TMV141

MINSTAKVADRATMETODEN (Least-squares problem)

L18

I många problem så vill man hitta en lösning till problem av typen $Ax=b$ med $A \in \mathbb{R}^{m \times n}$, $m>n$ (ofta $m>>n$).
Dessa system är ÖVERBESTÄMDA och saknar i allmänhet lösning.
Istället försöker man hitta en lösning $\bar{x}$ till problemet $A\bar{x}=\bar{b}$ där $\|b-\bar{b}\|$ är så litet som möjligt, dvs $\bar{b} = \mathrm{Proj}_{Col(A)}(b)$.

__Def__ Om $A \in \mathbb{R}^{m \times n}$ och $b \in \mathbb{R}^m$ så är MINSTAKVADRATLÖSNINGEN till $Ax=b$ en vektor $\bar{x} \in \mathbb{R}^n$, sådant att:

$$\|b-A\bar{x}\| \leq \|b-Ax\|, \ \forall x \in \mathbb{R}^n$$

Härledning av lösningsmetod

- Det måste gälla att $\bar{b} \in Col(A)$ för att $A\bar{x}=\bar{b}$ skall ha en lösning.
- För att $\bar{b}$ skall vara ~~en lösning~~ så nära b som möjligt krävs att deras skillnad är ortogonal mot varje vektor i ~~[illegible]~~ $v \in Col(A)$ dvs. $v \cdot (\bar{b}-b)=0$, $\forall v \in Col(A)$
  (detta enligt satsen om ortogonal uppdelning)
- Det räcker med att testa vektorer v som spänner $Col(A)$ nämligen a_i, $i=1,..,n$

  Alltså: $a_i^T(\bar{b}-b)=0$, $i=1,..,n$

  Det är samma sak som att: $A^T(\bar{b}-b)=0$ dvs $A^T\bar{b}=A^Tb$.

  Eftersom $\bar{b}=A\bar{x}$ så är det ekvivalent med att

$$A^TA\bar{x}=A^Tb$$

<u>Sats 6.5.13</u> Minstakvadrat-lösningarna till ekvationen $Ax=b$ är samma som lösningarna till den (konsistenta) normaliserade ekvationen $\underline{A^TAx = A^Tb}$.

<u>Sats 6.5.14</u> Minstakvadratproblemet till $Ax=b$ har en unik lösning $\bar{x}$ omm kolonnerna i A är linjärt oberoende.
Då och endast då är A^TA inverterbar och den unika lösningen ges av $\bar{x} = (A^TA)^{-1}A^Tb$.

<u>Ex</u>
Hitta minsta kvadratlösning till ek-systemet $Ax=b$.

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$$

$$A^TA = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$A^Tb = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \end{bmatrix}$$

Alt 1
$$\begin{aligned} 2x_1 + x_2 &= 6 \\ x_1 + 2x_2 &= 6 \end{aligned} \Rightarrow \begin{aligned} x_1 &= 2 \\ x_2 &= 2 \end{aligned} \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

Eller: Alt 2
A^TA är inverterbar och 2×2

$$(A^TA)^{-1} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}^{-1} = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$\bar{x} = (A^TA)^{-1}A^Tb = \underbrace{\frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 6 \\ 6 \end{bmatrix}}_{1/3\begin{bmatrix} 6 \\ 6 \end{bmatrix}} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

TILLÄMPNING AV MINSTA KVADRATSMETODEN
– LINJÄRA MODELLER –

En linjär modell beskrivs av en: ~~designmatris X~~

- DESIGN MATRIS X (som beskriver hur modellen byggs upp)
- OBSERVATIONS VEKTOR $\bar{y}$ (som beskriver observationer)
- PARAMETER VEKTOR β (som beskriver parameter i modellen som man vill optimera med avseende på).

LINJÄR MODELL: $X\beta = \bar{y}$ $\quad (Ax=b)$

Antag att vi har n parametrar och gör m observationer. Då gäller att X är en $m \times n$ matris och vi får ek.-sys. $X\beta = \bar{y}$

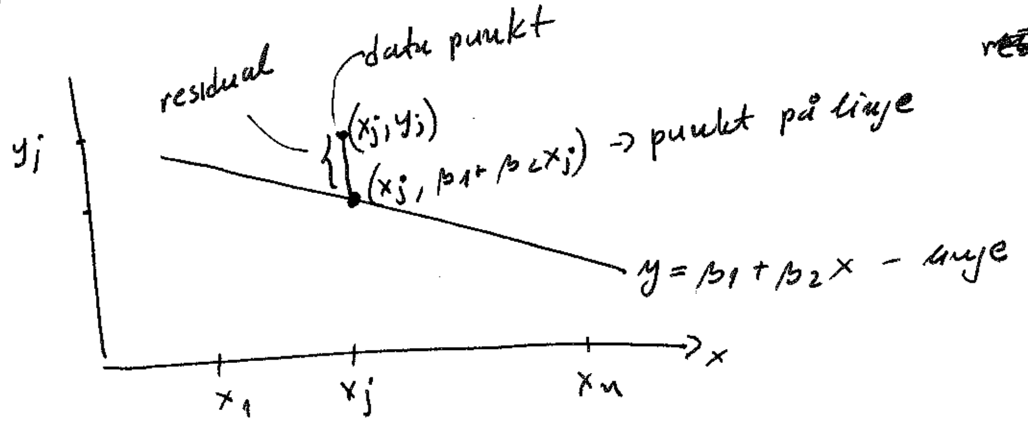
NORMALEKV. blir då $X^TX\beta = X^T\bar{y}$

En vanlig exampel är anpassning av en rät linje till m givna punkter i planet $(x_1, y_1), \dots, (x_m, y_m)$. Man letar efter en linje $y = \beta_1 + \beta_2 x$.

estimation of y-intercept

Här är alltså obekanta β_1, β_2 och $X = \begin{pmatrix} 1 & x_1 \\ 1 & \vdots \\ 1 & \\ \vdots & \\ 1 & x_m \end{pmatrix}$, $\bar{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$

~~residual = yobserved~~ y.



Minstkvadrat linje är linje $y = \beta_1 + \beta_2 x$ som minimisera "sum of squares of the residuals"

residual = $y_{observed} - y_{predicted}$

β_1, β_2 – linjär regression koeficienter.

Parameter vektor β minimisera avstånd mellan $X\beta$ och y

Ex Bestäm $y = \beta_1 + \beta_2 x$ som ~~de~~ bäst aproximerar punkter

$$\underbrace{(1,0)}_{x_1\ y_1},\ \underbrace{(2,1)}_{x_2\ y_2},\ \underbrace{(3,1)}_{x_3\ y_3}.$$

$$X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}$$

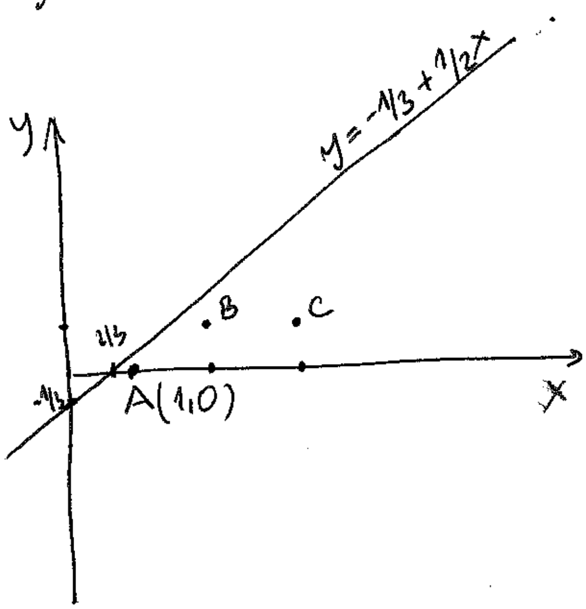
$$\bar{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

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

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

$$X^T X \beta = X^T \bar{y} \Leftrightarrow \begin{pmatrix} 3 & 6 \\ 6 & 14 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \end{pmatrix} \Rightarrow \begin{matrix} \beta_1 = -1/3 \\ \beta_2 = 1/2 \end{matrix}$$

Linjen som vi söker är: $y = -1/3 + 1/2 x$



A(1,0)
B(2,1)
C(3,1)