynek megoldására a klasszikus mechanika a változók szeparálásának, más néven normál-koordináták bevezetésének elvét alkalmazza.

Az alábbiakban egy egyszerű példán, a korpuszkuláris húr-modell tárgyalása kapcsán meg fogjuk mutatni, hogy a rendszer megoldását szolgáltató trigonometrikus matrix-függvény kanonikus előállítása éppen a fenti elveknek megfelelő megoldási formára vezet.

Korpuszkuláris húr-modellnek nevezzük rugalmas, tömegtelen fonál egyes pontjaihoz rögzített tömegpontok rendszerét. Legyen a tömegpontok száma $n-1$, mindegyiküknek tömege: $m=1$, a szomszédosak távolsága: $l=1$ és a fonál rugalmas feszültsége: $F=1$.

Ha a k-dik tömegpontnak az egyensúlyi helyzettől számított transzverzális elmozdulása t időpontban $x_k(t)$, akkor — kis kilengések esetén — a tömegpontok mozgását a

$$\ddot{x}_k + (-x_{k-1} + 2x_k - x_{k+1}) = 0 \qquad (k = 1, 2, \ldots n-1) \quad (20)$$

[15] A matrix-elméletnek fenti alkalmazására Rényi Alfréd volt szíves a figyelmemet felhívni.

differenciálegyenletrendszer és a végpontok rögzítettségét kifejező

$$x_0 \equiv 0\,;\; x_n \equiv 0 \tag{21}$$

kerületi feltétel határozza meg.

A (20) differenciálegyenletek *és* a (21) kerületi feltételek összefoglalhatók a kontinuáns típusú matrix-szal bíró

$$\frac{d^2}{dt^2}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \end{bmatrix} + \begin{bmatrix} 2 & -1 & \dots & 0 \\ -1 & 2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots -1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \end{bmatrix} = 0 \text{ ill. } \ddot{\mathbf{x}} + \mathbf{A}\mathbf{x} = 0$$

vektor-differenciálegyenletté.

A tömegpontoknak a $t=0$ pillanatban elfoglalt kezdőhelyzetei $x_{10}, x_{20}, \dots x_{n-1,0}$ komponensekkel bíró $\mathbf{x}_0$ vektorrá, kezdő sebességei az $\dot{x}_{10}, \dot{x}_{20}, \dots \dot{x}_{n-1,0}$ komponensekkel bíró $\dot{\mathbf{x}}_0$ vektorrá foglalhatók össze. A kezdeti feltételek tehát:

$$\mathbf{x}(0) = \mathbf{x}_0\,;\; \dot{\mathbf{x}}(0) = \dot{\mathbf{x}}_0.$$

(3) szerint a (20) differenciálegyenletnek a kezdeti feltételeket kielégítő megoldása:

$$\mathbf{x}(t;\, \mathbf{x}_0, \dot{\mathbf{x}}_0) = \cos(\sqrt{\mathbf{A}}t)\cdot \mathbf{x}_0 + \frac{\sin(\sqrt{\mathbf{A}}t)}{\sqrt{\mathbf{A}}}\cdot \dot{\mathbf{x}}_0. \tag{3}$$

7. §. Ahhoz, hogy ebből az igen tömör megoldási formulából a normálkoordinátákban kifejezett megoldáshoz jussunk, az itt szereplő matrixfüggvényeket kanonikus alakjukra kell hoznunk. IV. (15) szerint

$$\cos(\sqrt{\mathbf{A}}t) = \sum_{k=1}^{n-1}\cos(\sqrt{\lambda_k}t)\cdot L_k(\mathbf{A});\quad \frac{\sin(\sqrt{\mathbf{A}}t)}{\sqrt{\mathbf{A}}} = \sum_{k=1}^{n-1}\frac{\sin(\sqrt{\lambda_k}t)}{\sqrt{\lambda_k}}L_k(\mathbf{A}), \tag{22}$$

ahol $\lambda_1 \dots \lambda_{n-1}$ az $|\mathbf{A}-\lambda\mathbf{E}| = 0$ karakterisztikus egyenlet gyökei (melyek ez esetben mind különbözőknek fognak bizonyulni).

Az $|\mathbf{A}-\lambda\mathbf{E}|$ kontinuáns 0-helyeinek a kiszámításánál célszerű a $\lambda = 4\sin^2\frac{\theta}{2}$ relációval egy új θ parametert bevezetni. Ezzel a jelöléssel[16]

$$|\mathbf{A}-\lambda\mathbf{E}| = |\mathbf{A}-4\sin^2\frac{\theta}{2}\mathbf{E}| = \begin{vmatrix} 2\cos\theta & -1 & \dots & 0 \\ -1 & 2\cos\theta & \dots & 0 \\ \cdot & & & \\ 0 & 0 & \dots & 2\cos\theta \end{vmatrix}_{n-1} = \frac{\sin n\theta}{\sin\theta}. \tag{23}$$

Tehát a $D_{n-1}(\lambda) = |\mathbf{A}-\lambda\mathbf{E}| = 0$ karakterisztikus egyenlet gyökei a θ parameter $\theta_k = \frac{k\pi}{n}$ értékeinek felelnek meg, azaz

$$\lambda_k = 4\sin^2\frac{\theta_k}{2} = 4\sin^2\frac{k\pi}{n} \qquad (k = 1, 2, \dots n-1). \tag{23'}$$

[16] L. PL. E. Pascal, Die Determinanten. Leipzig. 1900. pp. 155—156.

Miután a gyökök mind egyszeresek, ezért a (22) előállításban szereplő, $L_k(\mathbf{A})$ matrixok mind elsőrangúak, vagyis egyetlen diádra redukálódnak.

Az $L_k(\mathbf{A})$ matrixok effektiv kiszámításához ezúttal célszerűbb azoknak az adj. $(\mathbf{A}-\lambda\mathbf{E})$ matrix-szal való IV. (21) kapcsolatát felhasználni.

Jelöljünk egy k-adrendű $|\mathbf{A}-\lambda\mathbf{E}|$ típusú kontinuánst $D_k(\lambda)$-val, továbbá adj. $(\mathbf{A}-\lambda\mathbf{E})$-nek az i-edik sorban és j-dik oszlopban álló elemét $A_{ij}(\lambda)$-val.

Ekkor könnyen belátható, hogy $i \leqq j$ esetén

$$A_{ij}(\lambda) = (-1)^{j-i} D_{i-1}(\lambda) D_{n-1-j}(\lambda). \tag{24}$$

Ebből az algebrai komplementum $(-1)^{i+j}$-nel való szorzással adódik, tehát (23) figyelembevételével

$$A_{ij}(\lambda) = \frac{\sin i\theta}{\sin\theta}\,\frac{\sin(n-j)\theta}{\sin\theta}; \quad \text{ha} \quad i \leqq j$$

és $\mathbf{A}-\lambda\mathbf{E}$ szimmetriája folytán

$$A_{ij}(\lambda) = \frac{\sin(n-i)\theta}{\sin\theta}\,\frac{\sin j\theta}{\sin\theta}, \quad \text{ha} \quad i \geqq j$$

IV. (21) szerint:

$$L_k(\mathbf{A}) = -\frac{1}{D'_{n-1}(\lambda_k)}\,\text{adj.}\,(\mathbf{A}-\lambda_k\mathbf{E}) = -\frac{1}{D'_{n-1}(\lambda_k)}\,[A_{ij}(\lambda_k)]$$

(24) alapján

$$A_{ij}(\lambda_k) = A_{ij}\left(4\sin^2\frac{k\pi}{n}\right) = \frac{\sin\frac{ki\pi}{n}}{\sin\frac{k\pi}{n}}\cdot\frac{\sin\frac{k(n-j)\pi}{n}}{\sin\frac{k\pi}{n}} = (-1)^k\,\frac{\sin\frac{ki\pi}{n}}{\sin\frac{k\pi}{n}}\cdot\frac{\sin\frac{kj\pi}{n}}{\sin\frac{k\pi}{n}},$$

továbbá

$$D'_{n-1}(\lambda_k) = \frac{d}{d\sin^2\frac{\theta}{2}}\,D_{n-1}\left(4\sin^2\frac{\theta}{2}\right)\Bigg|_{\theta=\frac{k\pi}{n}} (-=1)^k\,\frac{n}{2\sin^2\frac{k\pi}{n}},$$

tehát mindezek figyelembevételével

$$L_k(\mathbf{A}) = \mathbf{u}_k\mathbf{u}_k^*;\quad \mathbf{u}_k^* = \sqrt{\frac{2}{n}}\left[\sin\frac{k\pi}{n},\ \sin\frac{2k\pi}{n},\ldots \sin\frac{(n-1)k\pi}{n}\right]. \tag{25}$$

Könnyen verifikálható, hogy az így megszerkesztett vektorrendszer ortogonális, azaz kielégíti a $\mathbf{u}_h^*\mathbf{u}_l = \delta_{kl}$ relációkat.

Az $L_k(\mathbf{A})$ matrixok ezen kifejezéseit, valamint λ_k (23)-ből adódó értékeit a trigonometrikus matrix-függvények (22) képletébe, majd ezeket a (3) megoldási képletbe helyettesítve, a húr-modell differenciálegyenletrendszerének az alábbi explicit megoldását nyerjük:

$$\mathbf{x}(t;\mathbf{x}_0,\dot{\mathbf{x}}_0) = \frac{2}{n}\sum_{k=1}^{n-1}\mathbf{u}_k\mathbf{u}_k^*\left\{\mathbf{x}_0\cos\left(2\sin\frac{k\pi}{2n}t\right) + \dot{\mathbf{x}}_0\,\frac{\sin\left(2\sin\frac{k\pi}{2n}t\right)}{2\sin\frac{k\pi}{n}}\right\}$$

vagy részletes skalár írásmódban:

$$x_i(t) = \frac{2}{n}\sum_{k=1}^{n-1} \sin\frac{ki\pi}{n}\left\{(\mathbf{u}_k^*\mathbf{x}_0)\cos\left(2\sin\frac{k\pi}{n}t\right) + (\mathbf{u}_k^*\dot{\mathbf{x}}_0)\frac{\sin\left(2\sin\frac{k\pi}{2n}t\right)}{2\sin\frac{k\pi}{n}}\right\}.$$

Itt a kezdeti adatoktól való függést az $(\mathbf{u}_k^*\cdot\mathbf{x}_0)$ és $(\mathbf{u}_k^*\cdot\dot{\mathbf{x}}_0)$ skalár szorzatok juttatják kifejezésre. Ez a tény azzal függ össze, hogy az $X_k = \mathbf{u}_k^*\cdot\mathbf{x}$ skalárok a rendszer normálkoordinátái, vagyis, hogy ezekkel kifejezve a tömegpontrendszernek úgy a kinetikus, valamint a potenciális energiája négyzetösszegre redukálódik.

Valóban, a kinetikus energia:

$$\sum \dot{x}_k^2 = \dot{\mathbf{x}}^*\dot{\mathbf{x}} = \dot{\mathbf{x}}^*\left(\sum \mathbf{u}_k\mathbf{u}_k^*\right)\dot{\mathbf{x}} = \sum (\mathbf{u}_k^*\dot{\mathbf{x}})^2 = \sum \dot{X}_k^2$$

és a potenciális energia:

$$\mathbf{x}^*\mathbf{A}\mathbf{x} = \mathbf{x}^*\sum(\mathbf{u}_k\lambda_k\mathbf{u}_k^*)\mathbf{x} = \sum \lambda_k(\mathbf{u}_k^*\mathbf{x})^2 = \sum \lambda_k X_k^2,$$

tehát a

$$X_k = \mathbf{u}_k^*\mathbf{x} = \sqrt{\frac{2}{n}}\left\{x_1\sin\frac{k\pi}{n} + x_2\sin\frac{2k\pi}{n} + \ldots + x_{n-1}\sin\frac{(n-1)k\pi}{n}\right\}$$

új függő változók bevezetésével a (20) differenciálegyenletrendszer szétesik az egyes normálrezgéseket meghatározó

$$\ddot{X}_k + \lambda_k X_k = 0 \qquad (k = 1, 2, \ldots, n-1)$$

szeparált differenciálegyenletekre.