Lineær algebra
2. kursusgang

Mindste kvadraters linie

Find en linie $y = ax + b$ der bedst muligt tilnærmer punkterne

$$(x_1, y_1), \dots, (x_n, y_n).$$

Mindste kvadraters linien er linien hvor a og b er valgt sådan at

$$(ax_1 + b - y_1)^2 + \dots + (ax_n + b - y_n)^2$$

er mindst mulig.

a og b findes som mindste kvadraters løsning til ligningssystemet

$$\begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

Altså en vektor $\begin{bmatrix} b \\ a \end{bmatrix}$ som opfylder at afstanden mellem vektoren på venstre side og vektoren på højre side er mindst mulig.

$$A = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$$ kaldes designmatrix.

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$ kaldes observationsvektor.

Parametervektoren $\begin{bmatrix} b \\ a \end{bmatrix}$ bestemmes som løsning til normalligningen

$$A^\top A \begin{bmatrix} b \\ a \end{bmatrix} = A^\top \mathbf{y}.$$

Afsnit 6.3: ortogonal projektioner

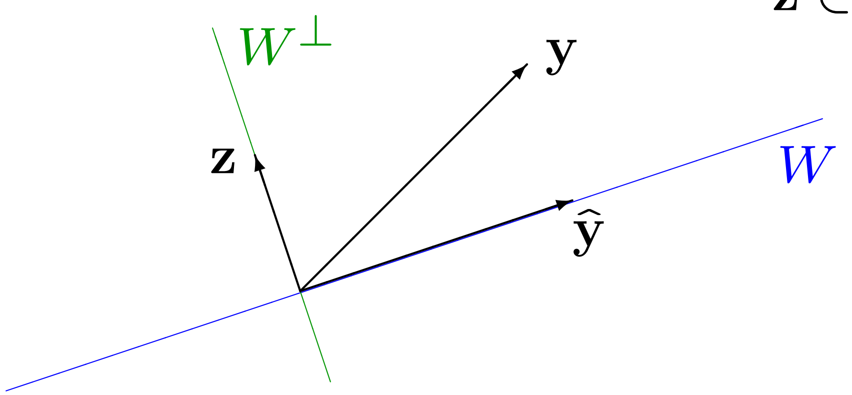
W: et underrum af $\mathbb{R}^n$
$\mathbf{y}$: en vektor i $\mathbb{R}^n$

Så findes der entydige vektorer $\hat{\mathbf{y}}$ og $\mathbf{z}$ så

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$$

$$\hat{\mathbf{y}} \in W$$

$$\mathbf{z} \in W^{\perp}$$



Thm 6.7. Hvis $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ er en ortonormal basis for W
(altså: $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ er en basis for W
og $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ for alle $i \neq j$ og $\mathbf{u}_i \cdot \mathbf{u}_1 = 1$ for alle i)
så er

$$\hat{\mathbf{y}} = \mathbf{y} \cdot \mathbf{u}_1 \mathbf{u}_1 + \mathbf{y} \cdot \mathbf{u}_2 \mathbf{u}_2 + \ldots + \mathbf{y} \cdot \mathbf{u}_p \mathbf{u}_p$$

og

$$\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}.$$

$\hat{\mathbf{y}}$ kaldes den ortogonale projektion af $\mathbf{y}$ på W, skrives $\hat{\mathbf{y}} = U_W(\mathbf{y})S$.

$\hat{\mathbf{y}}$ er den vektor i i W der er nærmest $\mathbf{y}$:
For alle $\mathbf{v} \in W$, $\mathbf{v} \neq \hat{\mathbf{y}}$ er dist$(\mathbf{y}, \mathbf{v}) >$ dist$(\mathbf{y}, \hat{\mathbf{y}})$.

Standard-matricen for den lineære transformation U_W betegnes P_W. Dette er en $n \times n$ matrix som opfylder

$$U_W(\mathbf{y}) = P_W\mathbf{y}.$$

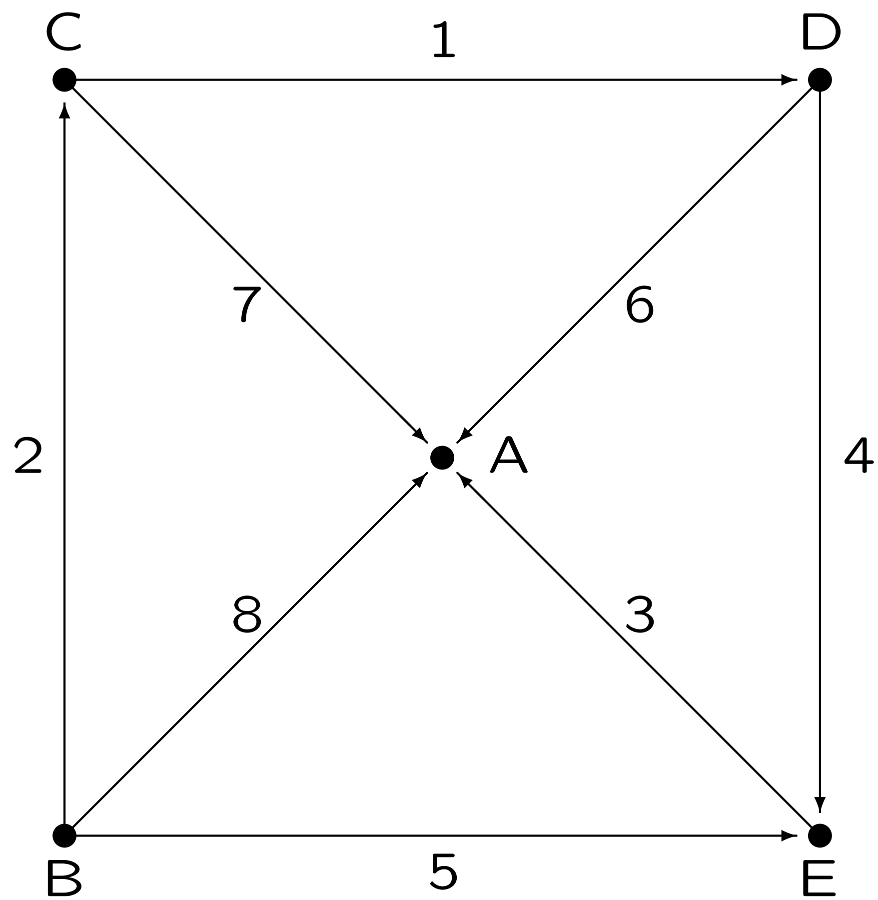
Thm 6.8 Hvis C er en matrix hvis søjler udgør en basis for W så er

$$P_W = C(C^TC)^{-1}C^T.$$

Hvis søjlerne i C udgør en ortonormalbasis for W så er

$$P_W = CC^T.$$

Eksempel: Højdeforskelle.



Afsnit 6.4, side 407

Vi skal finde $\mathbf{x}$ der (næsten) er løsning til ligningssystemet

$$A\mathbf{x} = \mathbf{b}.$$

En mindste kvadraters løsning er en vektor $\hat{\mathbf{x}}$ som opfylder

$$\text{dist}(\mathbf{b}, A\hat{\mathbf{x}}) \leq \text{dist}(\mathbf{b}, A\mathbf{x})$$

for alle $\mathbf{x} \in \mathbb{R}^n$.

$$||\mathbf{b} - A\hat{\mathbf{x}}|| \leq ||\mathbf{b} - A\mathbf{x}||$$

Hvis $\hat{\mathbf{x}}$ er en løsning til $A\mathbf{x} = \mathbf{b}$ så er $\hat{\mathbf{x}}$ også en mindste kvadraters løsning.

Col A består af alle vektorer på formen $A\mathbf{x}$.

Så hvis $A\mathbf{x} = \mathbf{b}$ har en løsning så ligger $\mathbf{b}$ i Col A.

Og hvis $A\mathbf{x} = \mathbf{b}$ ikke har en løsning så ligger $\mathbf{b}$ ikke i Col A.

Lad $\hat{\mathbf{b}}$ den ortogonale projektion af $\mathbf{b}$ på underrummet Col A. Så er

$$\text{dist}(\mathbf{b}, \hat{\mathbf{b}}) \leq \text{dist}(\mathbf{b}, \mathbf{c})$$

for enhver vektor $\mathbf{c} \in$ Col A.

Hvis $\hat{\mathbf{x}}$ er løsning til $A\mathbf{x} = \hat{\mathbf{b}}$ så er

$$\text{dist}(\mathbf{b}, A\hat{\mathbf{x}}) = \text{dist}(\mathbf{b}, \hat{\mathbf{b}}) \leq \text{dist}(\mathbf{b}, A\mathbf{x})$$

for alle $\mathbf{x} \in \mathbb{R}$. $\hat{\mathbf{x}}$ er altså en mindste kvadraters løsning.

Da $\hat{\mathbf{b}}$ er ortogonal projektionen af $\mathbf{b}$ på Col A er

$$\mathbf{b} - \hat{\mathbf{b}} = \mathbf{b} - A\hat{\mathbf{x}} \in (\text{Col } A)^{\perp}.$$

Altså: $\mathbf{b} - A\hat{\mathbf{x}}$ er ortogonal på søjlerne i $A = [\mathbf{a}_1 \ldots \mathbf{a}_n]$.
Derfor er

$$A^T(\mathbf{b} - A\hat{\mathbf{x}}) = \begin{bmatrix} \ldots & \mathbf{a}_1 & \ldots \\ \ldots & \mathbf{a}_2 & \ldots \\ & \vdots & \\ \ldots & \mathbf{a}_m & \ldots \end{bmatrix} (\mathbf{b} - A\hat{\mathbf{x}}) = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{0}.$$

Hvis $\hat{\mathbf{x}}$ er en mindste kvadraters løsning så har vi altså:

$$A^T(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}.$$

$$A^T\mathbf{b} - A^T A\hat{\mathbf{x}} = \mathbf{0}.$$

$$A^T A\hat{\mathbf{x}} = A^T\mathbf{b}.$$

Normalligningen for $A\mathbf{x} = \mathbf{b}$ er

$$A^T A\mathbf{x} = A^T\mathbf{b}.$$

$\hat{\mathbf{x}}$ er en mindste kvadraters løsning til $A\mathbf{x} = \mathbf{b}$
hvis og kun hvis
$\hat{\mathbf{x}}$ er løsning til normalligningen $A^T A\mathbf{x} = A^T\mathbf{b}$.

$A\mathbf{x} = \hat{\mathbf{b}}$ har altid mindst én løsning, $\hat{\mathbf{x}}$.

Hvis søjlerne i A er lineært uafhængige så har $A\mathbf{x} = \hat{\mathbf{b}}$ en entydig løsning.

Hvis søjlerne i A er lineært afhængige så er der frie variable og dermed uendeligt mange løsning til $A\mathbf{x} = \hat{\mathbf{b}}$.

Dette uddybes på næste side.

Hvis søjlerne i A er lineært uafhængige så er $A^T A$ invertibel og normalligningen

$$A^T A\mathbf{x} = A^T\mathbf{b}$$

har en entydig løsning:

$$\hat{\mathbf{x}} = (A^T A)^{-1} A^T\mathbf{b}.$$

Der er så altså en entydig mindste kvadraters løsning.

Hvis søjlerne i A er lineært afhængige så er $A^T A$ ikke invertibel.

Simpelt eksempel

Afstanden mellem to punkter har en (ukendt) værdi x.

Vi foretager en måling af afstanden og finder værdien b_1.
Vi har så $x = b_1 + \hat{r}_1$, hvor $\hat{r}_1$ angiver en fejl i målingen.

Vi gentager målingen ind til vi har n målinger:

$$x = b_1 + \hat{r}_1$$
$$x = b_2 + \hat{r}_2$$
$$\vdots$$
$$x = b_n + \hat{r}_n$$

Vi ønsker at bestemme x så $\hat{r}_1^2 + \ldots + \hat{r}_n^2$ er så lille som muligt.

Vi skal altså finde en mindste kvadraters løsning til ligningssystemet:

$$\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} x = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Vi anvender teorien med

$$A = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad \text{og} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Vi udregner

$$A^T A = n$$

$$A^T \mathbf{b} = b_1 + \ldots + b_n$$

Normalligningen er altså

$$nx = b_1 + \ldots + b_n.$$

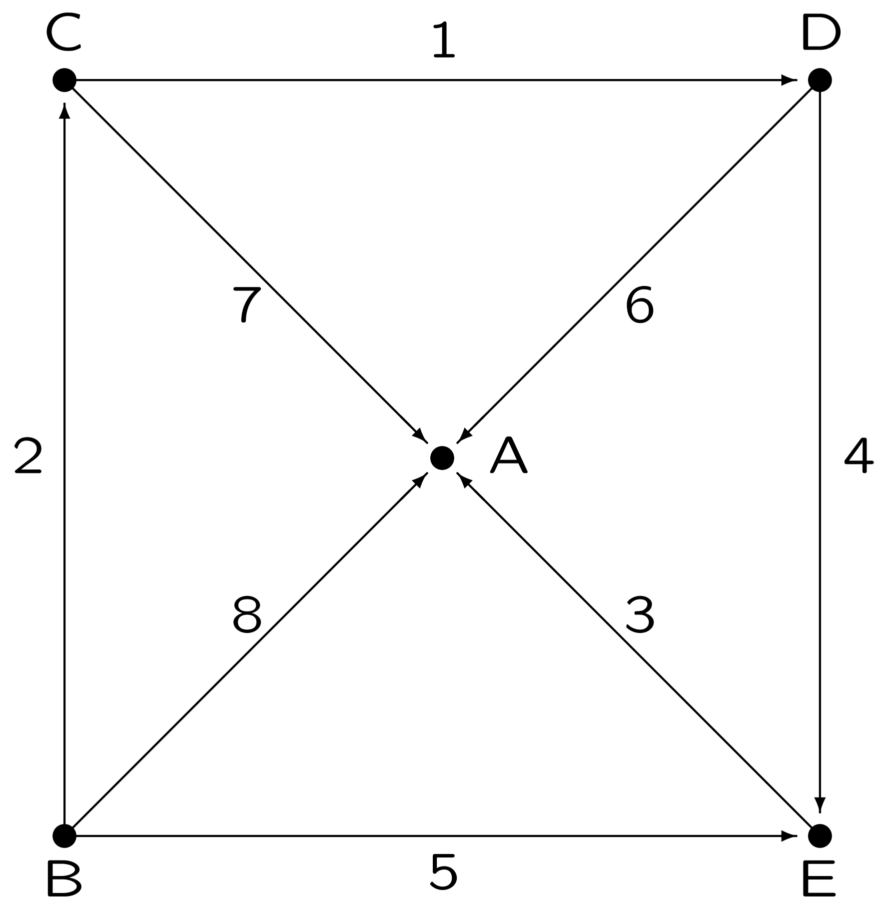
Mindste kvadraters løsningen er

$$\widehat{x} = \frac{b_1 + \ldots + b_n}{n}.$$

Vigtigt eksempel

Mellem fem punkter A, B, C, D, E er der målt følgende højdeforskelle:

Fra punkt	Til punkt	Højdeforskel
B	A	8
B	C	2
B	E	5
C	A	7
C	D	1
D	A	6
D	E	4
E	A	3


C
1
D
7
6
2
A
4
8
3
B
5
E

Højderne for de fem punkter er h_A, h_B, h_C, h_D, h_E.

Observationsligninger:

$$
\begin{aligned}
h_A - h_B &= 8 + \widehat{r}_1\\
h_C - h_B &= 2 + \widehat{r}_2\\
h_E - h_B &= 5 + \widehat{r}_3\\
h_A - h_C &= 7 + \widehat{r}_4\\
h_D - h_C &= 1 + \widehat{r}_5\\
h_A - h_D &= 6 + \widehat{r}_6\\
h_E - h_D &= 4 + \widehat{r}_7\\
h_A - h_E &= 3 + \widehat{r}_8
\end{aligned}
$$

På matrixform:

$$\begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 \\ 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} h_A \\ h_B \\ h_C \\ h_D \\ h_E \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \\ 5 \\ 7 \\ 1 \\ 6 \\ 4 \\ 3 \end{bmatrix} + \begin{bmatrix} \widehat{r}_1 \\ \widehat{r}_2 \\ \widehat{r}_3 \\ \widehat{r}_4 \\ \widehat{r}_5 \\ \widehat{r}_6 \\ \widehat{r}_7 \\ \widehat{r}_8 \end{bmatrix}$$

Sæt

$$A = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 \\ 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 & -1 \end{bmatrix} \quad \text{og} \quad \mathbf{b} = \begin{bmatrix} 8 \\ 2 \\ 5 \\ 7 \\ 1 \\ 6 \\ 4 \\ 3 \end{bmatrix},$$

og find mindste kvadraters løsning til

$$A\mathbf{x} = \mathbf{b}.$$

Udregn

$$A^TA = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 \\ 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 & -1 \end{bmatrix} =$$

$$\begin{bmatrix} 4 & -1 & -1 & -1 & -1 \\ -1 & 3 & -1 & 0 & -1 \\ -1 & -1 & 3 & -1 & 0 \\ -1 & 0 & -1 & 3 & -1 \\ -1 & -1 & 0 & -1 & 3 \end{bmatrix}$$

og

$$A^T\mathbf{b} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 8 \\ 2 \\ 5 \\ 7 \\ 1 \\ 6 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 24 \\ -15 \\ -6 \\ -9 \\ 6 \end{bmatrix}$$

Udvidet koefficientmatrix for normalligningen $A^T A\mathbf{x} = A^T\mathbf{b}$:

$$\begin{bmatrix} 4 & -1 & -1 & -1 & -1 & 24 \\ -1 & 3 & -1 & 0 & -1 & -15 \\ -1 & -1 & 3 & -1 & 0 & -6 \\ -1 & 0 & -1 & 3 & -1 & -9 \\ -1 & -1 & 0 & -1 & 3 & 6 \end{bmatrix} \xrightarrow{\texttt{rref}} \begin{bmatrix} 1 & 0 & 0 & 0 & -1 & 2.8 \\ 0 & 1 & 0 & 0 & -1 & -5.4 \\ 0 & 0 & 1 & 0 & -1 & -4 \\ 0 & 0 & 0 & 1 & -1 & -3.4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

og matricen omskrevet til reduceret række echelon form.

Mindste kvadraters løsninger:

$$\begin{bmatrix} h_A \\ h_B \\ h_C \\ h_D \\ h_E \end{bmatrix} = \begin{bmatrix} 2.8 + t \\ -5.4 + t \\ -4 + t \\ -3.4 + t \\ t \end{bmatrix},$$

hvor t er en fri variabel.

Antag nu at E et fikspunkt med højde 10. Vi kan så erstatte h_E med 10 i ligningerne.

Observationsligninger:

$$h_A - h_B = 8 + \widehat{r}_1$$
$$h_C - h_B = 2 + \widehat{r}_2$$
$$-h_B = -5 + \widehat{r}_3$$
$$h_A - h_C = 7 + \widehat{r}_4$$
$$h_D - h_C = 1 + \widehat{r}_5$$
$$h_A - h_D = 6 + \widehat{r}_6$$
$$-h_D = -6 + \widehat{r}_7$$
$$h_A = 13 + \widehat{r}_8$$

Sæt

$$A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad \text{og} \quad \mathbf{b} = \begin{bmatrix} 8 \\ 2 \\ -5 \\ 7 \\ 1 \\ 6 \\ -6 \\ 13 \end{bmatrix},$$

og find mindste kvadraters løsning til

$$A\mathbf{x} = \mathbf{b}.$$

Søjlerne i A er nu lineært uafhængige. Der er derfor en entydig mindste kvadraters løsning:

$$\begin{bmatrix} h_A \\ h_B \\ h_C \\ h_D \end{bmatrix} = (A^T A)^{-1} A^T \mathbf{b} = \begin{bmatrix} 12.8 \\ 4.6 \\ 6 \\ 6.6 \end{bmatrix}$$

Bemærk: Vi har ikke beregnet $\widehat{\mathbf{b}}$.

Hvis det ønskes kan den nu beregnes som

$$\widehat{\mathbf{b}} = A\hat{\mathbf{x}}$$