13.7 Tilnærmet løsning av likningssystemer (Lay 6.5, 6.6 og 6.8)

Teorem 1

La A være en $(m \times n)$-matrise, og la $\vec{b} \in \mathbb{R}^m$.
Anta at vi vil løse problemet

$$A\vec{x} \approx \vec{b} \qquad (1)$$

med ukjent $\vec{x} \in \mathbb{R}^n$ i den forstand at vi vil finne en vektor $\vec{x} \in \mathbb{R}^n$ som gjør avstanden

$$\|\vec{b} - A\vec{x}\|$$

minst mulig (gjerne 0 også!). En vektor $\hat{x} \in \mathbb{R}^n$ har denne avstandsminimaliserende egenskapen hviss

$$(A^T A)\hat{x} = A^T \vec{b} \qquad (2)$$

Det fins alltid løsninger $\vec{x}$ av (2). Hvis søylene i A er lineært uavhengige, er løsningen $\hat{x}$ entydig.

Bevis La $\mathcal{U} \subseteq \mathbb{R}^m$ være bildet til A, og la $\vec{x} \in \mathbb{R}^n$.
Anta at
$$A^T(A\vec{x}) = A^T\vec{b}$$
Dette er ekvivalent med at vektorene
$$A\vec{x} \quad \text{og} \quad \vec{b}$$
har samme indreprodukt med alle søylevektorene i A. Disse søylevektorene spenner ut $\mathcal{U}$, så det følger at
$$\text{proj}_{\mathcal{U}}(A\vec{x}) = \text{proj}_{\mathcal{U}}\vec{b}$$
Siden $A\vec{x} \in \mathcal{U}$, har vi
$$\text{proj}_{\mathcal{U}}(A\vec{x}) = A\vec{x}$$
Altså
$$A\vec{x} = \text{proj}_{\mathcal{U}}\vec{b} \qquad (*)$$
Fra teorem 13.3.2 vet vi at
$$\|\vec{b} - \text{proj}_{\mathcal{U}}\vec{b}\| < \|\vec{b} - \vec{y}\|$$
for alle $\vec{y} \in \mathcal{U}$. Innsatt med $(*)$ gir dette
$$\|\vec{b} - A\vec{x}\| < \|\vec{b} - \vec{y}\|$$
for alle $\vec{y} \neq A\vec{x}$ i $\mathcal{U}$. Altså har $\vec{x}$ den avstandsminimaliserende egenskapen.

Entydigheten: Hvis søylene i A er lineært uavhengige, er A injektiv som avbildning fra $\mathbb{R}^n$ til $\mathbb{R}^m$ (ved $\vec{x} \mapsto A\vec{x}$). Altså fins en entydig $\vec{x} \in \mathbb{R}^n$ slik at $(*)$ holder. $\square$

Kommentarer:

- Systemet (2) kalles normallikningen til problemet (1).
  Løsninger av (2) kalles ofte minste kvadraters løsninger
- Alternativ notasjon: $A\vec{x} \approx \vec{b}$ skrives ofte

$$\vec{b} = A\vec{x} + \vec{e} \qquad (**)$$

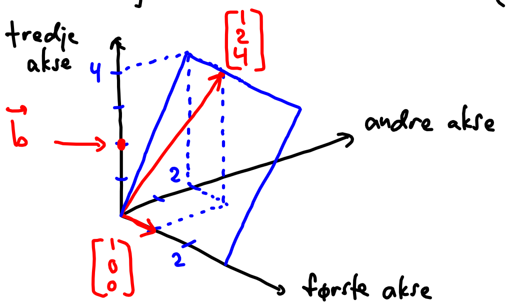
der $\vec{e}$ er en feilvektor som skal minimaliseres

$\vec{b}$ kalles observasjonsvektoren
$\vec{x}$ " parametervektoren
A " designmatrisen

eks. Skal finne minste kvadraters løsninger til $A\vec{x} = \vec{b}$,
dvs. løse $A\vec{x} \approx \vec{b}$, der

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 4 & 0 \end{bmatrix} \quad \text{og} \quad \vec{b} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Løsn. Avbildning: $A : \mathbb{R}^2 \to \mathbb{R}^3$ ved $A(\vec{x}) = A \cdot \vec{x}$. Figur:



Ser at systemet

$$\begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 4 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

er selvmotsigende.

Vi får

$$A^T A = \begin{array}{ccc|cc} & & & 1 & 1 \\ & & & 2 & 0 \\ & & & 4 & 0 \\ \hline 1 & 2 & 4 & 21 & 1 \\ 1 & 0 & 0 & 1 & 1 \end{array} = \begin{bmatrix} 21 & 1 \\ 1 & 1 \end{bmatrix}$$

$$A^T\vec{b} = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 0 \end{bmatrix}$$

Så normallikningen $(A^T A)\vec{x} = A^T\vec{b}$ blir altså

$$\begin{bmatrix} 21 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 0 \end{bmatrix}$$

Løser:
$$\begin{cases} 21x_1 + x_2 = 8 & \text{I} \\ x_1 + x_2 = 0 & \text{II} \end{cases}$$

II sier $x_2 = -x_1$

I sier da $20x_1 = 8$

$$x_1 = \frac{8}{20} = \frac{2}{5}$$

Løsning: $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{2}{5} \\ -\frac{2}{5} \end{bmatrix}$

Sjekk:
$$A\vec{x} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 4 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2/5 \\ -2/5 \end{bmatrix} = \begin{bmatrix} 0 \\ 4/5 \\ 8/5 \end{bmatrix}$$

$$\left\langle \begin{bmatrix} 0 \\ 4/5 \\ 8/5 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} - \begin{bmatrix} 0 \\ 4/5 \\ 8/5 \end{bmatrix} \right\rangle = \left\langle \begin{bmatrix} 0 \\ 4/5 \\ 8/5 \end{bmatrix}, \begin{bmatrix} 0 \\ -4/5 \\ 2/5 \end{bmatrix} \right\rangle$$

$$= 0 - \frac{16}{25} + \frac{16}{25} = 0 \quad \text{ok}$$

$\boxed{\text{13.8 Minste kvadraters metode}}$ (Lay 6.5, 6.8)

(klassisk utgave, spesialtilfelle av 13.7)

Gitt m punkter

$$(t_1, s_1), \ldots, (t_m, s_m) \in \mathbb{R}^2$$

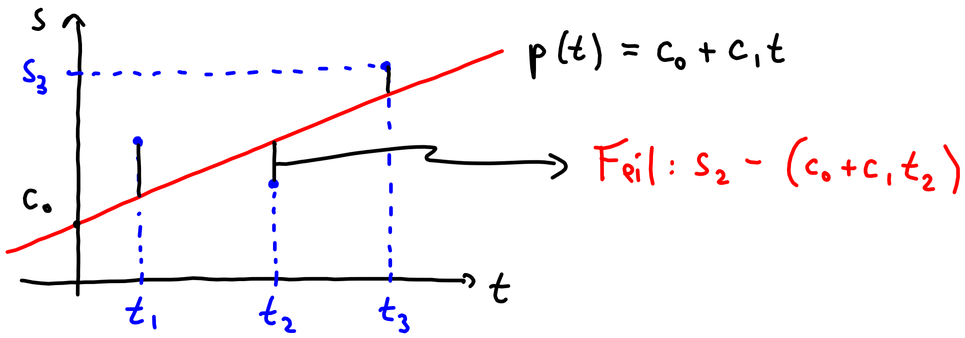
der $m > 2$ og ikke alle tallene $t_1, \ldots, t_m$ er like. Vi ønsker å finne c_0 og c_1 slik at funksjonen

$$p(t) = c_0 + c_1 t \qquad \text{(lineær funksjon)}$$

gjør summen

$$S = \sum_{i=1}^{m} \left[s_i - (c_0 + c_1 t_i) \right]^2$$

av kvadratavvikene minst mulig.



Teorem 1

I situasjonen ovenfor finner vi koeffisientene c_0 og c_1 som den entydige løsningen

$$\hat{\vec{x}} = \vec{x} = \begin{bmatrix} c_0 \\ c_1 \end{bmatrix}$$

av likningen $(A^T A)\vec{x} = A^T \vec{b}$, der

$$A = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} \quad \text{og} \quad \vec{b} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_m \end{bmatrix}$$

Bevis Kravene for at $p(t_i) = s_i$ for $i = 1, \dots, m$ er

$$\begin{cases} p(t_1) = c_0 + c_1 t_1 = s_1 \\ p(t_2) = c_0 + c_1 t_2 = s_2 \\ \quad \vdots \\ p(t_m) = c_0 + c_1 t_m = s_m \end{cases}$$

Som kan skrives

$$\left.\begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} \cdot \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_m \end{bmatrix} \right\} \text{dette vil "typisk" være selvmotsigende}$$

dvs. $A\vec{x} = \vec{b}$, der A, $\vec{b}$ og $\vec{x}$ er som i teoremet.

Ved teorem 13.7.1 (tilnærmet løsning av likningssystemer) vet vi at hvis $\vec{x}$ løser likningen

$$(A^T A)\vec{x} = A^T \vec{b}$$

tilhørende problemet $A\vec{x} \approx \vec{b}$, så har vi

$$\| A\vec{x} - \vec{b} \|^2 < \| \vec{y} - \vec{b} \|^2$$

for alle $\vec{y} \neq A\vec{x}$ som ligger i bildet til A.

Siden alle mulige kombinasjoner $\vec{x} = \begin{bmatrix} c_0 \\ c_1 \end{bmatrix}$ av koeffisienter tilsvarer et punkt i dette bildet, følger at løsningen $\hat{\vec{x}}$ minimaliserer

$$\left\| A \cdot \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} - \begin{bmatrix} s_1 \\ \vdots \\ s_m \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} \cdot \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} - \begin{bmatrix} s_1 \\ \vdots \\ s_m \end{bmatrix} \right\|^2$$

$$= \left\| \begin{bmatrix} c_0 + c_1 t_1 \\ \vdots \\ c_0 + c_1 t_m \end{bmatrix} - \begin{bmatrix} s_1 \\ \vdots \\ s_m \end{bmatrix} \right\|^2$$

$$= \left\| \begin{bmatrix} (c_0 + c_1 t_1) - s_1 \\ \vdots \\ (c_0 + c_1 t_m) - s_m \end{bmatrix} \right\|^2$$

$$= \sum_{i=1}^{m} \left[(c_0 + c_1 t_i) - s_i \right]^2 = S$$

Entydigheten av løsningen $\hat{\vec{x}}$ følger fordi ikke alle tallene t_i er like, slik at søylene i A er lineært uavhengige (jf. teoremet fra 13.7) □

Vektet minste kvadraters metode

I anvendelser kan det hende at man vil vektlegge noen elementer i datavektoren $\vec{b}$ mer enn andre, f. eks. pga. måleusikkerhet.

Istedenfor å løse

$$A\vec{x} \approx \vec{b}$$

kan man da bruke teorem 13.7.1 på problemet

$$W(A\vec{x}) \approx W\vec{b} \qquad (*)$$

der W er en diagonal vektmatrise med ikke-negative tall $w_1, \dots, w_m$ langs diagonalen. Vi har

$$\left(W(A\vec{x})\right)_i - \left(W\vec{b}\right)_i = w_i \left(A\vec{x}\right)_i - w_i \vec{b}_i$$

$$= w_i \left[(A\vec{x})_i - \vec{b}_i \right],$$

så dette betyr at vi ønsker å minimalisere kvadratsummen av vektede avvik, der avvik fra $\vec{b}_i$ vektes med w_i.

Likningen (*) kan skrives

$$(WA)\vec{x} \approx W\vec{b}$$

Erstatter vi A med WA og $\vec{b}$ med $W\vec{b}$ i normallikningen $(A^TA)\vec{x} = A^T\vec{b}$, får vi

$$(WA)^T (WA)\vec{x} = (WA)^T (W\vec{b})$$

Dette kan så løses med hensyn på $\vec{x}$, på vanlig måte.