6.5 Minste kvadraters metode

Anta at vi er gitt en $n \times m$ matrise

$A = [a_1, \dots, a_m]$ og $b \in \mathbb{R}^m$.

Betrakt det lineære systemet

$$Ax = b$$

Vi kan løse dette hvis og bare hvis

$$b \in W = \mathrm{Col}(A) = \mathrm{span}\{a_1, \dots, a_m\}.$$

Hva gjør vi når $b \notin W$?

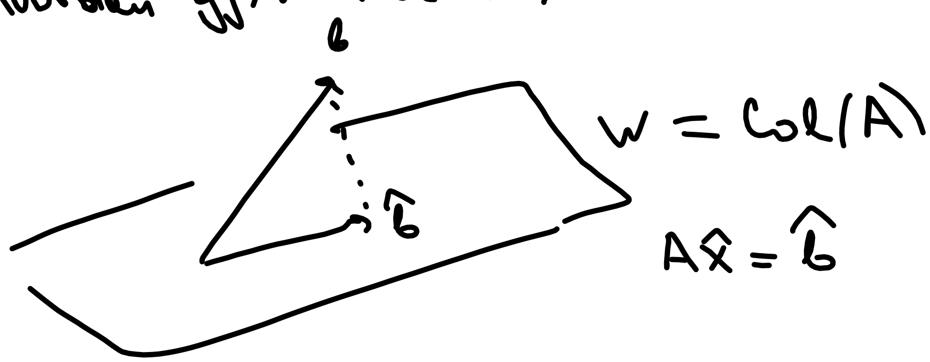
Vi kan prøve å finne den best approksimasjonen til en løsning: vi prøver å finne $\hat{x}$ slik at $\|A\hat{x} - b\|$ er minst mulig.

Dvs. vi prøver å minimisere

$$\sum_{i=1}^{m} ((A\hat{x})_i - b_i)^2.$$

Slik $\hat{x}$ kalles en <u>minste kvadraters løsning</u> av systemet $Ax = b$

Hvordan gjør vi dette?



<u>Løsning 1:</u>

Vi tar $\hat{b} = \mathrm{Proj}_W(b)$,

etterpå løser $A\hat{x} = \hat{b}$.

Vi kan skrive dette på følgende måte:

1) Vi finner en basis i $\mathrm{Col}(A)$.

2) Bruker Gram-Schmidt prosessen for å finne en ortogonal basis $\{v_1, \dots, v_k\}$ i $W = \mathrm{Col}(A)$.

3) $\hat{b} = \mathrm{Proj}_W(b) = \frac{\langle b, v_1\rangle}{\langle v_1, v_1\rangle} v_1 + \dots + \frac{\langle b, v_k\rangle}{\langle v_k, v_k\rangle} v_k$.

4) Finner løsningene til $A\hat{x} = \hat{b}$.

<u>Merk:</u> $\hat{x}$ er entydig bestemt hvis og bare hvis $\mathrm{Nul}(A) = 0$.

<u>Merk:</u> Hvis $b \in \mathrm{Col}(A)$, da er $\hat{x} = x$ en løsning av $Ax = b$.

<u>Løsning 2</u>

Observer at $b - \hat{b} \in W^{\perp} = (\mathrm{Col}\, A)^{\perp} = \mathrm{Nul}\, A^T$.

Derson $A^T(b - \hat{b}) = 0$,

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

Likningen $\boxed{A^T A \hat{x} = A^T b}$

kalles den <u>normale likningen</u>.

Anta at $\hat{x}$ er en løsning av den normale likningen. Da får at
$A^T(A\hat{x} - b) = 0$, dvs. $A\hat{x} - b \in \mathrm{Nul}\, A^T = W^{\perp}$

Det følger at $A\hat{x} = \mathrm{proj}_W(b) = \hat{b}$.

<u>Oppsummering</u>: en minste kvadraters løsning av $Ax = b$ er det samme som en (vanlig) løsning av $A^T A x = A^T b$.

Løsning 3
(for matriser A med lineært uavhengige kolonner)
Betrakt QR-faktoriseringen

$$A = QR$$

(hvor Q har ortonormale kolonnene, R en øvre triangulær matrise med strengt positive tall langs diagonalen)
Betrakt den normale likningen

$$A^T A x = A^T b,$$

dvs.

$$(QR)^T QRx = (QR)^T b = R^T Q^T b$$

$$R^T Q^T Q Rx \longleftarrow \text{fordi } Q^T Q = I$$

$$= R^T R x$$

Merk at R (og derfor R^T) er invertibel, $\det R = r_{11} \cdots r_{nn} > 0$. Multipliseres med $(R^T)^{-1}$ og får

$$Rx = Q^T b$$

Oppsummering: en minste kvadraters løsning av $Ax = b$ er det samme som en løsning av

$$Rx = Q^T b \quad (\Leftrightarrow x = R^{-1} Q^T b)$$

Eksempel

$$A = \begin{array}{cc} a_1 & a_2 \\ \end{array}$$
$$A=\begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad b=\begin{pmatrix} 5 \\ 1 \\ 0 \end{pmatrix}$$

Vi vil finne en minste kvadraters løsning av $Ax=b$

Løsning 1

1) Vi ser at a_1, a_2 er lineært uavhengige.

2) Bruker Gram-Schmidt prosessen for å finne en basis i $Col(A)$:

$$v_1 = a_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

$$v_2 = a_2 - \frac{\langle a_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} - \frac{3}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}$$

3) $$\hat{b} = \mathrm{Proj}_{Col(A)}(b) = \frac{\langle b, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 + \frac{\langle b, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2$$

$$= \frac{6}{3} v_1 + \frac{8}{8} v_2 = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 2 \end{pmatrix}$$

4) $$A\hat{x} = \hat{b}$$

$$\begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{pmatrix} (\hat{x}_1, \hat{x}_2) = \begin{pmatrix} 4 \\ 0 \\ 2 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 3 & 4 \\ 1 & -1 & 0 \\ 1 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Da får vi at $\hat{x} = (1,1)$.

Løsning 2

$$A^T A x = A^T b$$

$$A^T A = \begin{pmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 3 & 11 \end{pmatrix}$$

$$A^T b = \begin{pmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \end{pmatrix} \begin{pmatrix} 5 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ 14 \end{pmatrix}$$

Systemet
$$\begin{pmatrix} 3 & 3 \\ 3 & 11 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 14 \end{pmatrix}$$

Da får vi $(x_1, x_2) = (1, 1)$.

Løsning 3

$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}$ er en ortogonal basis for $\mathrm{Col}(A)$ (fra løsning 1).

Vi normaliserer:

$$u_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad u_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

$$Q = [u_1, u_2] = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & 0 \end{pmatrix},$$

$$R = Q^T A = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{3} & \sqrt{3} \\ 0 & 2\sqrt{2} \end{pmatrix}$$

$$Rx = Q^T b$$

$$\begin{pmatrix} \sqrt{3} & \sqrt{3} \\ 0 & 2\sqrt{2} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{pmatrix}\begin{pmatrix} 5 \\ 1 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} 2\sqrt{3} \\ 2\sqrt{2} \end{pmatrix}$$

$$(x_1, x_2) = (1, 1)$$

Anta at A har ortogonale kolonnene.

Den minste kvadraters løsning av $Ax = b$ er gitt av

$$x_1 = \frac{\langle b, a_1 \rangle}{\langle a_1, a_1 \rangle}, \ldots, x_n = \frac{\langle b, a_n \rangle}{\langle a_n, a_n \rangle},$$

fordi

$$\hat{b} = \mathrm{Proj}_{\mathrm{Col}(A)}(b) = \frac{\langle b, a_1 \rangle}{\langle a_1, a_1 \rangle} a_1 + \ldots + \frac{\langle b, a_n \rangle}{\langle a_n, a_n \rangle} a_n.$$

<u>Eksempel</u>

$$A = \begin{array}{c} \begin{matrix} a_1 & a_2 \end{matrix} \\ \begin{pmatrix} 1 & 2 \\ -1 & 4 \\ 1 & 2 \end{pmatrix} \end{array}, \quad b = \begin{pmatrix} 3 \\ -1 \\ 5 \end{pmatrix}$$

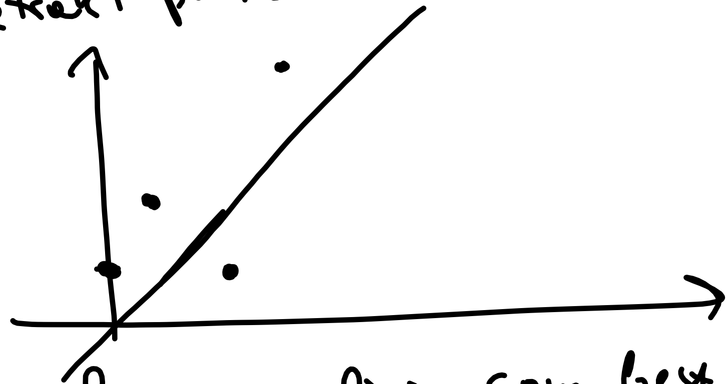
Vi ser at a_1 og a_2 er ortogonale.

Det følger at den minste kvadraters løsning av $Ax = b$ er

$$x_1 = \frac{\langle b, a_1 \rangle}{\langle a_1, a_1 \rangle} = \frac{8}{3}, \quad x_2 = \frac{\langle b, a_2 \rangle}{\langle a_2, a_2 \rangle} = \frac{12}{24} = \frac{1}{2}.$$

6.6 Lineære modeller

Eks:
Betrakt punktene $(0,1), (1,2), (2,1), (3,5)$ i $\mathbb{R}^2$:



Kan vi finne en linje som best approksimerer disse punktene?

Betrakt en linje $y = \beta_1 + \beta_2 x$

Først prøver vi å finne en linje som går gjennom alle 4 punktene:

$$\beta_1 + \beta_2 \cdot 0 = 1$$
$$\beta_1 + \beta_2 \cdot 1 = 2$$
$$\beta_1 + \beta_2 \cdot 2 = 1$$
$$\beta_1 + \beta_2 \cdot 3 = 5$$

Vi kan skrive dette som

$$\underbrace{\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}}_{X} \underbrace{\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}}_{\beta} = \underbrace{\begin{pmatrix} 1 \\ 2 \\ 1 \\ 5 \end{pmatrix}}_{y}$$

$$X\beta = y$$

Vi kan ikke løse dette, men vi kan finne minste kvadraters løsninger:
Vi skal bruke normale likninger (Løsning 2):

$$X^T X \beta = X^T y$$

$$X^T X = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix}$$

$$X^T y = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 1 \\ 5 \end{pmatrix} = \begin{pmatrix} 9 \\ 19 \end{pmatrix}$$

$$\begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 9 \\ 19 \end{pmatrix}$$

$$\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \frac{1}{20} \begin{pmatrix} 14 & -6 \\ -6 & 4 \end{pmatrix} \begin{pmatrix} 9 \\ 19 \end{pmatrix} = \frac{1}{20} \begin{pmatrix} 12 \\ 22 \end{pmatrix} = \begin{pmatrix} 0{,}6 \\ 1{,}1 \end{pmatrix}$$

Vi får linje $0{,}6 + 1{,}1x = y$.

Linjen kalles <u>minste kvadraters linje</u> for de gitte punktene:

$$\sum_{i=1}^{4} (0{,}6 + 1{,}1 x_i - y_i)^2 \quad \text{er minst mulig.}$$