

MINSTAKVADRATMETODEN (Least-squares problem)

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 $m \gg n$).

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 $\bar{b} = \text{Proj}_{\text{Col}(A)}(b)$.

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:

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

Härledning av lösningsmetod

- Det måste gälla att $\bar{b} \in \text{Col}(A)$ 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 ~~Col(A)~~ $u \in \text{Col}(A)$ dvs. $u \cdot (\bar{b} - b) = 0$, $\forall u \in \text{Col}(A)$
(detta enligt satsen om ortogonal uppdelning)

- Det räcker mest att testa vektorer u som spänner $\text{Col}(A)$ nämligen a_i , $i=1, \dots, n$

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

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

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

Sats 6.5.13 ~~Minsta~~ Minstakvadrat - Lösningarna till ekvationen $Ax=b$ är samma som lösningarna till den (konsistenta) normaliserade ekvationen $A^T A x = A^T b$.

Sats 6.5.14 Minstakvadratproblemet till $Ax=b$ har en unik lösning $\bar{x}$ om och endast om kolumnerna i A är linjärt oberoende.
Då och endast då är $A^T A$ invertierbar och den unika lösningen ges av $\bar{x} = (A^T A)^{-1} A^T b$.

Ex ~~lös~~ Hitta minsta kvadratlösning ~~lös~~ till ek-systemet $Ax=b$.

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

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

$$A^T b = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \end{bmatrix} \quad \text{Allt} \quad \begin{matrix} 2x_1 + x_2 = 6 \\ x_1 + 2x_2 = 6 \end{matrix} \Rightarrow \begin{matrix} x_1 = 2 \\ x_2 = 2 \end{matrix} \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

Eller: Alt 2
 $A^T A$ är invertierbar och 2×2

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

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

TILLÄMPNING AV MINSTA KVADRATSMETODEN

- LINJÄRA MODELLER -

En linjär modell beskrivs av en: ~~design matrix~~

- DESIGN MATRIX 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: $X\beta = \bar{y}$ ($Ax=b$)

Antag att vi har n parametrar och gör m observationer.

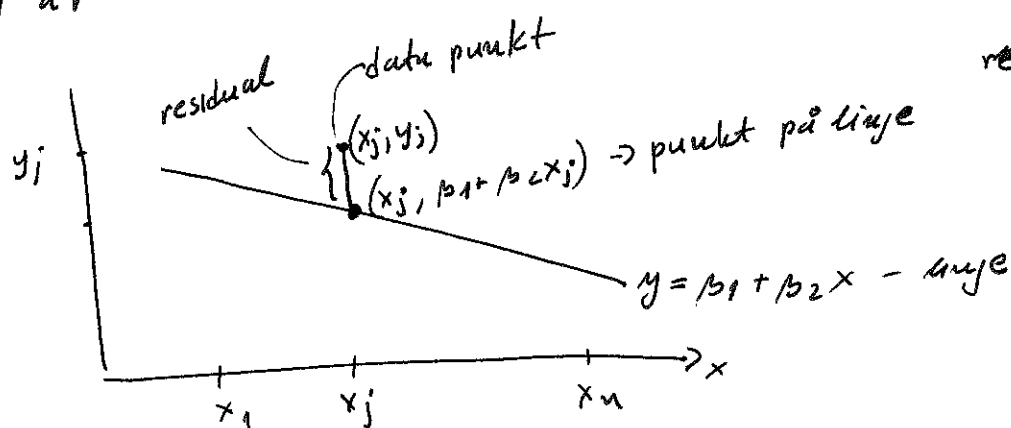
Då gäller att X är en $m \times n$ matris och vi får ek. sys. $X\beta = \bar{y}$

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

En vanlig exempel 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 $y = \beta_1 + \beta_2 x$.

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



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

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

β_1, β_2 - linjära regressions koefficienter.

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

Ex Bestäm $y = \beta_1 + \beta_2 x$ som bäst approximerar punkterna

$$(1,0), (2,1), (3,1).$$
$$\begin{array}{cc|cc|cc} x_1 & y_1 & x_2 & y_2 & x_3 & y_3 \end{array}$$

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

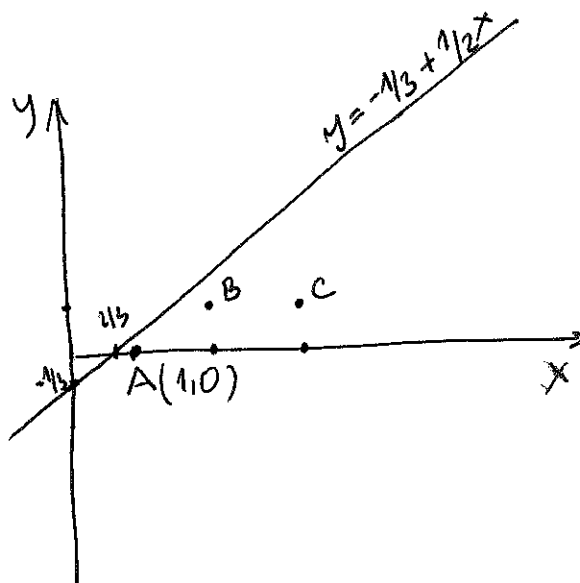
$$\bar{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

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

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

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

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



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