Föreläsning 3: Linjär algebra.

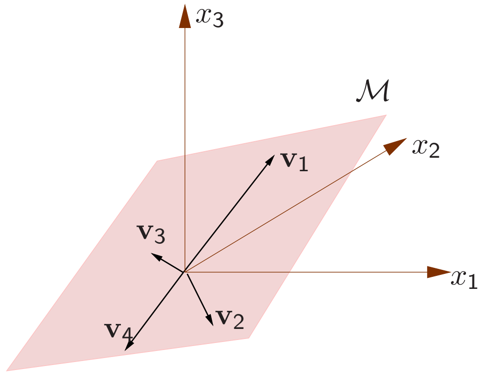
1. Underrum.
2. Ortogonalkomplement.
3. De fyra fundamentala underrummen.
4. Lösning av linjära ekvationssystem.
   - Linjära algebrans fundamentalsats.

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

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

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

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

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

Summor 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 θ 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 ortogonala underrum

Ett linjärt rum $\mathcal{M}$ (till exempel $\mathbf{R}^n$) är en (ortogonal) 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*

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

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

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

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

Nollrummet : $\mathcal{N}(\mathbf{A}) = \{\mathbf{x} \in \mathbf{R}^n | \mathbf{A}\mathbf{x} = \mathbf{0}\}$

$$= \{\mathbf{x} \in \mathbf{R}^n | \bar{\mathbf{a}}_i^\mathsf{T}\mathbf{x} = \mathbf{0}, \ i = 1, \dots, m\}$$

Radrummet : $\mathcal{R}(\mathbf{A}^\mathsf{T}) = \{\mathbf{A}^\mathsf{T}\mathbf{u} | \mathbf{u} \in \mathbf{R}^m\}$

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

Vänstra Nollrummet: $\mathcal{N}(\mathbf{A}^\mathsf{T}) = \{\mathbf{u} \in \mathbf{R}^m | \mathbf{A}^\mathsf{T}\mathbf{u} = \mathbf{0}\}$

$$\begin{aligned}&= \{\mathbf{u} \in \mathbf{R}^m | \hat{\mathbf{a}}_j^\mathsf{T}\mathbf{u} = \mathbf{0}, \ j = 1, \dots, n\}\\ &= \{\mathbf{u} \in \mathbf{R}^m | \mathbf{u}^\mathsf{T}\mathbf{A} = \mathbf{0}\}\end{aligned}$$

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

De fyra underrummens ortogonala komplement

Sats 1. *Följande ortogonalitetssamband är uppfyllda*

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

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

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

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

Bevis: *Läs Svanbergs gula häfte.*

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

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

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

Bevis: Då $\mathcal{N}(\mathbf{A}), \mathcal{R}(\mathbf{A}^{\mathsf{T}}) \subset \mathbf{R}^n$ enligt föregående sats uppfyller $\mathcal{R}(\mathbf{A}^{\mathsf{T}})^{\perp} = \mathcal{N}(\mathbf{A})$ 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

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{1}$$

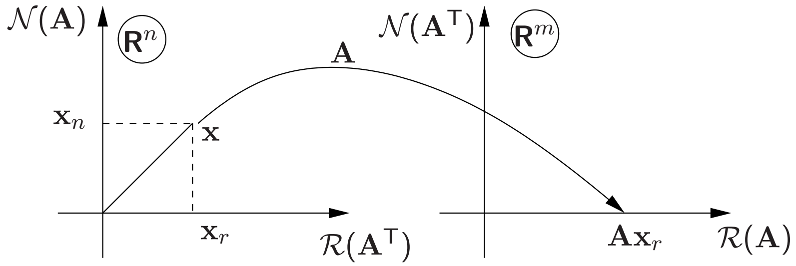
(A) När existerar det en lösning?

(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 $\mathbf{A}\mathbf{x} = \mathbf{b}$ har lösning omm $\mathbf{b} \in \mathcal{R}(\mathbf{A})$.
- Lösningen är unik om och endast om $\mathcal{N}(\mathbf{A}) = 0$. Annars kan varje lösning delas upp i två komponenter $\mathbf{x} = \mathbf{x}_r + \mathbf{x}_n$ där $\mathbf{x}_r \in \mathcal{R}(\mathbf{A}^{\mathsf{T}})$ uppfyller $\mathbf{A}\mathbf{x}_r = \mathbf{b}$ 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 $\mathbf{x}_r \in \mathcal{R}(\mathbf{A}^\mathsf{T})$ 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 $\|\mathbf{x}\| = \sqrt{\mathbf{x}^\mathsf{T}\mathbf{x}}$. 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 $\mathcal{N}(\mathbf{A}) \neq 0$ så väljer vi minsta norm lösningen som ges av $\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{u}$, där $\mathbf{u} \in \mathbf{R}^m$ löser $\mathbf{A}\mathbf{A}^\mathsf{T}\mathbf{u} = \mathbf{b}$. Om $\mathbf{A}\mathbf{A}^\mathsf{T}$ är inverterbar så får man det slutna uttrycket $\mathbf{x} = \mathbf{A}^\mathsf{T}(\mathbf{A}\mathbf{A}^\mathsf{T})^{-1}\mathbf{b}$.

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

(C) fortsättning

(iii) $\mathcal{N}(\mathbf{A}) = \mathbf{0}$ och $m > n$ (överbestämt system). Då erhålls lösningen som $\mathbf{x} = (\mathbf{A}^\mathsf{T}\mathbf{A})^{-1}\mathbf{A}^\mathsf{T}\mathbf{b}$.

Bevis: Beviset läses kursivt. Detta följer av att

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad \Leftrightarrow \quad \mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{b}. \tag{2}$$

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

Läsanvisningar

- Linjär Algebra för optimerare, Gula häftet, sidorna 19-29.