2009. november 3/4.

Lineáris algebra (A, B, C)
8. előadás
(vázlat)

[Egy másik ötlet a szorzástétel bizonyítására: Az $|AB|$-ban a mátrixszorzás definíciója alapján mindenütt n tagú összegek vannak, csak az a baj, hogy összeg determinánsát nem számolhatjuk tagonként. Egy sorban lévő összegekkel azonban már tudunk bánni (a többi marad változatlan), elég soronként más és más futóindexet használni, s a szummákat egymás után kivihetjük:

$$|AB| =$$

$$\begin{vmatrix} \sum_{k_1=1}^n a_{1k_1}b_{k_11} & \dots & \sum_{k_1=1}^n a_{1k_1}b_{k_1n} \\ \vdots & & \vdots \\ \sum_{k_n=1}^n a_{nk_n}b_{k_n1} & \dots & \sum_{k_n=1}^n a_{nk_n}b_{k_nn} \end{vmatrix} = \sum_{k_1=1}^n \dots \sum_{k_n=1}^n \begin{vmatrix} a_{1k_1}b_{k_11} & \dots & a_{1k_1}b_{k_1n} \\ \vdots & & \vdots \\ a_{nk_n}b_{k_n1} & \dots & a_{nk_n}b_{k_nn} \end{vmatrix}$$

$$= \sum_{k_1=1}^n \dots \sum_{k_n=1}^n a_{1k_1}\cdot \ldots \cdot a_{nk_n} \begin{vmatrix} b_{k_11} & \dots & b_{k_1n} \\ \vdots & & \vdots \\ b_{k_n1} & \dots & b_{k_nn} \end{vmatrix} = \sum_{k_1=1}^n \dots \sum_{k_n=1}^n a_{1k_1}\cdot \ldots \cdot a_{nk_n} \begin{vmatrix} {}_{k_1}[B] \\ \vdots \\ {}_{k_n}[B] \end{vmatrix}.$$

Az összeg olyan tagjai, amelyekben a $k_1, \dots, k_n$ között van két egyező, biztos nullát adnak (van két azonos sor!). A megmaradó tagok már az $1, 2, \dots, n$ permutációira való összegezéssel írhatók (becsempészve a hiányzó hatványt):

$$= \sum_{\substack{k_1,\dots,k_n \\ (1,\dots,n)}} (-1)^{I(k_1,\dots,k_n)} a_{1k_1}\cdot \ldots \cdot a_{nk_n} (-1)^{I(k_1,\dots,k_n)} \begin{vmatrix} {}_{k_1}[B] \\ \vdots \\ {}_{k_n}[B] \end{vmatrix}.$$

A $(-1)^{I(k_1,\dots,k_n)} \begin{vmatrix} {}_{k_1}[B] \\ \vdots \\ {}_{k_n}[B] \end{vmatrix}$ cserék — mint 7/2-n — révén épp $|B|$, így előtte már $|A|$ áll.]

A valós elemű négyzetes mátrixok determinánsáról szóló seregnyi tételünk még kevés az alkalmazásokhoz, legalább említenünk kell, hogy mi történik, ha a négyzetes mátrixunk elemeit nem $\mathbb{R}$-ből, hanem más struktúrából vesszük. Ha pl. $\mathbb{C}$-ből vesszük az elemeket, akkor hasonló definícióval dolgozhatunk, s ugyanolyan tételeket nyerhetünk. Ugyanez mondható, ha számtestből vesszük az elemeket (**számtest**: a $\mathbb{C}$ olyan részhalmaza, melyben benne van a 0 és az 1, továbbá zárt a $\mathbb{C}$-ben tanult összeadásra, kivonásra, szorzásra és nullától különböző számmal való osztásra). Ha az elemeket egy $\mathcal{R}$ **gyűrű**ből vesszük, akkor egységelem hiányában már a determináns definíciójával is lehet gond, de ezt még megoldhatjuk azzal, hogy amikor -1 szorzó kellene, akkor az utána lévő szorzat ellentettjét vesszük; de az egységmátrix azért nagyon hiányozna. A kommutativitás hiánya még kellemetlenebb lenne: $\begin{vmatrix} a & a \\ b & b \end{vmatrix} = ab - ab = 0$, de $\begin{vmatrix} a & b \\ a & b \end{vmatrix} = ab - ba$, ami lehet 0-tól különböző egy nem kommutatív gyűrűben, így már sajnos $|A^\top| \overset{\acute{a}lt}{\neq} |A|$. Ha az elemeket egy $\mathcal{R}$ **egységelemes, kommutatív és nullosztómentes gyűrű**ből vesszük, akkor már érvényesek lesznek a korábbi eredményeink közül mindazok, ahol nem használtunk sem osztást, sem osztásra épülő fogalmat a bizonyításban. [A két egyező sor esetére vonatkozó tétel (6/3) bizonyítása már elakad, egy jóval hosszabb bizonyítás mégis működik.]

Gyakori fogás bizonyításokban egy valós számhoz valamit hozzáadni, és ugyanazt le is vonni. Ennek multiplikatív megfelelője: $a, d \in \mathbb{R}$, $d \neq 0$ esetén $a = add^{-1} = d^{-1}ad$, hisz a valós számok szorzása kommutatív. Mátrixoknál már más a helyzet: $A, D \in \mathbb{R}^{n\times n}$, $\det(D) \neq 0$ esetén $\exists D^{-1}$ és $A = ADD^{-1}$, de A és $D^{-1}AD$ különbözhet egymástól, persze nem mindig: $A = D^{-1}AD \iff DA = AD$. Ha tehát az adott A-hoz létezik olyan **invertálható** D, amellyel az A nem cserélhető fel, akkor $A \neq D^{-1}AD$.

DEFINÍCIÓ: $A, B \in \mathbb{R}^{n\times n}$ esetén azt mondjuk, hogy az A hasonló $\mathbb{R}$ felett a B-hez (jelölés: $A \sim_{\mathbb{R}} B$), ha létezik olyan invertálható $D \in \mathbb{R}^{n\times n}$, melyre $B = D^{-1}AD$.

Mivel a mátrixszorzás nem kommutatív, ezért sok esetben változni fog a mátrix, amit esetleg kihasználhatunk csúnya számolás rövidítésére.
Tegyük fel, hogy adott egy csúnya (a főátlón kívül sok nullától különböző elemet tartalmazó) $A \in \mathbb{R}^{n\times n}$ mátrixunk, melyet egy nagy k hatványra kell emelnünk. Ekkor az A-hoz $\mathbb{R}$ felett hasonló mátrixok között keresgélünk szebb alakút, pl. diagonális mátrixot (ha szerencsénk van!!!). Ekkor az A az $\mathbb{R}$ felett hasonló egy diagonális mátrixhoz.

DEFINÍCIÓ: Az A diagonalizálható $\mathbb{R}$ felett, ha $\mathbb{R}$ felett hasonló egy diagonális mátrixhoz.

Ha tehát szerencsénk volt, akkor van olyan invertálható D, hogy

$B = D^{-1}AD = \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & \lambda_n \end{bmatrix}$. Ekkor $D^{-1}ADD^{-1}AD = D^{-1}A^2D$ mintájára

$B^k = (D^{-1}AD)^k = D^{-1}A^kD$, azaz $A^k = D(D^{-1}AD)^kD^{-1} = DB^kD^{-1} =$

$$D\left(\begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & \lambda_n \end{bmatrix}\right)^k D^{-1} = D\begin{bmatrix} \lambda_1^k & 0 & ... & 0 \\ 0 & \lambda_2^k & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & \lambda_n^k \end{bmatrix} D^{-1}.$$

A haszon tehát látszik, de mit is jelenthet, hogy szerencsénk volt? Az $\mathbb{R}^n$ $\mathbf{e}_1, \ldots, \mathbf{e}_n$ triviális bázisát tekintve, a fenti diagonális B-re teljesül, hogy $B\mathbf{e}_1 = \lambda_1\mathbf{e}_1, \ldots, B\mathbf{e}_n = \lambda_n\mathbf{e}_n$, azaz $B\mathbf{e}_j = \lambda_j\mathbf{e}_j$ $(j = 1, \ldots, n)$. Ezt úgy emlegetjük majd, hogy $\mathbf{e}_j$ „sajátvektora" B-nek λ_j „sajátértékkel". A definíció elvileg több formában is megadható, legjobb, ha együtt szerepel a kétféle „sajátvektor"–„sajátérték" definíciója.

DEFINÍCIÓ: Legyen n pozitív egész, $A \in \mathbb{R}^{n\times n}$.
Az $\mathbf{x} \in \mathbb{R}^n$ **jobb oldali sajátvektora** A-nak, ha
1./ $\mathbf{x} \neq \mathbf{0}$,
2./ $\exists\, \lambda_0 \in \mathbb{R}\,:\, A\mathbf{x} = \lambda_0\mathbf{x}$.
Ilyenkor a λ_0 az $\mathbf{x}$ j.o. sajátvektorhoz tartozó j.o. sajátértéke az A-nak.

DEFINÍCIÓ: Legyen n pozitív egész, $A \in \mathbb{R}^{n\times n}$.
Az $\mathbf{y} \in \mathbb{R}^{1\times n}$ **bal oldali sajátvektora** A-nak, ha
1./ $\mathbf{y} \neq \mathbf{0} = [0, ..., 0]$,
2./ $\exists\, \mu_0 \in \mathbb{R}\,:\, \mathbf{y}A = \mu_0\mathbf{y}$.
Ilyenkor a μ_0 az $\mathbf{y}$ b.o. sajátvektorhoz tartozó b.o. sajátértéke az A-nak.

Gyorsan redukáljuk a kétféle keresést egyre: $\mathbf{y}A = \mu_0\mathbf{y} \Leftrightarrow A^\top\mathbf{y}^\top = \mu_0\mathbf{y}^\top$ alapján az A bal oldali sajátvektorainak és sajátértékeinek keresése visszavezethető az $A^\top$ jobb oldali sajátvektorainak és sajátértékeinek keresésére.

Ezentúl jobb oldali sajátvektorokat és sajátértékeket, röviden sajátvektorokat és sajátértékeket keresünk.

Visszatérve a fenti szép diagonális B-re, a $B\mathbf{e}_j = \lambda_j \mathbf{e}_j$ $(j = 1, \ldots, n)$ észrevételt átfogalmazzuk a („csúnya”) A-ra ($B = D^{-1}AD,\ A = DBD^{-1}$):
$D^{-1}AD\mathbf{e}_j = \lambda_j \mathbf{e}_j$ $(j = 1, \ldots, n)$, azaz $A(D\mathbf{e}_j) = \lambda_j (D\mathbf{e}_j)$ $(j = 1, \ldots, n)$.
$D = [\mathbf{d}_1, \ldots, \mathbf{d}_n]$ esetén $D = DI_n = D[\mathbf{e}_1, \ldots, \mathbf{e}_n] = [D\mathbf{e}_1, \ldots, D\mathbf{e}_n]$ miatt $\mathbf{d}_j = D\mathbf{e}_j$ $(j = 1, \ldots, n)$. Tehát $A\mathbf{d}_j = \lambda_j \mathbf{d}_j$ $(j = 1, \ldots, n)$; továbbá D^{-1} létezése miatt $\varrho(D) = n$, így $\mathbf{d}_1, \ldots, \mathbf{d}_n$ L $\mathbb{R}^n$-ben, végül $\mathbf{d}_1, \ldots, \mathbf{d}_n$ bázis $\mathbb{R}^n$-ben, s eme bázis minden vektora sajátvektora az A-nak. Ezzel igazoltuk a következő tétel $\Rightarrow$ irányát, a másik pedig fordított sorrendben olvasható ki a fentiekből:

TÉTEL: Legyen $A \in \mathbb{R}^{n \times n}$.
A diagonalizálható $\mathbb{R}$ felett $\iff$ létezik $\mathbb{R}^n$-ben az A sajátvektoraiból álló bázis (röviden: SB).

Adott sajátérték esetén tekintsük az összes ilyen sajátértékkel rendelkező sajátvektort. Ha ezekhez most hozzávesszük a nullvektort, akkor $\mathbb{R}^n$ egy alterét kapjuk, az adott sajátértékhez tartozó ún. sajátalteret:
DEFINÍCIÓ: Legyen $A \in \mathbb{R}^{n \times n}$, $\lambda_0 \in \mathbb{R}$ pedig egy (jobb oldali) sajátértéke az A-nak. A λ_0-hoz tartozó **sajátaltér**: $W_{\lambda_0} \stackrel{\text{def}}{=} \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n,\ A\mathbf{x} = \lambda_0 \mathbf{x}\}$.

Legyen pl. $n = 2$, $A = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix}$. Az A sajátvektorait keressük $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2,\ \mathbf{x} \neq \mathbf{0}$ alakban. $A\mathbf{x} = \lambda\mathbf{x} \iff \{3x_1 = \lambda x_1,\ x_1 + 2x_2 = \lambda x_2\} \iff \{0 = (\lambda - 3)x_1,\ x_1 = (\lambda - 2)x_2\}$. I. eset: $x_1 \neq 0$: $\lambda = 3, x_1 = x_2$. II. eset: $x_1 = 0$: $(\lambda - 2)x_2 = 0$. Itt már $x_2 \neq 0$, különben $\mathbf{x} = \mathbf{0}$ volna; így most $\lambda = 2$. Összefoglalva: az A sajátértékei: $\lambda_1 = 3,\ \lambda_2 = 2$. A megfelelő sajátalterek: $W_3 = \left\{ \begin{bmatrix} x \\ x \end{bmatrix} \mid x \in \mathbb{R} \right\}$, $W_2 = \left\{ \begin{bmatrix} 0 \\ x \end{bmatrix} \mid x \in \mathbb{R} \right\}$. Létezik SB, pl. $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ SB. Ebből készítve a D oszlopait: $D = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, D invertálható, $D^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$, végül $D^{-1}AD = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$. Nagy méret esetén ez a favágó módszer rengeteg esetszétválasztást kívánhat.
A következő ekvivalencia-sorozat azt mutatja majd, hogy először a lehetséges λ_0 értékeket célszerű megkeresni (ez lesz a problémás), azután már csak homogén lineáris egyenletrendszer (nemtriviális) megoldását kívánja azon sajátvektorok megkeresése, amelyekhez a λ_0 sajátérték tartozik.
$\lambda_0 \in \mathbb{R}$ esetén
λ_0 (jobb oldali) sajátértéke A-nak $\iff \exists\, \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0} \,:\, A\mathbf{x} = \lambda_0 \mathbf{x} \iff$
$\exists\, \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0} \,:\, A\mathbf{x} - \lambda_0 \mathbf{x} = \mathbf{0} \iff$
$\exists\, \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0} \,:\, (A - I_n\lambda_0)\mathbf{x} = \mathbf{0} \iff |A - I_n\lambda_0| = 0$.
Itt

$$|A - I_n\lambda_0| = \begin{vmatrix} a_{11} - \lambda_0 & a_{12} & \ldots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} - \lambda_0 & \ldots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \ldots & a_{n-1,n-1} - \lambda_0 & a_{n-1,n} \\ a_{n1} & a_{n2} & \ldots & a_{n,n-1} & a_{nn} - \lambda_0 \end{vmatrix} =$$

$(-1)^n \lambda_0^n + (-1)^{n-1}(a_{11} + a_{22} + a_{33} + \ldots + a_{nn})\lambda_0^{n-1} +$„CSÚF”$+ |A|$, ahol $a_{11} + a_{22} + \ldots + a_{nn} = \mathrm{Tr}(A)$, az A mátrix nyoma (trace, Spur), a főátlóban lévő elemek összege, a CSÚF jelző

pedig arra utal, hogy $n \geqq 3$ és $n-2 \geqq j \geqq 1$ esetén a λ_0^j együtthatóját nehéz felírni, mert nemcsak a főátlóból származhat.
Az már kiderült, hogy először a sajátértékeket kell megkeresnünk. Egy $\mathbb{R}$ feletti $n \times n$-es A mátrix jobb oldali sajátértékeit úgy találhatjuk meg, hogy az $|A - I_n\lambda_0| = 0$ valós gyökeit keressük meg. Mivel $|A^\top - I_n\lambda_0| = |A - I_n\lambda_0|$, azért az A bal oldali és jobb oldali sajátértékei megegyeznek. Az is leolvasható, hogy pl. felső háromszög mátrix sajátértékei a főátlóban lévő elemek.

$$|A - I_n\lambda| \stackrel{\text{def}}{=} \begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \dots & a_{n-1,n-1}-\lambda & a_{n-1,n} \\ a_{n1} & a_{n2} & \dots & a_{n,n-1} & a_{nn}-\lambda \end{vmatrix} =$$

$$(-1)^n\lambda^n + (-1)^{n-1}(a_{11}+a_{22}+a_{33}+\dots+a_{nn})\lambda^{n-1} + \text{„CSÚF"} + |A|.$$

Ezzel a λ-nak egy n-edfokú polinomja adódott.
DEFINÍCIÓ: Legyen n pozitív egész, $A \in \mathbb{R}^{n\times n}$.
Az A mátrix **karakterisztikus polinom**ja: $k_A(\lambda) \stackrel{\text{def}}{=} |A - I_n\lambda|$.

Itt valós együtthatós, λ határozatlanú f polinomon egy $a_0 + a_1\lambda + a_2\lambda^2 + \dots + a_k\lambda^k$ alakú kifejezést értünk, ahol az a_i együtthatók valós számok, s az ilyen polinomok között a szokásos műveleteket és tulajdonságaikat ismertnek tételezzük fel. A határozatlan szó helyett változót mondani nem helyes, mert akkor keveredik a polinom és a polinomfüggvény fogalma. Most azonban ez elviselhető hiba, f helyett is $f(\lambda)$-t írunk, ez sem helyes, de még mindig jobb, mintha valaki az f-et számnak képzeli. Ha a legnagyobb kitevőjű λ-hatvány együtthatója, az $a_k \neq 0$, akkor azt mondjuk, hogy a polinom foka k: $\deg f(\lambda) = k$. A λ a lineáris algebrában szokásos jelölés, akit idegesít, nyugodtan használhat helyette pl. x-et.

A polinomokról az érdeklődők tájékozódhatnak KISS EMIL: BEVEZETÉS AZ ALGEBRÁBA c. könyve (TYPOTEX, 2007) 2. és 3. fejezetéből.
Most illik azonban jelezni, hogy a fenti definíciónál és a következő tétel bizonyítása során olyasmiket is felhasználunk, amiket korábban nem bizonyítottunk, csupán utaltunk rájuk a determinánsok elméletének legvégén (8/1).

TÉTEL: $A, B \in \mathbb{R}^{n\times n}$ és $A \sim_{\mathbb{R}} B$ esetén $k_A(\lambda) = k_B(\lambda)$.

[Biz: $k_B(\lambda) = |B - I_n\lambda| = |D^{-1}AD - I_n\lambda| = |D^{-1}(A - I_n\lambda)D| = |D^{-1}||A - I_n\lambda||D| = |A - I_n\lambda||D^{-1}||D| = |A - I_n\lambda| = k_A(\lambda)$.]

Legyen most $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, ennek karakterisztikus polinomja $\begin{vmatrix} 0-\lambda & -1 \\ 1 & 0-\lambda \end{vmatrix} = \lambda^2 + 1$, melynek nincs valós gyöke! Látszólag tehát megsemmisültek a 8/2 oldali reményeink a hatványozás egyszerűsítésére. Ez az a pont, ahol már nem szabad megelégednünk a valós számokkal, s bizony az összes eddigi vizsgálatot át kellene gondolnunk pl. a $\mathbb{C}$ esetére, vagy akár tetszőleges számtestre. (A $\mathbb{C}$-ben a fenti polinomnak van gyöke: i és $-i$.) Az algebra alaptétele szerint minden legalább elsőfokú komplex együtthatós polinomnak van gyöke a komplex számok körében, ebből gyöktényezős alak is nyerhető, de 4-nél nagyobb fokszám esetén általában csak közelítő eljárással határozhatók meg a gyökök.