2016. október 3/5.

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

A 3. előadás legvégén szereplő mátrixszorzás definíciójánál az összeszorozhatóság méretfeltételei alapján persze nem számíthatunk kommutatív műveletre: lehet, hogy AB értelmezve van (A oszlopainak száma megegyezik B sorainak számával), de BA nincs értelmezve. De ez még nem a fő baj: $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times m}$ esetén létezik AB is, BA is, csakhogy az első $m \times m$-es, a második $n \times n$-es, tehát $m \neq n$ esetén biztosan nem lehetnek egyenlők. Az igazi gond az, hogy általában még $m = n$ esetén sem lesznek azonosak, pedig az alak már rendben volna. Az $m = n = 2$ esetre nézzük a következő példát: Legyen $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Ekkor $A, B \in \mathbb{R}^{2\times 2}$, tehát értelmezve van AB is, BA is, mindkettő 2×2-es, de $AB = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $BA = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, tehát $AB \neq BA$. A fenti példában $B = AB \neq BA = \mathbf{0}$, egyúttal tehát azt is látjuk, hogy két nullmátrixtól különböző mátrix szorzata adhat nullmátrixot, de nem biztos, hogy más sorrendben is ezt ad. Ezek után talán meglepő lehet, hogy a fenti sor–oszlop szorzás asszociatív és az összeadással disztributív kapcsolatban áll, majd nemsokára pontosan meg is fogalmazzuk ezeket. [[Aki már találkozott más lineáris algebrai felépítéssel, esetleg a sor–sor, oszlop–oszlop, oszlop–sor szorzást is meg tudja fogalmazni, de csalódni fog: ezek egyike sem kommutatív, sőt még csak nem is asszociatív.]] Később mátrixokat fogunk használni bizonyos transzformációk leírására. Aki ezt meglepőnek találja, megnézheti a manapság mindenhol használt Postscript nyelv reference manuálját:
http://www.adobe.com/products/postscript/pdfs/PLRM.pdf
A 187–188. oldalon kiderül, hogy mátrixokkal adja meg azokat a koordinátatranszformációkat, amiket használ (például, amikor a lapon egy ábrát forgatni, tükrözni, egy betűt dönteni kell).
Beígértük már a mátrixszorzás asszociativitását és a disztributivitásokat, de a pontos kimondás előtt célszerű egyszerűbb dolgokon gyakorolni a mátrixszorzást.

DEFINÍCIÓ: $I_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \in \mathbb{R}^{n\times n}$ az $n \times n$-es **egységmátrix**, $\quad {}_i[I_n]_j = \delta_{ij} = \begin{cases} 1, & \text{ha } i = j; \\ 0, & \text{ha } i \neq j. \end{cases}$ (A δ_{ij} egyik szokásos elnevezése: Kronecker-szimbólum.)

Az egységmátrix elnevezést az indokolja, hogy az I_n a szorzásnál minden olyan mátrixot változatlanul hagy, amellyel az adott sorrendben össze lehet szorozni:
TÉTEL: $A \in \mathbb{R}^{m\times n}$ esetén $I_m A = A$ és $AI_n = A$.
[Az első állítás igazolásánál először vegyük észre, hogy $I_m A$ értelmezve van, és $m \times n$-es, mint az A; ezután nézzük minden szóbajövő i, j esetére az i-edik sor j-edik elemét:

$${}_i[I_m A]_j = \sum_{\ell=1}^{m} {}_i[I_m]_\ell \, {}_\ell[A]_j = \sum_{\ell=1}^{m} \delta_{i\ell} a_{\ell j} = \delta_{ii} a_{ij} = a_{ij} = {}_i[A]_j.$$

A másik állítás hasonlóan igazolható.]
A szorzás kommutativitásának hiányában igen fontos szerepe lesz a transzponálásnak:

DEFINÍCIÓ: $A \in \mathbb{R}^{m\times n}$ esetén az A mátrix **transzponáltja**: $A^\top \in \mathbb{R}^{n\times m}$, melyre minden szóbajövő i, j-re ${}_i[A^\top]_j = {}_j[A]_i$.
TÉTEL (A transzponálás kapcsolata az eddigi műveletekkel):
$A, B \in \mathbb{R}^{m\times n} \Rightarrow (A+B)^\top = A^\top + B^\top$;
$\lambda \in \mathbb{R}$, $A \in \mathbb{R}^{m\times n} \Rightarrow (\lambda A)^\top = \lambda A^\top$;
$A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times k} \Rightarrow (AB)^\top = B^\top A^\top$.
[Csak az utolsó állítás bizonyítása szorul részletezésre: a létezés és az alakok egyezése rögtön látszik (épp fordítva volna zűrös), továbbá minden szóbajövő i, j-re

$$ {}_i[(AB)^\top]_j = {}_j[AB]_i = \sum_{\ell=1}^{n} {}_j[A]_\ell\, {}_\ell[B]_i = $$

$$ \sum_{\ell=1}^{n} {}_\ell[A^\top]_j\, {}_i[B^\top]_\ell = \sum_{\ell=1}^{n} {}_i[B^\top]_\ell\, {}_\ell[A^\top]_j = {}_i[B^\top A^\top]_j, $$

hisz számok szorzásának sorrendjét szabad cserélnünk.]

A mátrixszorzás asszociativitásának vizsgálatánál azt is tisztáznunk kell, hogy melyik oldal mikor létezik!
TÉTEL: $A \in \mathbb{R}^{m\times n_1}$, $B \in \mathbb{R}^{n_2\times k_2}$, $C \in \mathbb{R}^{k_3\times s}$ esetén

$$ \exists\,(AB)C \iff \underbrace{\{n_1 = n_2 \text{ és } k_2 = k_3\}}_{\Downarrow} \iff \exists\, A(BC) $$
$$ (AB)C = A(BC). $$

[A létezésekkel kapcsolatos állítások a mátrixszorzás definíciójából adódnak.
$\Downarrow$: $n_1 = n_2 (= n)$ és $k_2 = k_3 (= k)$ esetén $(AB)C$, $A(BC) \in \mathbb{R}^{m\times s}$, továbbá minden szóbajövő i, j-re

$$ {}_i[(AB)C]_j = \sum_{p=1}^{k} {}_i[AB]_p\, {}_p[C]_j = \sum_{p=1}^{k} \left(\sum_{q=1}^{n} {}_i[A]_q\, {}_q[B]_p \right) {}_p[C]_j = \sum_{q=1}^{n} \sum_{p=1}^{k} {}_i[A]_q\, {}_q[B]_p\, {}_p[C]_j $$

$$ = \sum_{q=1}^{n} {}_i[A]_q \left(\sum_{p=1}^{k} {}_q[B]_p\, {}_p[C]_j \right) = \sum_{q=1}^{n} {}_i[A]_q\, {}_q[BC]_j = {}_i[A(BC)]_j, $$

tehát $(AB)C = A(BC)$.]

TÉTEL: $A \in \mathbb{R}^{m\times n_1}$, $B \in \mathbb{R}^{n_2\times k_2}$, $C \in \mathbb{R}^{n_3\times k_3}$ esetén

$$ \exists\, A(B+C) \iff \underbrace{\{n_1 = n_2 = n_3 \text{ és } k_2 = k_3\}}_{\Downarrow} \iff \exists\, AB + AC $$
$$ A(B+C) = AB + AC. $$

[A bizonyítás az előbbi minta alapján nem okozhat gondot.]
HF: A másik disztributivitás kimondása.
TÉTEL: $\lambda \in \mathbb{R}$, $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times k} \Rightarrow \lambda(AB) = (\lambda A)B = A(\lambda B)$.
[A bizonyítás triviális.]

Most tekintsünk egy $A \in \mathbb{R}^{m\times n}$ mátrixot, s osszuk fel néhány részre vizszintesen is, függőlegesen is, pl. vizszintesen két részre, függőlegesen három részre, ha $m \geqq 2$, $n \geqq 3$:

$$ A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix} $$

Azt mondjuk, hogy így az A mátrixot partícionáltuk, vagy blokkokra bontottuk, ahol a blokkok maguk is mátrixok, mégpedig olyan méretűek, hogy az egy sorban elhelyezkedő blokkok ugyanannyi sorból állnak, az egy oszlopban lévők pedig ugyanannyi oszlopot tartalmaznak. Blokkokra bontott mátrixok közötti műveletek sokszor áttekinthetőbben végezhetők el. Ehhez pl. a mátrixösszeadásnál nyilván azonos módon kell partícionálni az összeadandó mátrixokat, sokkal fontosabb azonban a mátrixszorzás áttekinthetősége: itt az lehet a cél, hogy a blokkokra bontott mátrixokat esetleg úgy szorozhassuk össze, mintha számmátrixok volnának (a szorzás során azonban a tényezők sorrendjét tilos változtatni, hisz a mátrixszorzás nem kommutatív). Ez a cél el is érhető, mindössze azt kell biztosítanunk, hogy a megfelelő blokkok összeszorozhatók legyenek. Legyen most $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times k}$, és tekintsük a következő partícionálásukat:

$$\begin{array}{ccc} \wedge & \sqcap & \Yup \\ \multicolumn{3}{c}{\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}} \end{array} \qquad \begin{array}{c} < \\ \sqsubset \\ \prec \end{array} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \\ B_{31} & B_{32} \end{bmatrix}$$

A fenti példában a hasonló jellel jelzett méretek megegyezése (az összeszorozhatóságnál szereplő n sorrendben is azonos felbontása pozitív egészek összegére) elég ahhoz, hogy a szorzást úgy végezhessük, mintha számmátrixokról lenne szó, s a szorzatmátrixnak négy blokkja lesz: $\begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} + A_{13}B_{31} & A_{11}B_{12} + A_{12}B_{22} + A_{13}B_{32} \\ A_{21}B_{11} + A_{22}B_{21} + A_{23}B_{31} & A_{21}B_{12} + A_{22}B_{22} + A_{23}B_{32} \end{bmatrix}$. A fentiek jogosságát általában nem bizonyítjuk, de szabad használni a mátrixszorzás áttekinthetőbbé tételére, s ha valakinek kétsége támadna, egy konkrét esetet mindig könnyű ellenőrizni.
Most az $\mathbb{R}^m$ triviális bázisát használva az $\mathbf{a}_1, ..., \mathbf{a}_n \in \mathbb{R}^m$ definíciójában (3/2), megkapjuk az $A = \begin{bmatrix} a_{11} & ... & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & ... & a_{mn} \end{bmatrix}$ mátrix oszlopvektorokra való partícionálását:
$A = [\mathbf{a}_1, ..., \mathbf{a}_n] \in \mathbb{R}^{m\times n}$, melyre tetszőleges $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ esetén teljesül, hogy $A\mathbf{x} = \mathbf{a}_1 x_1 + ... + \mathbf{a}_n x_n$, így leolvasható $A\mathbf{x} \in \text{Span}\,(\mathbf{a}_1, ..., \mathbf{a}_n)$ is. Az $A\mathbf{x} = \mathbf{b}$ LER esetén az elemi bázistranszformáció segítségével már el tudtuk dönteni a megoldhatóságot, meg tudtuk határozni a megoldásszámot, s megoldhatóság esetén a táblázat segítségével meg is tudtunk adni egy megoldást. Ha azonban végtelen sok megoldás van, ezek áttekintése még nem történt meg. A korábbi jelölésekkel (3/3) – mivel az $r = 0$ eset triviális –, azzal a feltevéssel dolgozunk, hogy van megoldás és $0 < r < n$. INDEXCSERÉVEL (azaz az A oszlopai sorrendjének s egyúttal az ismeretlenek sorrendjének megváltoztatásával) elérhetjük, hogy a bevitt vektorok $\mathbf{a}_1, ..., \mathbf{a}_r$ legyenek [a végső formula után persze vissza kell cserélni].

$$\begin{array}{c|ccc|ccc|c|} & \mathbf{a}_1 & ... & \mathbf{a}_r & \mathbf{a}_{r+1} & ... & \mathbf{a}_n & \mathbf{b} \\ \hline \mathbf{a}_1 & & & & & & & \\ \vdots & & I_r & & & D & & \mathbf{d} \\ \mathbf{a}_r & & & & & & & \\ \hline \mathbf{e}'_{r+1} & & & & & & & \\ \vdots & & \mathbf{0} & & & \mathbf{0} & & \mathbf{0} \\ \mathbf{e}'_m & & & & & & & \end{array}$$

Legyen most $A_1 = [\mathbf{a}_1, ..., \mathbf{a}_r]$, $A_2 = [I_r, D]$. Ekkor könnyen ellenőrizhető a táblázat alapján, hogy $A_1\mathbf{d} = \mathbf{b}$ és $A = A_1A_2$. Így $\mathbf{x} \in \mathbb{R}^n$ esetén
$A\mathbf{x} = \mathbf{b} \Longleftrightarrow A_1A_2\mathbf{x} = A_1\mathbf{d} \Longleftrightarrow A_1(A_2\mathbf{x} - \mathbf{d}) = \mathbf{0} \Longleftrightarrow A_2\mathbf{x} - \mathbf{d} = \mathbf{0}$
teljesül, ahol ki kell emelnünk, hogy az utolsó ekvivalencia $\Rightarrow$ iránya egyáltalán nem nyilvánvaló: nem elég hozzá, hogy $A_1 \neq \mathbf{0}$, mert nullmátrixtól különböző mátrixok szorzata lehet nullmátrix! A $\Rightarrow$ abból adódik, hogy az A_1 oszlopvektorai, $\mathbf{a}_1, ..., \mathbf{a}_r$ lineárisan független vektorrendszert alkotnak! A folytatásban a partícionálás alkalmazásához célszerű az $\mathbf{x}$ megfelelő partícionálása, ahol az indexekkel jelezzük a méretet. Ekkor azonban, ha netán $r = n - r$, akkor jogtalan specializálás lenne, ha a fölső és az alsó rész eleve azonos volna. Ezt a zavart vesszőkkel kerüljük el: `Tetszőleges` $\mathbf{x} = \begin{bmatrix} \mathbf{x}'_r \\ \mathbf{x}''_{n-r} \end{bmatrix} \in \mathbb{R}^n$`-re`

$$A\mathbf{x} = \mathbf{b} \Longleftrightarrow A_2\mathbf{x} = \mathbf{d} \Longleftrightarrow [I_r, D] \begin{bmatrix} \mathbf{x}'_r \\ \mathbf{x}''_{n-r} \end{bmatrix} = \mathbf{d} \Longleftrightarrow \mathbf{x}'_r = \mathbf{d} - D\mathbf{x}''_{n-r},$$

`amivel megkaptuk az általános megoldást:` $\mathbf{x}''_{n-r}$ `komponensei a szabad` (vagy független) `változók`, $\mathbf{x}'_r$ `komponensei a kötött` (vagy függő) `változók`. Mivel az elemi bázistranszformáció sokféleképpen végezhető, ezek helye nem egyértelmű, de darabszámuk egyértelmű. [Most kellene még visszacserélni az indexeket az eredeti lineáris egyenletrendszer általános megoldásának felírásához.]
Az oszlopvektorrendszer rangjának segítségével definiálhatjuk tetszőleges mátrix oszloprangját, majd sorrangját:
DEFINÍCIÓ: $A = [\mathbf{a}_1, ..., \mathbf{a}_n] \in \mathbb{R}^{m\times n}$ **oszloprangja**: $\varrho_{\mathcal{O}}(A) = r(\mathbf{a}_1, ..., \mathbf{a}_n)$ [emlékeztető: $=$ dim Span $(\mathbf{a}_1, ..., \mathbf{a}_n)$]; **sorrangja** `pedig` $\varrho_{\mathcal{S}}(A) = \varrho_{\mathcal{O}}(A^\top)$.
Rövidesen be fogjuk bizonyítani, hogy tetszőleges A mátrix oszloprangja és sorrangja megegyezik. Ennek egyik segédeszköze (s egyúttal a partícionálás természetes alkalmazása) szorzat oszloprangjának vizsgálata:
TÉTEL: `Legyenek` $C = [\mathbf{c}_1, ..., \mathbf{c}_n]$ `és` $D = [\mathbf{d}_1, ..., \mathbf{d}_k]$ `ebben a sorrendben összeszorozható` $\mathbb{R}$ `feletti mátrixok. Ekkor` $\varrho_{\mathcal{O}}(CD) \leqq \varrho_{\mathcal{O}}(C)$.
[A bizonyítás egyetlen ötlete, hogy először csak a D partícionálását írjuk be: $CD = C[\mathbf{d}_1, ..., \mathbf{d}_k] = [C\mathbf{d}_1, ..., C\mathbf{d}_k]$. Mivel $C\mathbf{d}_i \in$ Span $(\mathbf{c}_1, ..., \mathbf{c}_n)$ $(i = 1, ..., k)$, ezért teljesül az is, hogy Span $(C\mathbf{d}_1, ..., C\mathbf{d}_k) \subseteqq$ Span $(\mathbf{c}_1, ..., \mathbf{c}_n)$, az utóbbiak dimenzióira vonatkozó egyenlőtlenség pedig már a bizonyítandó állítás.]
TÉTEL: `Tetszőleges` $\mathbb{R}$ `feletti` A `mátrixra` $\varrho_{\mathcal{O}}(A) = \varrho_{\mathcal{S}}(A)$ [ezentúl $= \varrho(A)$, `az` A **`rangja`**; a ϱ helyett használatos ρ vagy akár r is].
[Bizonyítás: Az állítás nyilvánvaló, ha $A = \mathbf{0}$. $A \in \mathbb{R}^{m\times n}$, $A \neq \mathbf{0}$ esetén használni szeretnénk a korábbi (4/4) elemi bázistranszformációs számolásokat. Ehhez meg kell említeni, hogy az ottani indexcserétől (oszlopcserétől) nem változik sem az oszloprang, sem a sorrang; továbbá az $r = n$ esetet is nézni kell, ekkor legyen $A_2 = I_n$. Így $0 < r \leqq n$ esetén is teljesül az $A = A_1A_2$, és $\varrho_{\mathcal{S}}(A) = \varrho_{\mathcal{O}}((A_1A_2)^\top) = \varrho_{\mathcal{O}}(A_2^\top A_1^\top) \leqq \varrho_{\mathcal{O}}(A_2^\top) = r = \varrho_{\mathcal{O}}(A)$. Miután tetszőleges mátrixra beláttuk a $\varrho_{\mathcal{S}}(A) \leqq \varrho_{\mathcal{O}}(A)$ egyenlőtlenséget, most felírhatjuk ugyanezt az $A^\top$ mátrixra is, ami az A-ra adja a hiányzó ellenkező irányú egyenlőtlenséget.]

Ezután az $A\mathbf{x} = \mathbf{b}$ általánosításaként $AX = B$ és $YC = D$ alakú mátrixegyenletekkel foglalkozunk (adott A, B illetve C, D mátrixok esetén), melynek motivációja lehet például a korábbi problémás eset, amikor jó lett volna az A_1-gyel olyan alapon „egyszerűsíteni”, hogy valami alkalmas mátrixszal balról szorzunk. $YC = D \Longleftrightarrow C^\top Y^\top = D^\top$ miatt elég az $AX = B$-vel foglalkoznunk természetes méretfeltétel mellett: $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{m\times k}$.