



## Föreläsning 3: Linjär algebra I.

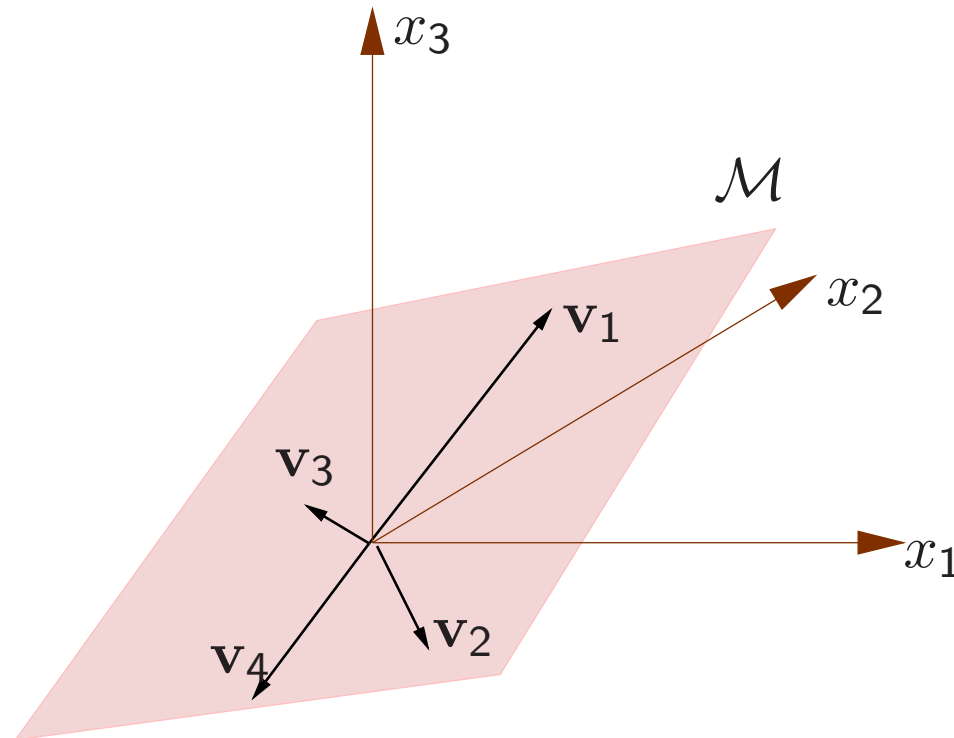
1. Underrum.
2. Ortogonalkomplement.
3. De fyra fundamentala underrummen.
4. Lösning av linjära ekvationssystem.
  - Linjära algebrans fundamentalsats.
5. Beräkning av de fundamentala underrummen,

## Underrum

**Definition 1.** En delmängd  $\mathcal{M} \subset \mathbf{R}^n$  kallas underrum i  $\mathbf{R}^n$  om för alla  $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{M}$  och  $\alpha_1, \alpha_2 \in \mathbf{R}$  det gäller att

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 \in \mathcal{M}$$

Notera att  $0 \in \mathcal{M}$  alltid gäller.



## Baser till underrum

Vektorerna  $\mathbf{v}_1, \dots, \mathbf{v}_k$  är linjärt oberoende om det inte finns skalärer  $\alpha_1, \dots, \alpha_k$  (inte alla noll) sådana att  $\sum_{l=1}^k \alpha_l \mathbf{v}_l = \mathbf{0}$ .

**Definition 2.**  $\mathcal{M}$  spänns upp av  $\mathbf{v}_1, \dots, \mathbf{v}_k$  om för varje  $\mathbf{v} \in \mathcal{M}$  finns  $\alpha_1, \dots, \alpha_k \in \mathbf{R}$  så att

$$\sum_{l=1}^k \alpha_l \mathbf{v}_l = \mathbf{v}$$

Om  $\mathbf{v}_1, \dots, \mathbf{v}_k$  är linjärt oberoende utgör de en bas till  $\mathcal{M}$  och dimensionen för  $\mathcal{M}$  är  $k$ , vilket betecknas  $\dim \mathcal{M} = k$ .

**Notation:**

$$\mathcal{M} = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\} = \left\{ \sum_{l=1}^k \alpha_l \mathbf{v}_l : \alpha_l \in \mathbf{R}; l = 1, \dots, k \right\}.$$

## Summer av underrum och ortogonala underrum

Givet två underrum  $\mathcal{M}_1$  och  $\mathcal{M}_2$ , så är delmängden

$\mathcal{M} = \{\mathbf{v}_1 + \mathbf{v}_2 : \mathbf{v}_1 \in \mathcal{M}_1, \mathbf{v}_2 \in \mathcal{M}_2\}$  ett underrum och vi skriver  $\mathcal{M} = \mathcal{M}_1 + \mathcal{M}_2$ .

Två vektorer  $\mathbf{v}_1$  och  $\mathbf{v}_2$  är ortogonala om  $\mathbf{v}_1^T \mathbf{v}_2 = \|\mathbf{v}_1\| \|\mathbf{v}_2\| \cos \theta = 0$ , d.v.s., om vinkeln  $\theta$  mellan dem är  $\pi/2$ , och vi skriver  $\mathbf{v}_1 \perp \mathbf{v}_2$ .

Två underrum  $\mathcal{M}_1$  och  $\mathcal{M}_2$  är ortogonala om  $\mathbf{v}_1 \perp \mathbf{v}_2$  för alla  $\mathbf{v}_1 \in \mathcal{M}_1$  och  $\mathbf{v}_2 \in \mathcal{M}_2$ , och vi skriver  $\mathcal{M}_1 \perp \mathcal{M}_2$ .

Om  $\dim \mathcal{M} = \dim \mathcal{M}_1 + \dim \mathcal{M}_2$  så kan varje  $\mathbf{v} \in \mathcal{M}$  skrivas unikt på formen  $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$ , där  $\mathbf{v}_1 \in \mathcal{M}_1$  och  $\mathbf{v}_2 \in \mathcal{M}_2$ .

## Direkt summa av underrum

Ett linjärt rum  $\mathcal{M}$  (till exempel  $\mathbf{R}^n$ ) är en direkt summa av underrummen  $\mathcal{M}_1, \mathcal{M}_2$  (d.v.s  $\mathcal{M}_1, \mathcal{M}_2 \subset \mathcal{M}$ ) om  $\mathcal{M}_1 \perp \mathcal{M}_2$  och varje  $\mathbf{v} \in \mathcal{M}$  unikt kan skrivas  $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$  för något  $\mathbf{v}_1 \in \mathcal{M}_1$  och  $\mathbf{v}_2 \in \mathcal{M}_2$ . Detta betecknas  $\mathcal{M} = \mathcal{M}_1 \oplus \mathcal{M}_2$ ,

**Exempel 1.** *Antag*

$$\mathcal{M}_1 = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$$

$$\mathcal{M}_2 = \text{span}\{\mathbf{v}_{k+1}, \dots, \mathbf{v}_l\}$$

$$\mathcal{M} = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_l\}$$

där  $\mathbf{v}_1, \dots, \mathbf{v}_l$  är linjärt oberoende vektorer. Då gäller att

$$\mathcal{M} = \mathcal{M}_1 \oplus \mathcal{M}_2$$

om  $\mathbf{v}_i \perp \mathbf{v}_j$  för alla  $i = 1, \dots, k$  och  $j = k + 1, \dots, l$ .

# Ortogonalkomplement

Ortogonalkomplementet till ett underrum  $\mathcal{M} \subset \mathbf{R}^n$  definieras som

$$\mathcal{M}^\perp = \{\mathbf{w} \in \mathbf{R}^n : \mathbf{w}^T \mathbf{v} = 0, \forall \mathbf{v} \in \mathcal{M}\}$$

Följande gäller

- $\mathcal{M}^\perp$  är ett underrum.
- $\mathbf{R}^n = \mathcal{M} \oplus \mathcal{M}^\perp$ .
- Om  $\dim \mathcal{M} = k$  så är  $\dim \mathcal{M}^\perp = n - k$ .
- $(\mathcal{M}^\perp)^\perp = \mathcal{M}$ .

## De fyra fundamentala underrummen

Betrakta den linjära operatorn  $A : \mathbf{R}^n \rightarrow \mathbf{R}^m$ ,

$$\mathbf{A} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{a}}_1 & \hat{\mathbf{a}}_2 & \dots & \hat{\mathbf{a}}_n \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{a}}_1^T \\ \bar{\mathbf{a}}_2^T \\ \vdots \\ \bar{\mathbf{a}}_m^T \end{bmatrix}$$

$$\text{Bildrummet } \mathcal{R}(\mathbf{A}) = \{ \mathbf{Ax} \mid \mathbf{x} \in \mathbf{R}^n \} = \left\{ \sum_{j=1}^n \hat{\mathbf{a}}_j x_j \mid x_j \in \mathbf{R}, j = 1, \dots, n \right\}$$

$$\begin{aligned} \text{Nollrummet : } \mathcal{N}(\mathbf{A}) &= \{ \mathbf{x} \in \mathbf{R}^n \mid \mathbf{Ax} = \mathbf{0} \} \\ &= \{ \mathbf{x} \in \mathbf{R}^n \mid \bar{\mathbf{a}}_i^T \mathbf{x} = 0, i = 1, \dots, m \} \end{aligned}$$

$$\text{Radrummet : } \mathcal{R}(\mathbf{A}^\top) = \{ \mathbf{A}^\top \mathbf{u} \mid \mathbf{u} \in \mathbf{R}^m \}$$

$$= \left\{ \sum_{i=1}^m \bar{\mathbf{a}}_i u_i \mid u_i \in \mathbf{R}, i = 1, \dots, m \right\}$$

$$\text{Vänstra Nollrummet: } \mathcal{N}(\mathbf{A}^\top) = \{ \mathbf{u} \in \mathbf{R}^m \mid \mathbf{A}^\top \mathbf{u} = \mathbf{0} \}$$

$$= \{ \mathbf{u} \in \mathbf{R}^m \mid \hat{\mathbf{a}}_j^\top \mathbf{u} = 0, j = 1, \dots, n \}$$

$$= \{ \mathbf{u} \in \mathbf{R}^m \mid \mathbf{u}^\top \mathbf{A} = \mathbf{0} \}$$

- $\mathcal{R}(\mathbf{A})$  och  $\mathcal{N}(\mathbf{A}^\top)$  är underrum i  $\mathbf{R}^m$
- $\mathcal{R}(\mathbf{A}^\top)$  och  $\mathcal{N}(\mathbf{A})$  är underrum i  $\mathbf{R}^n$



# De fyra underrummens ortogonala komplement

**Sats 1.** *Följande ortogonalitetssamband är uppfyllda*

$$(i) \mathcal{R}(\mathbf{A})^\perp = \mathcal{N}(\mathbf{A}^\top).$$

$$(ii) \mathcal{R}(\mathbf{A}^\top)^\perp = \mathcal{N}(\mathbf{A}).$$

$$(iii) \mathcal{N}(\mathbf{A})^\perp = \mathcal{R}(\mathbf{A}^\top).$$

$$(iv) \mathcal{N}(\mathbf{A}^\top)^\perp = \mathcal{R}(\mathbf{A}).$$

**Bevis:** *Kapitel 25.5 i OK (Gamla materialet: gula häftet).*

## Linjär algebrans fundamentalsats

**Sats 2.** Låt  $\mathbf{A} \in \mathbf{R}^{m \times n}$ . Då har man följande direkta summor

$$(a) \mathbf{R}^n = \mathcal{R}(\mathbf{A}^T) \oplus \mathcal{N}(\mathbf{A})$$

$$(b) \mathbf{R}^m = \mathcal{R}(\mathbf{A}) \oplus \mathcal{N}(\mathbf{A}^T)$$

där  $\mathcal{R}(\mathbf{A}^T) \perp \mathcal{N}(\mathbf{A})$  och  $\mathcal{R}(\mathbf{A}) \perp \mathcal{N}(\mathbf{A}^T)$ .

Vidare gäller att  $\dim \mathcal{R}(\mathbf{A}) = \dim \mathcal{R}(\mathbf{A}^T)$ .

**Bevis:** Då  $\mathcal{N}(\mathbf{A}), \mathcal{R}(\mathbf{A}^T) \subset \mathbf{R}^n$  enligt föregående sats uppfyller  $\mathcal{R}(\mathbf{A}^T)^\perp = \mathcal{N}(\mathbf{A})$  så gäller att  $\mathbf{R}^n = \mathcal{R}(\mathbf{A}^T) \oplus \mathcal{N}(\mathbf{A})$ . Beviset av (b) följer av samma resonemang.

# Lösning av linjära ekvationssystem

Låt  $\mathbf{A} \in \mathbf{R}^{m \times n}$  och  $\mathbf{b} \in \mathbf{R}^m$  och betrakta linjära ekvationssystemet

$$\mathbf{Ax} = \mathbf{b} \quad (1)$$

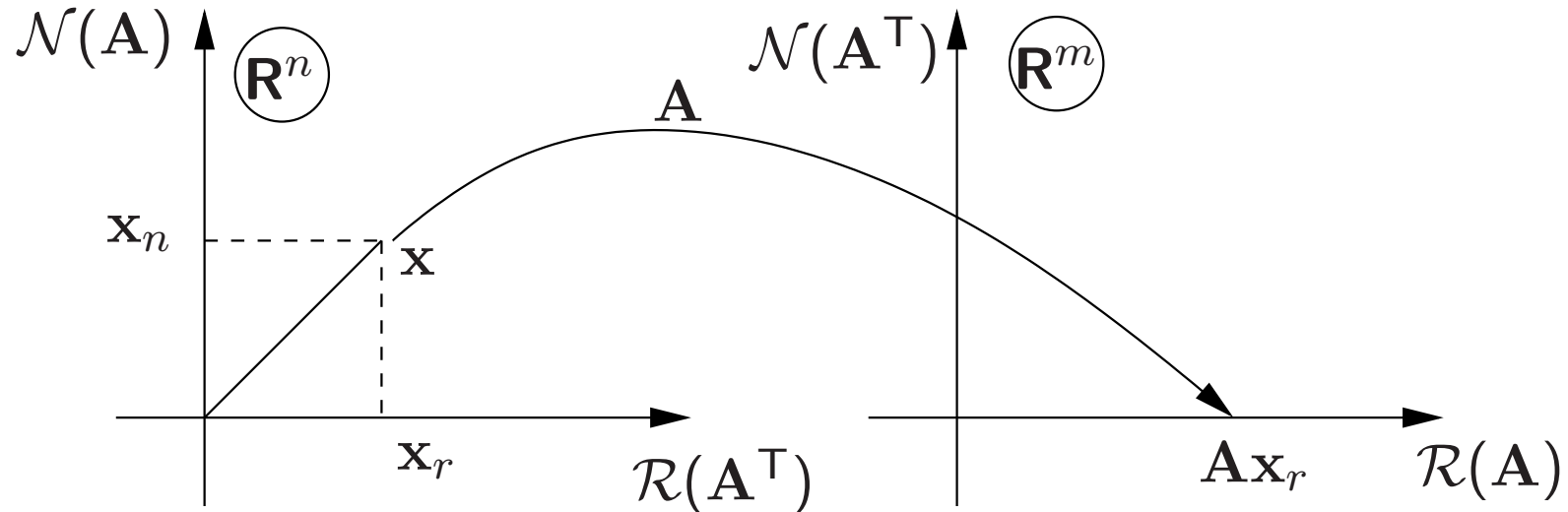
(A) När existerar det en lösning?

(B) Är lösningen unik? Vilken lösning väljer vi annars?

(C) Hur konstruerar vi lösningen?

Linjär algebrans fundamentalsats ger svar på ovanstående frågor.

## Geometrisk illustration av linjära algebrans huvudsats.



Diagrammet ger svar på frågorna (A) – (C) ovan.

- Ekvationssystemet  $\mathbf{Ax} = \mathbf{b}$  har lösning om  $\mathbf{b} \in \mathcal{R}(\mathbf{A})$ .
- Lösningen är unik om och endast om  $\mathcal{N}(\mathbf{A}) = \mathbf{0}$ . Annars kan varje lösning delas upp i två komponenter  $\mathbf{x} = \mathbf{x}_r + \mathbf{x}_n$  där  $\mathbf{x}_r \in \mathcal{R}(\mathbf{A}^T)$  uppfyller  $\mathbf{Ax}_r = \mathbf{b}$  och  $\mathbf{x}_n \in \mathcal{N}(\mathbf{A})$ .  
Ofta väljer vi den kortaste lösningen, dvs  $\mathbf{x}_n = \mathbf{0}$ .

De följande tre OH bilderna diskuterar frågorna (A) – (C) i mer detalj.

(A) Det finns en lösning till (1) om  $\mathbf{b} \in \mathcal{R}(\mathbf{A})$

(B) Lösningen är unik om och endast om  $\mathcal{N}(\mathbf{A}) = \mathbf{0}$ . Detta följer av att varje  $\mathbf{x} \in \mathbf{R}^n$  unikt kan skrivas  $\mathbf{x} = \mathbf{x}_r + \mathbf{x}_n$  där  $\mathbf{x}_r \in \mathcal{R}(\mathbf{A}^T)$  och  $\mathbf{x}_n \in \mathcal{N}(\mathbf{A})$ . Vi har  $\mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{x}_r + \mathbf{A}\mathbf{x}_n = \mathbf{A}\mathbf{x}_r$ . Komponenten  $\mathbf{x}_n \in \mathcal{N}(\mathbf{A})$  är således godtycklig och lösningen är unik om och endast om  $\mathcal{N}(\mathbf{A}) = \mathbf{0}$ .

Vilken lösning väljer vi om  $\mathcal{N}(\mathbf{A}) \neq \mathbf{0}$ ? Ett förslag är att välja den lösning som har kortast längd mätt i Euklidiska normen

$\|\mathbf{x}\| = \sqrt{\mathbf{x}^T \mathbf{x}}$ . Det följer att vi skall välja  $\mathbf{x}_n = \mathbf{0}$ .

(C) Antag  $\mathbf{b} \in \mathcal{R}(\mathbf{A})$ . Det finns tre fall för konstruktion av lösning.

(i)  $\mathcal{N}(\mathbf{A}) = \mathbf{0}$  och  $m = n$  (kvadratisk matris). Då erhålls lösningen som  $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ .

(ii) Om  $\mathcal{N}(\mathbf{A}) \neq \mathbf{0}$  så väljer vi minsta norm lösningen som ges av  $\mathbf{x} = \mathbf{A}^T \mathbf{u}$ , där  $\mathbf{u} \in \mathbf{R}^m$  löser  $\mathbf{A}\mathbf{A}^T \mathbf{u} = \mathbf{b}$ . Om  $\mathbf{A}\mathbf{A}^T$  är inverterbar så får man det slutna uttrycket  $\mathbf{x} = \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{b}$ .

**Bevis:** Minsta normlösningen  $\mathbf{x} = \mathbf{x}_r \in \mathcal{R}(\mathbf{A}^T)$  kan uttryckas  $\mathbf{x} = \mathbf{A}^T \mathbf{u}$  där  $\mathbf{u} \in \mathbf{R}^m$  bestäms av att  $\mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{A}^T \mathbf{u} = \mathbf{b}$ .

(C) fortsättning

(iii)  $\mathcal{N}(\mathbf{A}) = \mathbf{0}$  och  $m > n$  (överbestämt system).

Då erhålls lösningen som  $\mathbf{x} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b}$ .

**Bevis:** *Beviset läses kursivt.* Detta följer av att

$$\mathbf{Ax} = \mathbf{b} \quad \Leftrightarrow \quad \mathbf{A}^\top \mathbf{Ax} = \mathbf{A}^\top \mathbf{b}. \quad (2)$$

Implikationen åt höger är uppenbar medan implikationen åt vänster följer eftersom  $\mathbf{A}^\top (\mathbf{Ax} - \mathbf{b}) = \mathbf{0}$  betyder att  $\mathbf{Ax} - \mathbf{b} \in \mathcal{N}(\mathbf{A}^\top)$  samtidigt som vi antagit att  $\mathbf{Ax} - \mathbf{b} \in \mathcal{R}(\mathbf{A})$ . Då  $\mathcal{N}(\mathbf{A}^\top) = \mathcal{R}(\mathbf{A})^\perp$  betyder detta att  $\mathbf{Ax} - \mathbf{b} \in \mathcal{R}(\mathbf{A}) \cap \mathcal{R}(\mathbf{A})^\perp = \mathbf{0}$ . Detta visar implikationen åt vänster i (2). Matrisen  $\mathbf{A}^\top \mathbf{A} \in \mathbf{R}^{m \times m}$  och man kan enkelt visa att  $\mathcal{N}(\mathbf{A}^\top \mathbf{A}) = \mathbf{0}$ . Den är därmed inverterbar och högra likheten i (2) ger att  $\mathbf{x} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b}$ .

# Beräkning av de fyra fundamentala underrummen

Två metoder för beräkning av de fundamentala underrummen

- Singulärvärdesfaktorisering. Bäst för numerisk beräkning.
- Gauss-Jordans metod. Vi fokuserar på denna.





Steg 1 (fortsättning) Trappstegskolumnerna är indicerade  $\beta_1, \dots, \beta_r$ , där  $r = \text{rang}(\mathbf{A}) = \dim \mathcal{R}(\mathbf{A})$  (rangén för matrisen).

Övriga kolumner indiceras  $\nu_1, \dots, \nu_l$ ,  $l = n - r$ .

Steg 2 Definiera

$$\begin{aligned}\mathbf{A}_\beta &= \begin{bmatrix} \hat{a}_{\beta_1} & \dots & \hat{a}_{\beta_r} \end{bmatrix} \\ \mathbf{U}_\beta &= \begin{bmatrix} u_{\beta_1} & \dots & u_{\beta_r} \end{bmatrix} = \mathbf{I}_r \\ \mathbf{U}_\nu &= \begin{bmatrix} u_{\nu_1} & \dots & u_{\nu_l} \end{bmatrix} \\ \mathbf{S} &= \mathbf{P}^{-1}\end{aligned}$$

Notera att  $\mathbf{U}_\beta$  (trappstegskolumnerna i  $\mathbf{U}$ ) är en identitetsmatris.

Steg 2 (fortsättning) Då  $\mathbf{S} = \mathbf{P}^{-1}$  kan (3) ekvivalent skrivas

$$\mathbf{A} = \mathbf{S}\mathbf{T} = \begin{bmatrix} \mathbf{S}_1 & \mathbf{S}_2 \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ \mathbf{O} \end{bmatrix} = \mathbf{S}_1\mathbf{U}$$

Om vi betraktar kolonnerna med index  $\beta_1, \dots, \beta_r$  får vi

$$\mathbf{A}_\beta = \mathbf{S}_1\mathbf{U}_\beta = \mathbf{S}_1$$

vilket ger  $\mathbf{A} = \mathbf{A}_\beta\mathbf{U}$ .

Steg 3 Faktoriseringen  $\mathbf{A} = \mathbf{A}_\beta \mathbf{U}$  är av formen  $\mathbf{A} = \mathbf{B}\mathbf{C}$  där

$$\begin{array}{c}
 \left. \begin{array}{c} \text{m rader} \\ \left[ \begin{array}{c} \mathbf{A} \end{array} \right] \end{array} \right\} = \left[ \begin{array}{c} \mathbf{B} \end{array} \right] \left[ \begin{array}{c} \mathbf{C} \end{array} \right] \left. \vphantom{\left[ \begin{array}{c} \mathbf{C} \end{array} \right]} \right\} \text{r linjärt oberoende rader} \\
 \underbrace{\hspace{10em}}_{\text{n kolonner}} \quad \underbrace{\hspace{10em}}_{\text{r linjärt oberoende kolonner}}
 \end{array}$$

Enligt kap. 26 i OK (sektion 6.1 i gula häftet) gäller

$$\begin{aligned}
 \mathcal{R}(\mathbf{A}) &= \mathcal{R}(\mathbf{A}_\beta \mathbf{U}) = \mathcal{R}(\mathbf{A}_\beta) = \text{span}\{\hat{a}_{\beta_1}, \dots, \hat{a}_{\beta_r}\} \\
 \mathcal{N}(\mathbf{A}) &= \mathcal{N}(\mathbf{A}_\beta \mathbf{U}) = \mathcal{N}(\mathbf{U}) = \{\mathbf{x} \in \mathbf{R}^n : \mathbf{U}\mathbf{x} = \mathbf{0}\} \\
 &= \{\mathbf{x} \in \mathbf{R}^n : \mathbf{U}_\beta \mathbf{x}_\beta + \mathbf{U}_\nu \mathbf{x}_\nu = \mathbf{0}\} \\
 &= \{\mathbf{x} \in \mathbf{R}^n : \mathbf{x}_\beta = -\mathbf{U}_\nu \mathbf{x}_\nu; \mathbf{x}_\nu \in \mathbf{R}^l\}
 \end{aligned}$$

Steg 3 (fortsättning) Vidare har vi med hjälp av ovanstående faktoreringsresultat

$$\mathcal{R}(\mathbf{A}^T) = \mathcal{R}(\mathbf{U}^T \mathbf{A}_\beta^T) = \mathcal{R}(\mathbf{U}^T)$$

Slutligen

$$\mathcal{N}(\mathbf{A}^T) = \mathcal{R}(\mathbf{A})^\perp = \{\mathbf{y} \in \mathbf{R}^m : \mathbf{A}_\beta^T \mathbf{y} = \mathbf{0}\}$$

vilket motsvarar ett ekvationssystem som skall lösas.

# Lösning av linjära ekvationssystem

Existens och unikheter av lösning till det linjära ekvationssystemet

$$\mathbf{Ax} = \mathbf{b}$$

kan bestämmas med

1. linjär algebrans huvudsats
  - Ger teoretisk insikt samt lösningsformler (se föreläsning 3).
2. Gauss-Jordans metod
  - Praktisk metod för handräkning.

# Gauss-Jordans metod för lösning av ekvationssystem

Transformera ekvationssystemet med hjälp av elementära radoperationer

$$\mathbf{Ax} = \mathbf{b}$$

$$\Leftrightarrow \mathbf{PAx} = \mathbf{Pb}$$

$$\Leftrightarrow \begin{bmatrix} \mathbf{U} \\ \mathbf{0} \end{bmatrix} \mathbf{x} = \begin{bmatrix} \bar{\mathbf{b}}_r \\ \bar{\mathbf{b}}_n \end{bmatrix}$$

där

$$\mathbf{U} = \begin{bmatrix} \begin{array}{cccc} 1 & * & 0 & * \\ & 1 & * & 0 \\ & & 1 & * \\ & & & 1 \end{array} & \begin{array}{cccc} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{array} \\ \mathbf{0} & \begin{array}{cccc} & & & \\ & & & \\ & & & \\ & & & \end{array} \\ \beta_1 & \beta_2 & \beta_3 & \beta_r \end{bmatrix}$$

Observera att

1. Lösning saknas om, och endast om,  $\bar{\mathbf{b}}_n \neq 0$
2. Om  $\bar{\mathbf{b}}_n = 0$ , så finns lösningar, och de kan alla skrivas på formen  $\mathbf{x} = \mathbf{x}_r + \mathbf{x}_n$  där  $\mathbf{x}_n \in \mathcal{N}(\mathbf{A})$  och

$$\mathbf{x} = \begin{cases} \mathbf{x}_{\beta_l}, & \text{för koefficienter motsvarande trappstegskolumnerna} \\ 0, & \text{annars} \end{cases}$$

där

$$\mathbf{x}_{\beta} = \begin{bmatrix} x_{\beta_1} \\ \vdots \\ x_{\beta_k} \end{bmatrix} = \bar{\mathbf{b}}_r$$



# Läsanvisningar

I optimeringskompendiet: Kap. 23-25

Gamla materialet:

- Linjär Algebra för optimerare, Gula häftet, sidorna 19-35.