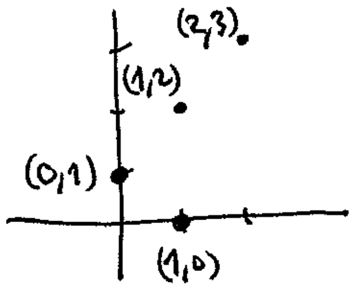
MINSTA KVADRATMETODEN

- ger approximativ lösning till ekv.-system som saknar lösning, t ex överbestämda ekv. system

<u>Ex</u>) Hitta en linje som går genom datapunkterna $(0,1)$, $(1,0)$, $(1,2)$, $(2,3)$!



Att punkterna ligger på linjen $l: y = a + bx$ innebär att ekv. systemet

$$(*)\quad \begin{cases} a + b\cdot 0 = 1 & (0,1) \in l \\ a + b\cdot 1 = 0 & (1,0) \in l \\ a + b\cdot 1 = 2 & (1,2) \in l \\ a + b\cdot 2 = 3 & (2,3) \in l \end{cases}$$

har lösning.

Vi kan skriva $(*)$ på matrisform som

$$\underbrace{\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} 1 \\ 0 \\ 2 \\ 3 \end{bmatrix}}_{\mathbf{b}}$$

<u>Obs</u> från bilden (eller ekv. systemet $(*)$) att lösning saknas – det finns ingen linje genom alla punkterna.

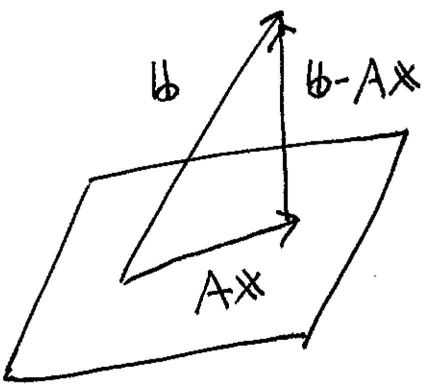
- Minsta kvadratmetoden går ut på att hitta $\mathbf{x}$ som minimerar $|A\mathbf{x} - \mathbf{b}|$

Antag att vi har ett ekvationssystem

(*) $A\mathbf{x} = \mathbf{b}$

där A är en $m\times n$-matris och $\mathbf{b}\in\mathbb{R}^m$.

<u>Obs</u> $A\mathbf{x} \in$ Kolonnrummet till A som är ett delrum till $\mathbb{R}^m$. (*) har lösning om $\mathbf{b} \in \text{Kolonn}(A)$.



<u>Ex</u>, I exemplet ovan är $\text{Kolonn}(A) = \text{span}((1,1,1,1),(0,1,1,2))$
$\mathbf{b} = (1,0,2,3) \notin \text{Kolonn}(A)$

<u>Minns</u> att $|A\mathbf{x}-\mathbf{b}|$ är minimal om $A\mathbf{x}-\mathbf{b}$ är ortogonal mot $\text{Kolonn}(A)$.

För att minimera $|A\mathbf{x}-\mathbf{b}|$ söker vi alltså $\mathbf{x}$ som uppfyller att $\mathbf{a}_j\cdot(A\mathbf{x}-\mathbf{b}) = 0 \quad j=1,\dots,n,$ (**)
där $\mathbf{a}_1,\dots,\mathbf{a}_n$ är kolonnerna i A, dvs $A = \begin{bmatrix} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{bmatrix}$. Detta (**) kan vi uttrycka som

$$\begin{bmatrix} \text{---} & \mathbf{a}_1 & \text{---} \\ & \vdots & \\ \text{---} & \mathbf{a}_n & \text{---} \end{bmatrix} \begin{bmatrix} | \\ A\mathbf{x}-\mathbf{b} \\ | \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{dvs } A^T(A\mathbf{x}-\mathbf{b}) = \mathbf{0}$$

Vi vill alltså hitta $\mathbf{x}\in\mathbb{R}^n$ som uppfyller

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

Denna ekvation kallas normalekvationen.

Proposition 1

Om $\bar{x}$ är en lösning till

$$A^T A x = A^T b \quad (*)$$

så är $|A\bar{x} - b| = \min_{x \in \mathbb{R}^n} |Ax - b|$

Bevis

Minns $|u+v|^2 = |u|^2 + 2u\cdot v + |v|^2$

Tag $x \in \mathbb{R}^n$. Då

$$\underline{\underline{|Ax - b|^2}} = |\underbrace{(Ax - A\bar{x})}_{A(x-\bar{x})} + (A\bar{x} - b)|^2 =$$

$$|A(x-\bar{x})|^2 + 2\left(A(x-\bar{x}) \cdot (A\bar{x} - b)\right) + |A\bar{x} - b|^2 =$$

Obs om identifierar $u \in \mathbb{R}^n$ med kolonnmatrisen $\begin{bmatrix} | \\ u \\ | \end{bmatrix}$ och skalärer med 1×1-matriser kan vi skriva $u \cdot v = u^T v$.

$$|A(x-\bar{x})|^2 + 2\underbrace{(A(x-\bar{x}))^T (A\bar{x} - b)}_{(x-\bar{x})^T \underbrace{A^T(A\bar{x}-b)}_{=0 \text{ ty } \bar{x} \text{ uppfyller } (*)}} + |A\bar{x} - b|^2 =$$

$$\underbrace{|A(x-\bar{x})|^2}_{\geq 0} + |A\bar{x} - b|^2 \geq \underline{\underline{|A\bar{x} - b|^2}}$$

Slutsats $|Ax - b| \geq |A\bar{x} - b|$

(med likhet omm $A(x-\bar{x}) = 0$), dvs

$|A\bar{x} - b| = \min_{x \in \mathbb{R}^n} |Ax - b|$. $\square$

Ex) Hitta en linje som är anpassad efter datapunkterna i exemplet ovan i minsta kvadratbemärkelse!

Vi vill lösa

$$A^T A \hat{x} = A^T b \quad (*)$$

där $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$ $b = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 3 \end{bmatrix}$.

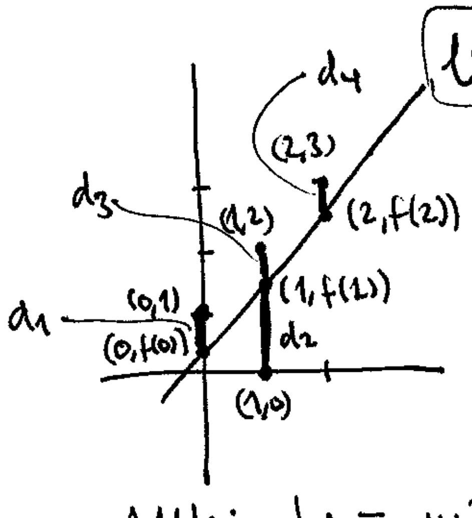
Nu är $A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 4 & 6 \end{bmatrix}$

Obs $A^T A$ är inverterbar. Man kan kolla att inversen är $(A^T A)^{-1} = \frac{1}{4} \begin{bmatrix} 3 & -2 \\ -2 & 2 \end{bmatrix}$.

Det följer att $(*)$ har lösning

$$\begin{bmatrix} a \\ b \end{bmatrix} = \bar{x} = (A^T A)^{-1} A^T b = \frac{1}{4} \begin{bmatrix} 3 & -2 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 2 \\ 3 \end{bmatrix} = \ldots = \begin{bmatrix} 1/2 \\ 1 \end{bmatrix}$$

Detta motsvarar linjen $l : y = 1/2 + x$



I vilken bemärkelse är l anpassad till punkterna?

Obs $A\bar{x} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1/2 \\ 1 \end{bmatrix} = \begin{bmatrix} f(0) \\ f(1) \\ f(1) \\ f(2) \end{bmatrix}$ $b = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 3 \end{bmatrix}$

Alltså $|A\bar{x} - b|^2 = |(f(0)-1, f(1)-0, f(1)-2, f(2)-3)|^2 =$

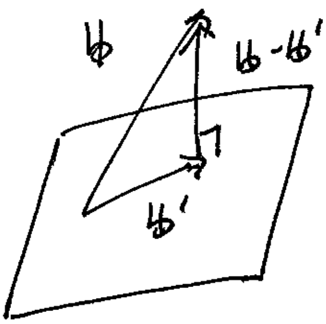
$|f(0)-1|^2 + |f(1)-0|^2 + |f(1)-2|^2 + |f(2)-3|^2 = d_1^2 + d_2^2 + d_3^2 + d_4^2$

Att $|A\hat{x} - b|$ minimeras betyder alltså att (roten ur) summan av kvadraterna av de vertikala avstånden d_i minimeras.

Proposition 2

Ekvationen $A^TA\mathbb{x}=A^T\mathbb{b}$ har alltid lösning.

Bevis: Låt $\mathbb{b}'$ vara den ortogonala projektionen av $\mathbb{b}$ på Kolonn(A). Då är $\mathbb{b}'-\mathbb{b}$ ortogonal mot Kolonn$(A) = \text{span}(a_1, \dots, a_n)$



Kolonn(A) om $A=\begin{bmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{bmatrix}$.

Eftersom $\mathbb{b}'\in$ Kolonn(A) har ekvationen $A\mathbb{x}=\mathbb{b}'$ en lösning. Låt $\mathbb{x}$ vara en sådan lösning. Då

$$A^TA\mathbb{x}-A^T\mathbb{b}=A^T(A\mathbb{x}-\mathbb{b})=A^T\Big[\underbrace{(A\mathbb{x}-\mathbb{b}')}_{=\mathbb{0}\text{ enligt antagande}}+(\mathbb{b}'-\mathbb{b})\Big]=$$

$$A^T(\mathbb{b}'-\mathbb{b})=$$

$$\begin{bmatrix} — & a_1 & — \\ & \vdots & \\ — & a_n & — \end{bmatrix}\begin{bmatrix} | \\ \mathbb{b}'-\mathbb{b} \\ | \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbb{0},$$

$(\mathbb{b}'-\mathbb{b})$ ortogonal mot $a_1, \dots, a_n$

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

Alltså har vi visat att $A^TA\mathbb{x}=A^T\mathbb{b}$ alltid är lösbar.

□

4/10 16

Proposition 3 (A $m \times n$)

Antag att $A\mathbb{x}=\mathbb{b}$ har (åtminstone) en lösning.

Då

$\mathbb{x}$ är en lösn. till $A\mathbb{x}=\mathbb{b}$ $\Longleftrightarrow$ $\mathbb{x}$ är en lösn till $A^TA\mathbb{x}=A^T\mathbb{b}$

För beviset behöver vi följande lemma.

Lemma $\text{Noll}(A)=\text{Noll}(A^TA)$

Bevis: **Obs** • om $A\mathbb{x}=\mathbb{0}$, så följer att $A^TA\mathbb{x}=\mathbb{0}$.

Alltså $\text{Noll}(A)=\{A\mathbb{x}=\mathbb{0}\}\subseteq\{A^TA\mathbb{x}=\mathbb{0}\}=\text{Noll}(A^TA)$

• om $A^TA\mathbb{x}=\mathbb{0}$, så följer att

$$0=\mathbb{x}^TA^TA\mathbb{x}=(A\mathbb{x})^TA\mathbb{x}=$$

$$(A\mathbb{x})\cdot(A\mathbb{x})=|A\mathbb{x}|^2$$

Alltså $|A\mathbb{x}|=0\Rightarrow A\mathbb{x}=\mathbb{0}$

Alltså $\text{Noll}(A^TA)\subseteq\text{Noll}(A)$.

Slutsats: $\text{Noll}(A)=\text{Noll}(A^TA)$ $\square$

Bevis Prop 3:

Minns att om $A\mathbb{x}=\mathbb{b}$ har en lösning så är mängden av alla lösningar

$$\mathbb{x}_{p_1}+\text{Noll}(A)=:\Pi_1,$$

där $\mathbb{x}_{p_1}$ är en lösning (partikulärlösning) till $A\mathbb{x}=\mathbb{b}$. Geometriskt är Π_1 ett plan av dim nolldim(A) i $\mathbb{R}^n$.

4/10 16

På samma sätt är mängden av lösningar till $A^T A \mathbf{x} = A^T \mathbf{b}$

$$\mathbf{x}_{p_2} + \mathrm{Null}(A^T A) \underset{\text{Lemma}}{=} \mathbf{x}_{p_2} + \mathrm{Null}(A) =: \Pi_2$$

Även Π_2 är ett nulldim(A)-dimensionellt plan i $\mathbb{R}^n$ (parallellt med Π_1).

Obs att om $\mathbf{x}$ löser $A\mathbf{x} = \mathbf{b}$, så klart att $\mathbf{x}$ löser $A^T A \mathbf{x} = A^T \mathbf{b}$. Det följer att $\Pi_1 \subseteq \Pi_2$.

Eftersom Π_1 och Π_2 är plan i $\mathbb{R}^n$ av samma dimension medför $\Pi_1 \subseteq \Pi_2$ att $\Pi_1 = \Pi_2$, d v s $\mathbf{x}$ löser $A\mathbf{x} = \mathbf{b}$ omm $\mathbf{x}$ löser $A^T A \mathbf{x} = A^T \mathbf{b}$. $\square$