IDAG: Matrisop. som linjära avbildningar

Blockmatriser

LU-faktorisering

Mål: Bättre förståelse av $A+B$, AB, A^{-1}

Förenkla räkningar genom blockuppdelning

Se fördelar med faktoriseringar, bestämma $A=LU$.

Geometrisk tolkning av matrisoperationer

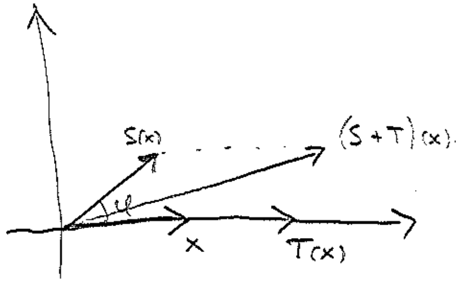
Vi införde $A+B$, AB och A^{-1} utan mycket mer motivering än att det är "naturligt". Via tolkningen som linjär avbildning kan vi dock ge en tydlig geometrisk motivering:

Låt de lin. avb. S och T ges av matriserna A och B.

Då ges standardmatrisen för

Summan $S+T$ av $A+B$

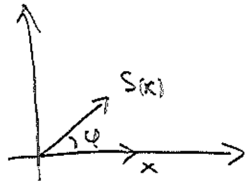
Ex.



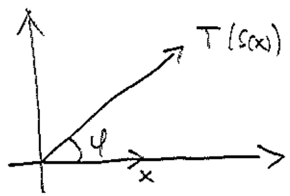
Rotation (S) plus skalning (T).

Kompositionen TS av BA

$$x \xrightarrow{S} S(x) \xrightarrow{T} T(S(x))$$

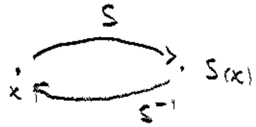


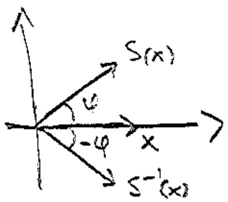
Rotation



Rotation och sen skalning

Inversen $S^{-1}: Sx \mapsto x$ av A^{-1}





Rotation i omvänd riktning.

Inversen av en lin. avb. existerar inte alltid.

Därför har vi formellt

<u>Def.</u> En lin. avb. $T: \mathbb{R}^n \to \mathbb{R}^n$ är <u>inverterbar</u> om det finns en lin. avb. $S: \mathbb{R}^n \to \mathbb{R}^n$ sådan att

$$S(T(x)) = x \quad \text{och} \quad T(S(x)) \text{ för alla } x \in \mathbb{R}^n.$$

Vi kallar S <u>inversen</u> till T och skriver den som T^{-1}.

<u>Sats 2.3.9</u> Låt $T: \mathbb{R}^n \to \mathbb{R}^n$ ha standardmatrisen A. Då är T inv. omm A är inv. I så fall har T^{-1} standardmatrisen A^{-1}.

<u>Bevis</u>: Lay.

<u>Ex.</u> Låt $T: \mathbb{R}^n \to \mathbb{R}^n$ vara en <u>surjektiv</u> lin. avb. med st. matris A. Då ger sats 2.3.8 att A är inv. $(A \sim I_n)$ så T är inv.

Detta betyder också att T är injektiv.

<u>Ex. forts.</u>

Likaså får vi att T <u>injektiv</u> $\Rightarrow$ T inv. $\Rightarrow$ T surjektiv.

Dvs. en lin. avb. <u>från $\mathbb{R}^n$ till $\mathbb{R}^n$</u> (viktigt) är antingen både injektiv och surjektiv eller inget dera.

Blockmatriser

Om vi har två matriser $A = (a_{ij})$ och $B = (b_{ij})$ med samma antal rader kan vi bilda en större matris $[A\ B]$. På samma sätt kan vi bilda $\begin{bmatrix} A \\ B \end{bmatrix}$ eller $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ om dimensionerna passar ihop.

Vi kan såklart även göra tvärtom:

<u>Ex.</u> Låt $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 4 \end{bmatrix}$. Då kan vi skriva

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \text{ där } \begin{array}{l} A_{11} = I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ A_{12} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ A_{21} = 0_{1\times 2} = [0\ 0], \ A_{22} = [4]. \end{array}$$

Poängen med detta är att vi kan förenkla matrisräkningar. Om

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$ med samma blockindelning

så är

$$A + B = \begin{bmatrix} A_{11}+B_{11} & A_{12}+B_{12} \\ A_{21}+B_{21} & A_{22}+B_{22} \end{bmatrix}.$$

Vi kan alltså göra beräkningarna parallellt, t.ex. på olika processorer i en dator.

Via rad-kolonn-regeln kan vi också se liknande för matrismultiplikation:

Ex. Låt A vara som ovan och $B = \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix}$ med

$$B_{11} = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}, \quad B_{21} = \begin{bmatrix} 9 & 10 \end{bmatrix}. \text{ Då är}$$

$$AB = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \\ 9 & 10 \end{bmatrix} = \begin{bmatrix} 5+2\cdot 9 & 6+2\cdot 10 \\ 7+3\cdot 9 & 8+3\cdot 10 \\ 0+4\cdot 9 & 0+4\cdot 10 \end{bmatrix} =$$

forts.

$$= \begin{bmatrix} 5 & 6 \\ 7 & 8 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 2\cdot 9 & 2\cdot 10 \\ 3\cdot 9 & 3\cdot 10 \\ 4\cdot 9 & 4\cdot 10 \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} \\ A_{21}B_{11} \end{bmatrix} + \begin{bmatrix} A_{12}B_{21} \\ A_{22}B_{22} \end{bmatrix}$$

$$= \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} \\ A_{21}B_{11} + A_{22}B_{22} \end{bmatrix}$$

Dvs. vi kan utföra multiplikationen precis som om blocken var vanliga tal!

OBS: Blockens dimensioner måste såklart passa ihop.

På liknande sätt ser vi att

$$A^T = \begin{bmatrix} A_{11}^T & A_{21}^T \\ A_{12}^T & A_{22}^T \end{bmatrix}.$$

Notera att A_{12} och A_{21} transponerats och bytt plats. (Övn.: visa detta!)

Vi kan även (ibland) bestämma inverser enklare:

Ex. Låt $A = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}$ vara inverterbar och (med A_{11} och A_{22} kvadratiska!)

ansätt $A^{-1} = B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$ med samma indelning. →

Då har vi

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

$A_{11}B_{11} = I$ ger (Sats 2.3.8) att A_{11}^{-1} existerar, så vi får $\boxed{B_{11} = A_{11}^{-1}}$, och $A_{11}B_{12} = 0 \Rightarrow \boxed{B_{12} = 0}$

$$\Rightarrow \quad I = A_{21}B_{12} + A_{22}B_{22} = A_{22}B_{22}$$

$\Rightarrow$ A_{22}^{-1} existerar och $\boxed{B_{22} = A_{22}^{-1}}$

Slutligen får vi från $A_{21}B_{11} + A_{22}B_{21} = 0$ att

$$\boxed{B_{21}} = -A_{22}^{-1}A_{21}B_{11} = \boxed{-A_{22}^{-1}A_{21}A_{11}^{-1}}, \quad \text{dvs.}$$

$$A^{-1} = \begin{bmatrix} A_{11}^{-1} & 0 \\ -A_{22}^{-1}A_{21}A_{11}^{-1} & A_{22}^{-1} \end{bmatrix},$$

så det räcker att vi kan invertera A_{11} och A_{22}.

Matrisfaktoriseringar

En *faktorisering* av en matris A betyder att det finns matriser B och C så att $A = BC$.

Idén är att B, C skall vara "bättre" än A på något sätt.

Ex. Låt $B = [1 \ \tfrac{1}{2} \ \cdots \ \tfrac{1}{n}]^T \in \mathbb{R}^{n\times 1}$ och
$C = [1 \ \tfrac{1}{2} \ \cdots \ \tfrac{1}{n}] \in \mathbb{R}^{1\times n}$.

Då är

$$A = BC = \begin{bmatrix} 1 \\ 1/2 \\ \vdots \\ 1/n \end{bmatrix} [1 \ \tfrac{1}{2} \cdots \tfrac{1}{n}] = \begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{4} & \cdots & \frac{1}{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{2n} & \cdots & \frac{1}{n^2} \end{bmatrix}.$$

Att (naivt) beräkna Ax, $x \in \mathbb{R}^n$, tar $2n^2$ operationer, men att beräkna Cx och sen $B(Cx)$ tar bara $4n$.

Enorm skillnad för stora n! T.ex. $n = 1000$ ger $2 \cdot 10^6$ eller $4 \cdot 10^3$ operationer.

Det finns många olika faktoriseringar som ofta kan användas. En av de viktigaste ges av

<u>Def.</u> En <u>LU-faktorisering</u> av $A \in \mathbb{R}^{n\times n}$ ges av en <u>undertriangulär</u> matris L med värdet 1 på diagonalen, och en <u>övertriangulär</u> matris U sådana att $A = LU$.

<u>Def.</u> En matris $A = (a_{ij})$ är

<u>undertriangulär</u> om $a_{ij} = 0$ då $j > i$

<u>övertriangulär</u> om $a_{ij} = 0$ då $j < i$

Dvs.
$$\begin{pmatrix} * & 0 & 0 \\ * & * & 0 \\ * & * & * \end{pmatrix} \text{ alternativt } \begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \end{pmatrix}.$$

Om $A = LU$ kan vi lösa $Ax = b$ effektivt genom en tvåstegsprocess: Låt $y = Ux$ vara en ny (okänd) variabel.

Då har vi
$$\begin{cases} Ly = b \\ Ux = y \end{cases} \qquad (Ax = LUx = b)$$

Att lösa $Ly = b$ är lätt, då det har formen

$$\begin{pmatrix} 1 & & & \\ * & 1 & & \\ * & * & 1 & \\ * & * & * & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix}$$
(Detta kallas <u>framåt-substitution</u>.)

När vi har y kan vi lika lätt lösa $Ux = y$:

$$\begin{pmatrix} * & * & * & * \\ & * & * & * \\ & & * & * \\ & & & * \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$$
(Detta kallas <u>bakåt-substitution</u>.)

Mycket billigare ($\sim 2n^2$ oper.) än att lösa $Ax = b$ direkt ($\sim \frac{2}{3}n^3$ oper.).

Så hur bestämmer vi en LU-faktorisering?

Vi har faktiskt redan sett det!

I F4 visade vi att om $A \in \mathbb{R}^{n \times n}$ kan vi skriva $E_p \cdots E_1 A = U$

där E_k är elementära matriser och U är övertriangulär (trappstegsformen av A, $= I_n$ om A inv.).

Antag nu att vi aldrig behöver byta plats på rader när vi radreducerar. Då är varje E_k <u>undertriangulär</u>.

<u>Ex.</u>
$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -c_3 & 0 & 1 \end{pmatrix}}_{E_2} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ -c_2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{E_1} \underbrace{\begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \end{pmatrix}}_{A} = \begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & * & * \end{pmatrix}$$

Men då är $E_p \cdots E_1$ också undertriangulär!

Bevis: Antag A, B undertriang, dvs. $a_{ij} = b_{ij} = 0$ om $j > i$.

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{i} a_{ik} b_{kj}, \text{ och}$$

$j > i \Rightarrow j > k \Rightarrow b_{kj} = 0 \Rightarrow (AB)_{ij} = 0$. $\square$

Eftersom E_k är inv. är $E_p \cdots E_1$ också inv., så

$$A = LU \quad \text{där} \quad L = (E_p \cdots E_1)^{-1} = E_1^{-1} \cdots E_p^{-1},$$

där L är undertriangulär eftersom L^{-1} är det. (Övn. 2.5.19!)
(Ettor på diagonalen eftersom varje E_k har det.)

För att bestämma L, U kan vi radreducera A och multiplicera ihop L via E_k^{-1}. Mer effektivt är att observera att $E_p \cdots E_1 L = I$

och direkt sätta in element i L så att detta gäller.

<u>Ex.</u> $A = \begin{bmatrix} 2 & 4 & -1 \\ -4 & -5 & 3 \\ 2 & 10 & 4 \end{bmatrix}$

$E_2 E_1 A = \begin{bmatrix} 2 & 4 & -1 \\ 0 & 3 & 1 \\ 0 & 6 & 5 \end{bmatrix}$ och de nästkommande

E_k:na kommer inte röra första kolonnen. Alltså måste första kolonnen i L vara $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ (dela kolonnen med pivotelementet för att få en etta)

$$E_3 E_2 E_1 A = \begin{bmatrix} 2 & 4 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix},$$ så vi måste

ha $\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$ i 2:a kolonnen av L (dela $\begin{pmatrix} 0 \\ 3 \\ 6 \end{pmatrix}$ med pivotelementet 3)

Slutligen blir (såklart) sista kolonnen $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, dvs.

$$L = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad \text{och} \quad U = \begin{bmatrix} 2 & 4 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}.$$

(Kontrollera att $A = LU$!)

Som jämförelse (ej på föreläsningen) kan vi skriva upp

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$$

och deras inverser:

$$E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}.$$

Alltså blir

$$L = E_1^{-1} E_2^{-1} E_3^{-1} = E_1^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix}.$$