2016. november 21/23.

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

Amikor egy valós szimmetrikus mátrixhoz SONB-t keresünk, ezt végezhetjük sajátalterenként (majd összerakva őket), mivel a különböző sajátalterek ortogonálisak:
TÉTEL: Legyen $A = A^\top \in \mathbb{R}^{n\times n}$. Ha λ és μ az A különböző sajátértékei, továbbá $\mathbf{x} \in W_\lambda,\ \mathbf{y} \in W_\mu$, akkor $\langle \mathbf{x}, \mathbf{y}\rangle = 0$.
[Bizonyítás: $\lambda\langle \mathbf{x}, \mathbf{y}\rangle = \langle \lambda\mathbf{x}, \mathbf{y}\rangle = \langle A\mathbf{x}, \mathbf{y}\rangle = (A\mathbf{x})^\top \mathbf{y} = \mathbf{x}^\top A^\top \mathbf{y} = \mathbf{x}^\top A\mathbf{y} = \langle \mathbf{x}, A\mathbf{y}\rangle = \langle \mathbf{x}, \mu\mathbf{y}\rangle = \mu\langle \mathbf{x}, \mathbf{y}\rangle$ miatt $(\lambda-\mu)\langle \mathbf{x}, \mathbf{y}\rangle = 0$, $\lambda - \mu \neq 0$ adja az állítást.]

Lineáris leképezések

Legyen most $A \in \mathbb{R}^{m\times n}$ és vizsgáljuk meg a következő leképezés kapcsolatát a vektortérműveletekkel: $\varphi : \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{x} \in \mathbb{R}^n$ esetén $\varphi(\mathbf{x}) = A\mathbf{x}$.
$\mathbf{x}, \mathbf{y} \in \mathbb{R}^n \Rightarrow \varphi(\mathbf{x} + \mathbf{y}) = A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \varphi(\mathbf{x}) + \varphi(\mathbf{y})$;
$\lambda \in \mathbb{R}, \mathbf{x} \in \mathbb{R}^n \Rightarrow \varphi(\lambda\mathbf{x}) = A(\lambda\mathbf{x}) = \lambda A\mathbf{x} = \lambda\varphi(\mathbf{x})$.
A példában összeg képe a képek összege volt, számszoros képe pedig a kép számszorosa.
DEFINÍCIÓ: Legyen $\varphi : \mathbb{R}^n \to \mathbb{R}^m$. φ **vektortérhomomorfizmus** [vagy *homogén lineáris leképezés*, vagy **lineáris leképezés**, vagy *műveletartó leképezés* $+$, $\lambda\cdot$*-ra*], ha
1./ $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n \Rightarrow \varphi(\mathbf{u} + \mathbf{v}) = \varphi(\mathbf{u}) + \varphi(\mathbf{v})$;
2./ $\lambda \in \mathbb{R}$, $\mathbf{u} \in \mathbb{R}^n \Rightarrow \varphi(\lambda\mathbf{u}) = \lambda\varphi(\mathbf{u})$.
Ha egy lineáris leképezés, azaz vektortérhomomorfizmus netán bijektív, akkor vektortérizomorfizmusnak hívjuk.
Egy lineáris leképezés bázison tetszőlegesen előírható, ez viszont már meg is határozza:
EGYÉRTELMŰ KITERJESZTÉSI TÉTEL: Legyen $\mathbf{e}_1, ..., \mathbf{e}_n$ bázis $\mathbb{R}^n$-ben; $\mathbf{w}_1, ..., \mathbf{w}_n$ tetszőleges vektorok $\mathbb{R}^m$-ben. Ekkor $\exists!\,\varphi : \mathbb{R}^n \to \mathbb{R}^m$ vektortérhomomorfizmus, melyre $\varphi(\mathbf{e}_i) = \mathbf{w}_i\ (i = 1, ..., n)$.
[Biz: !: Tegyük fel, hogy φ jó, s használjuk 1./ és 2./-t. Legyen $\mathbf{x} \in \mathbb{R}^n, \mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + ... + x_n\mathbf{e}_n$. Ekkor $\varphi(\mathbf{x}) = \varphi(x_1\mathbf{e}_1 + (x_2\mathbf{e}_2 + ... + x_n\mathbf{e}_n)) = \varphi(x_1\mathbf{e}_1) + \varphi(x_2\mathbf{e}_2 + ... + x_n\mathbf{e}_n) = ... = \varphi(x_1\mathbf{e}_1) + \varphi(x_2\mathbf{e}_2) + ... + \varphi(x_n\mathbf{e}_n) = x_1\varphi(\mathbf{e}_1) + x_2\varphi(\mathbf{e}_2) + ... + x_n\varphi(\mathbf{e}_n) = x_1\mathbf{w}_1 + x_2\mathbf{w}_2 + ... + x_n\mathbf{w}_n$. $\exists$: Legyen $\mathbf{x} \in \mathbb{R}^n, \mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + ... + x_n\mathbf{e}_n$. Ekkor $\varphi(\mathbf{x}) \stackrel{\text{def}}{=} x_1\mathbf{w}_1 + x_2\mathbf{w}_2 + ... + x_n\mathbf{w}_n$ megfelelő lesz.]
A bevezető példához hasonlóan mátrixokat is használhatunk a vektortérhomomorfizmusok megadására [a baloperátoros írásmódhoz a következő definíció illik]:
DEFINÍCIÓ: Ha $\mathbf{e}_1, ..., \mathbf{e}_n$ bázis $\mathbb{R}^n$-ben; $\mathbf{f}_1, ..., \mathbf{f}_m$ bázis $\mathbb{R}^m$-ben; $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ vektortérhomomorfizmus, akkor a φ mátrixa az $\mathbf{e}; \mathbf{f}$ bázispárban
$[\varphi]^{\mathbf{e};\mathbf{f}} \stackrel{\text{def}}{=} [[\varphi(\mathbf{e}_1)]_\mathbf{f}, ..., [\varphi(\mathbf{e}_n)]_\mathbf{f}] \in \mathbb{R}^{m\times n}$.
A definícióból adódik, hogy tetszőleges $\mathbf{x} \in \mathbb{R}^n$ esetén $[\varphi(\mathbf{x})]_\mathbf{f} = [\varphi]^{\mathbf{e};\mathbf{f}}[\mathbf{x}]_\mathbf{e}$.
DEFINÍCIÓ: $\mathcal{H}om(\mathbb{R}^n, \mathbb{R}^m)$ jelölje a $\mathbb{R}^n$-ből $\mathbb{R}^m$-be képező vektortérhomomorfizmusok halmazát.
$\varphi, \psi \in \mathcal{H}om(\mathbb{R}^n, \mathbb{R}^m)$ esetén legyen $\varphi + \psi : \mathbb{R}^n \to \mathbb{R}^m$ úgy, hogy $\mathbf{u} \in \mathbb{R}^n$-re $(\varphi + \psi)(\mathbf{u}) = \varphi(\mathbf{u}) + \psi(\mathbf{u})$. $\lambda \in \mathbb{R}$ és $\varphi \in \mathcal{H}om(\mathbb{R}^n, \mathbb{R}^m)$ esetén legyen $\lambda\varphi : \mathbb{R}^n \to \mathbb{R}^m$ úgy, hogy $\mathbf{u} \in \mathbb{R}^n$-re $(\lambda\varphi)(\mathbf{u}) = \lambda(\varphi(\mathbf{u}))$.
Könnyen látható, hogy az imént definiált összeg és számszoros is $\mathcal{H}om(\mathbb{R}^n, \mathbb{R}^m)$-ben van, továbbá az is, hogy $\mathcal{H}om(\mathbb{R}^n, \mathbb{R}^m)$ vektortér az $\mathbb{R}$ felett a fenti műveletekre.
Vektortérhomomorfizmusok szorzatát is könnyen definiálhatjuk kompozícióként, csak a baloperátoros írásmód miatt a sorrendre figyelni kell, ezt az indexek jobban mutatják, mint a "kitevők":

DEFINÍCIÓ: Legyen $V_1 = \mathbb{R}^n$, $V_2 = \mathbb{R}^m$ és $V_3 = \mathbb{R}^s$; $\varphi \in \mathcal{H}om(V_1, V_2)$, $\psi \in \mathcal{H}om(V_2, V_3)$.
Legyen $\psi\varphi : V_1 \to V_3$ úgy, hogy $\mathbf{u} \in V_1$-re $(\psi\varphi)(\mathbf{u}) = \psi(\varphi(\mathbf{u}))$.
Könnyen látható, hogy a definiált $\psi\varphi \in \mathcal{H}om(V_1, V_3)$.

Tapasztaltuk már, hogy az azonos dimenziójú vektorterekkel ugyanúgy tudtunk dolgozni, a magyarázat: van köztük vektortérizomorfizmus (ilyenkor a két vektortér közé $\cong$ jelet teszünk). Bizonyítani fogjuk, hogy $\mathbb{R}^n \cong \mathbb{R}^m \Leftrightarrow \dim \mathbb{R}^n = \dim \mathbb{R}^m$. Előbb még olyan fogalmakat definiálunk, melyek "mérni" tudják, hogy pl. egy $\varphi \in \mathcal{H}om(\mathbb{R}^n, \mathbb{R}^m)$ mennyire nem izomorfizmus.
DEFINÍCIÓ: Legyen $\varphi \in \mathcal{H}om(\mathbb{R}^n, \mathbb{R}^m)$.
φ **képtere**: $\mathcal{I}m\,\varphi \stackrel{\text{def}}{=} \{\mathbf{v}' \mid \mathbf{v}' \in \mathbb{R}^m\ \exists \mathbf{u} \in \mathbb{R}^n\ \varphi(\mathbf{u}) = \mathbf{v}'\}$;
φ **magtere**: $\mathcal{K}er\,\varphi \stackrel{\text{def}}{=} \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\ \varphi(\mathbf{x}) = \mathbf{0}'\}$.
Jegyezzük meg, hogy $\varphi(\mathbf{0}) = \varphi(0\mathbf{0}) = 0\varphi(\mathbf{0}) = \mathbf{0}'$. Könnyen igazolható, hogy $\mathcal{I}m\,\varphi \leqq \mathbb{R}^m$ és $\mathcal{K}er\,\varphi \leqq \mathbb{R}^n$. Nyilvánvaló, hogy $\mathcal{I}m\,\varphi = \mathbb{R}^m \Leftrightarrow \varphi$ szürjektív. Mivel $\varphi(\mathbf{x}) = \varphi(\mathbf{y}) \Leftrightarrow \varphi(\mathbf{x}-\mathbf{y}) = \mathbf{0}' \Leftrightarrow \mathbf{x}-\mathbf{y} \in \mathcal{K}er\,\varphi$, ezért az is igaz, hogy φ injektív $\Leftrightarrow \mathcal{K}er\,\varphi = \{\mathbf{0}\}$. Később igazolni fogjuk, hogy érvényes a következő **dimenzióösszefüggés**:
$\dim \mathcal{I}m\,\varphi + \dim \mathcal{K}er\,\varphi = \dim \mathbb{R}^n$.
[Itt φ rangja: $r(\varphi) = \dim \mathcal{I}m\,\varphi$; φ defektusa: $d(\varphi) = \dim \mathcal{K}er\,\varphi$.] Foglalkoztunk már $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ lineáris leképezésekkel. Különösen fontosak lesznek azok a lineáris leképezések, amelyeknél $\mathbb{R}^n = \mathbb{R}^m$, ezek az $\mathbb{R}^n$ vektortér LINEÁRIS TRANSZFORMÁCIÓi. Ilyenkor megállapodunk abban, hogy a mátrix definíciójában mindkét helyre ugyanazt a bázist vesszük: Ha $\mathbf{e}_1, ..., \mathbf{e}_n$ bázis $\mathbb{R}^n$-ben, $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ lineáris transzformáció, akkor $[\varphi]^{\mathbf{e}} \stackrel{\text{def}}{=} [\varphi]^{\mathbf{e};\mathbf{e}}$.
Nemsokára majd bebizonyítjuk, hogy egy $\mathbb{R}$ feletti vektortér lineáris transzformációjának különböző bázisokban vett mátrixai hasonlók $\mathbb{R}$ felett. Most vizsgáljuk meg, mennyire nagy szerencse egy n dimenziós vektortér φ lineáris transzformációjához olyan bázist találni, amelyben a mátrixa diagonális: az $\mathbf{e}'_1, ..., \mathbf{e}'_n$ bázisnak azt kell tudnia, hogy $\varphi(\mathbf{e}'_i) = \lambda_i \mathbf{e}'_i$, azaz minden báziselem saját számszorosába képeződjön. Persze, $\varphi(\mathbf{0}) = \mathbf{0}$, de a $\mathbf{0}$ nem lehet báziselem! Ha síkon $(n = 2)$ a kezdőpont körüli $+90°$-os forgatást nézzük, erről belátható, hogy lineáris transzformáció, de egyetlen nullvektortól különböző vektort sem visz a saját számszorosába, nemhogy egy bázist. Először tehát azt kellene tisztázni, mikor van egyáltalán jó vektor (kis szerencse), aztán ilyenekből bázis (nagy szerencse).
DEFINÍCIÓ: Legyen $\varphi \in \mathcal{H}om(\mathbb{R}^n, \mathbb{R}^n)$.
Az $\mathbf{u} \in \mathbb{R}^n$ **sajátvektora** φ-nek, ha
1./ $\mathbf{u} \neq \mathbf{0}$,
2./ $\exists\, \lambda_0 \in \mathbb{R} : \varphi(\mathbf{u}) = \lambda_0 \mathbf{u}$.
Ilyenkor a λ_0 az $\mathbf{u}$ sajátvektorhoz tartozó sajátértéke a φ-nek.
Legyen $\mathbf{e}_1, ..., \mathbf{e}_n$ bázis $\mathbb{R}^n$-ben; $\varphi \in \mathcal{H}om(\mathbb{R}^n, \mathbb{R}^n)$; $A = [\varphi]^{\mathbf{e}}$. $\lambda_0 \in \mathbb{R}$ esetén
λ_0 sajátértéke φ-nek $\Leftrightarrow \exists\, \mathbf{u} \in \mathbb{R}^n$, $\mathbf{u} \neq \mathbf{0} : \varphi(\mathbf{u}) = \lambda_0 \mathbf{u} \Leftrightarrow$
$\exists\, \mathbf{u} \in \mathbb{R}^n$, $\mathbf{u} \neq \mathbf{0} : [\varphi(\mathbf{u})]_{\mathbf{e}} = \lambda_0 [\mathbf{u}]_{\mathbf{e}} \Leftrightarrow \exists\, \mathbf{u} \in \mathbb{R}^n$, $\mathbf{u} \neq \mathbf{0} : [\varphi]^{\mathbf{e}} [\mathbf{u}]_{\mathbf{e}} = \lambda_0 [\mathbf{u}]_{\mathbf{e}} \Leftrightarrow$
$\exists\, \mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0} : A\mathbf{x} = \lambda_0 \mathbf{x} \Leftrightarrow \lambda_0$ jobb oldali sajátértéke A-nak $\Leftrightarrow |A - I_n \lambda_0| = 0$.

Próbálkozhatunk φ karakterisztikus polinomjának definíciójával is: $k_\varphi(\lambda) \stackrel{\text{def}}{=} k_{[\varphi]^{\mathbf{e}}}(\lambda)$, de felmerül a kérdés, hogy értelmes-e (nem függ-e a bázis választásától)? Szerencsére nincs gond, hisz φ különböző bázisokban vett mátrixai $\mathbb{R}$ felett hasonlók, azt pedig már bizonyítottuk, hogy hasonló mátrixok karakterisztikus polinomja megegyezik.
Következőnek elégséges feltételt adunk majd SB létezésére a különböző sajátértékű sajátvektorok függetlenségének igazolása révén, előbb azonban mátrixok partícionálásával kap-

csolatban teszünk egy eddig halogatott kiegészítést.

Komoly mese a mátrixszorzás műveletigényéről

Az $n\times n$-es mátrixok szorzásánál a nagyságrendileg pontos műveletigény meghatározása jelenleg reménytelen. Az n kitevőjében a triviális 3 értéket 1969-ben V. Strassen [*Numer. Math.* **13** (1969), 354–356] 2,81-re csökkentette, ugyanő 1986-ban [*J. reine angew. Math.* **375/376** (1987), 406–443] 2,479-re, majd ezt 1990-ben D. Coppersmith és S. Winograd [*J. Symbolic Comput.* **9** (1990), 251–280] 2,376-ra, 2011-ben D. V. Zhdanovich [*J. Math. Sci. (N. Y.)* **186** (2012), 599–643] 2,373-ra javította. Messze van még a sejtett 2.

Most vázoljuk a $cn^{2,81}$-es becslés alapgondolatát (a 2,81 a $\log_2 7$ közelítő értéke, $3 = \log_2 8$). Legyen $n = 2^k$ alakú, A, B $n\times n$-es mátrixok, $C = AB$. Az A, B mátrixokat partícionáljuk a következőképpen:

$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, $B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$, ahol a blokkok mind $n/2\times n/2$-es mátrixok.

Ekkor $C = AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}$.

A blokkokból számítsuk ki a következő **hét** mátrixot:

$Q_1 = (A_{11} + A_{22})(B_{11} + B_{22})$,
$Q_2 = (A_{21} + A_{22})B_{11}$,
$Q_3 = A_{11}(B_{12} - B_{22})$,
$Q_4 = A_{22}(B_{21} - B_{11})$,
$Q_5 = (A_{11} + A_{12})B_{22}$,
$Q_6 = (A_{21} - A_{11})(B_{11} + B_{12})$,
$Q_7 = (A_{12} - A_{22})(B_{21} + B_{22})$.

Könnyen verifikálható, hogy ekkor

$$C = AB = \begin{bmatrix} Q_1 + Q_4 - Q_5 + Q_7 & Q_3 + Q_5 \\ Q_2 + Q_4 & Q_1 + Q_3 - Q_2 + Q_6 \end{bmatrix}.$$

Ezzel az eredeti *8 darab mátrixszorzás* és 4 darab mátrixösszeadás helyett **7 darab mátrixszorzás** és 18 darab mátrixösszeadás/kivonás kell.

Az igényelt skalárszorzások $M(n)$ számára a triviális becslés a mátrixszozás definíciójából $M(n) \le n^3$, a skalárösszeadások/kivonások $S(n)$ számára pedig $S(n) \le n^2(n-1)$, tehát $M(n) + S(n) \le 2n^3 - n^2$. Most az eredeti $M(n) \le 8M(n/2)$, $S(n) \le 4S(n/2)$ helyett

$$M(n) \le 7M(n/2), \quad S(n) \le 7S(n/2) + 18(n/2)^2, \text{azaz } S(n) + 6n^2 \le 7\left(S(n/2) + 6(n/2)^2\right)$$

adódik, végül $M(1) = 1, S(1) = 0$ figyelembevételével [$n = 2^k$ esetére]

$$M(n) \le 7^k = 2^{k\log_2 7} = n^{\log_2 7}, \quad S(n) \le 6\cdot 7^k - 6\cdot 4^k, \quad M(n) + S(n) \le 7n^{\log_2 7}.$$

Amikor az n nem 2-hatvány, a mátrixokat kiegészíthetjük a következő 2-hatvány méretig a főátlóban 1-gyel, a többi helyen 0-val, így az általános esetben $M(n)+S(n) \le 7(2n)^{\log_2 7}$ $= 49n^{\log_2 7}$ megfelelő.