

Idag: *

- * Minstakvadratmetoden 6.5
- * Linjär regression 6.6
- * Intro till Datorövning:
- * Skalärprodukt rum 6.7
(fortsätter nästa vecka)

6.5 Ett linjärt ekvationssystem $Ax = b$ är inkonsistent om $b \notin \text{Col}(A)$.

$A \in \mathbb{R}^{m \times n}$
 $b \in \mathbb{R}^m, x \in \mathbb{R}^n$

$\nearrow b$

Då finns ingen lösning, men vi kan göra bästa möjliga genom att hitta $x \in \mathbb{R}^n$ så att $\|Ax - b\|$ blir minimal.

Obs: vi minimerar $\|Ax - b\| = \sqrt{\sum_{i=1}^m (Ax - b)_i^2}$ därför "minsta kvadrat".

Satsen om "bästa approximation"
(Sats 9) säger att bästa approx.
till b i $\text{Col}(A)$ är

$$\hat{b} = \text{proj}_{\text{Col}(A)}(b).$$

Men $\hat{b} \in \text{Col}(A)$ så det finns
 $\hat{x} \in \mathbb{R}^n$ så att

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

Jalur berakna $\hat{X}^Z$.

Vi vet att $b - \hat{b} \perp \text{Col}(A) = \text{span}\{a_1, \dots, a_T\}$
där a_j är kolonnerna i A .

DWS $a_j^T (b - \hat{b}) = 0 \quad \forall a_j$

$$\Leftrightarrow a_j^T (b - A\hat{x}) = 0 \quad \forall a_j$$

$$\Leftrightarrow A^T(b - A\hat{x}) = 0$$

$$\Leftrightarrow A^T A \hat{x} = A^T b$$

Alltså: vi ska lösa de så kallade
normalekvationerna:

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

för att få $\hat{x}$.

Obs: (*) kan ha flera lösningar
om det finns fria variabler.

Def Låt $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$.

En vektor $\hat{x} \in \mathbb{R}^n$ kallas minsta-
kvadratlösning till $Ax = b$

om

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

Sats 13 Mängden av minsta-
kvadratlösningar till $Ax = b$
är lika med lösningsmängden
till normalekvationerna

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

Vi har visat att varje minstakvadrat-
lösning $\hat{x}$ är lösning till normal-
ekv. Återstår att visa att varje lösning
till normalekv. är minstakvadrat-
lösning, dvs minimerar residualen.
Vi hoppar över detta.

Obs: $A^T A$ är symmetrisk, ty

$$(A^T A)^T = A^T (A^T)^T = A^T A$$

Obs: 1. Vad menas med minsta kvadratlösning
till $Ax = b$? Svar: $\|A\hat{x} - b\| \leq \|Ax - b\| \quad \forall x \in \mathbb{R}^n$

2. Hur beräknas man minsta kvadrat-
lösning? Svar: $A^T A x = b$

6.6 linjär regression

Datorövning vecka 6:

Nortons modell för kryppning

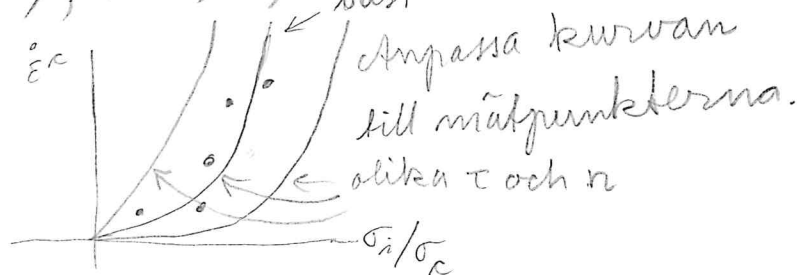
$$\dot{\epsilon}^c = \frac{1}{\tau} \left(\frac{\sigma}{\sigma_c} \right)^n \quad (\text{icke linjär modell})$$

$\dot{\epsilon}^c$ kryptöjningshastighet och spänningen

σ . σ_c är en referensspänning, som man väljer. experiment: lägg på σ och mät $\dot{\epsilon}^c$.

Two parametrar τ och n ska bestämmas från mätpunkter

$$\left(\frac{\sigma_i}{\sigma_c}, \dot{\epsilon}_i^c \right), i = 1, \dots, N \text{ bäst}$$

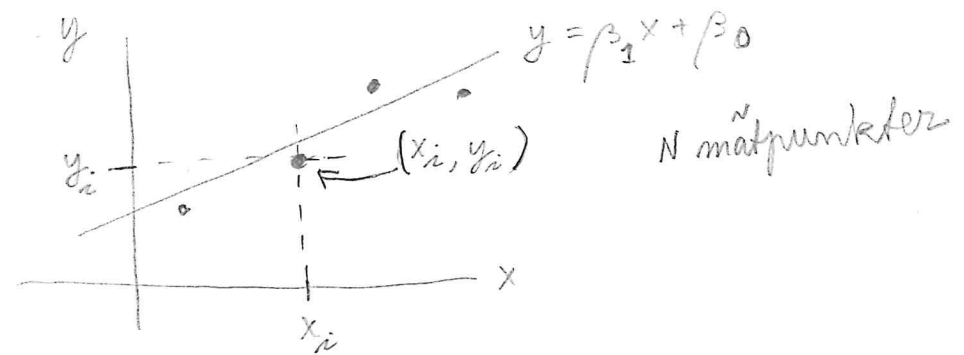


Icke linjär funktion, transformera till linjär funktion genom att ta logaritm:

$$\ln(\dot{\epsilon}^c) = \ln\left(\frac{1}{\tau} \left(\frac{\sigma}{\sigma_c}\right)^n\right)$$

$$\underbrace{\ln \dot{\epsilon}^c}_{=y} = \underbrace{n \ln\left(\frac{\sigma}{\sigma_c}\right)}_{=x} - \underbrace{\ln(\tau)}_{=\beta_0}$$

$$y = \beta_1 x + \beta_0 \quad (\text{linjär modell})$$



$$\begin{cases} y_1 = \beta_1 x_1 + \beta_0 \\ y_2 = \beta_1 x_2 + \beta_0 \\ \vdots \\ y_N = \beta_1 x_N + \beta_0 \end{cases} \quad \begin{array}{l} N \text{ ekv.} \\ 2 \text{ obekanta } \beta_0, \beta_1 \end{array}$$

$$y = X\beta$$

med $X = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix}$ design-matrisen
(man väljer $x_i = \ln(\frac{\sigma_i}{\kappa_c})$
när man designar experimentet)

$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$ parametervektorn
(ska bestämmas)

$y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}$ observationsvektorn
(man mäter/observerar
föjnings hastigheterna)

Inkonsistent system, minsta-
kvadratmetoden:

$$\underbrace{X^T X}_{2 \times 2} \underbrace{\beta}_{2 \times 1} = \underbrace{X^T y}_{2 \times 1}$$

$$\boxed{\begin{array}{l} X \quad N \times 2 \\ \beta = 2 \times 1 \\ y = N \times 1 \end{array}}$$

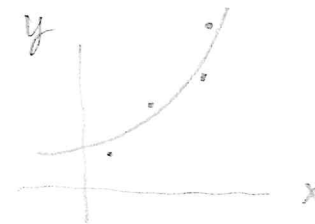
Detta kallas linjär regression,
dvs bestämma parametrar i linjär modell.

Mer allmän modell

$$y = \beta_0 f_0(x) + \beta_1 f_1(x) + \dots + \beta_P f_P(x)$$

kan behandlas med denna metod.

T.ex. $y = \beta_0 + \beta_1 x + \beta_2 x^2$ (kvadratisk
modell)



$$\begin{cases} y_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 \\ \vdots \\ y_N = \beta_0 + \beta_1 x_N + \beta_2 x_N^2 \end{cases}$$

$$\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 \\ \vdots & \vdots & \vdots \\ 1 & x_N & x_N^2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

$$y = X\beta$$

6.7 Lite om allmänna skalär- produktrum.

Def Låt V vara vektorrum.

En skalärprodukt på V är en reell-värd funktion som till varje par u, v tilldelar ett tal $\langle u, v \rangle \in \mathbb{R}$ så att

$$1. \langle u, v \rangle = \langle v, u \rangle$$

$$2. \langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$$

$$3. \langle \lambda u, v \rangle = \lambda \langle u, v \rangle \quad (\lambda \in \mathbb{R})$$

$$4. \langle u, u \rangle \geq 0 \text{ och } \langle u, u \rangle = 0 \Leftrightarrow u = 0$$

Ett vektorrum med skalärprodukt kallas skalärproduktrum.

Ex $\mathbb{R}^n$ med $\langle u, v \rangle = u^T v$.

Ex $V = \mathcal{C}([0, 1]) = \{ \text{kontinuerliga funktioner} \}$

med $\langle f, g \rangle = \int_0^1 f(x)g(x) dx$

Kolla 1, 2, 3 enkelt.

Villkor 4:

$$\langle f, f \rangle = \int_0^1 \underbrace{f(x)^2}_{\geq 0} dx \geq 0 \quad \forall f, \text{ klart,}$$

och $\langle f, f \rangle = 0$ om $f = 0$. Antag nu:

$$0 = \langle f, f \rangle = \int_0^1 f(x)^2 dx$$

$$\Rightarrow f(x)^2 = 0 \quad \forall x, \text{ ty annars } f(x)^2 > 0 \text{ på något litet intervall} \Rightarrow \int_0^1 f(x)^2 dx > 0.$$

$$\Rightarrow f(x) = 0 \quad \forall x$$

$$\Rightarrow f = 0 \text{ (noll-funktionen)}$$

Då kan vi definiera norm:

$$\|u\| = \sqrt{\langle u, u \rangle}$$

Ex $\|f\| = \left(\int_0^1 f(x)^2 dx \right)^{1/2}$.

Allt som vi gjort kan nu göras
i allmänt skalärproduktrum.

* Ortogonalitet: $u \perp v \Leftrightarrow \langle u, v \rangle = 0$

* Ortogonal bas, ortonormal bas.

* Ortogonal projektion:

$$y_W = \text{proj}_W(y) =$$

$$= \frac{\langle y, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1 + \dots + \frac{\langle y, u_p \rangle}{\langle u_p, u_p \rangle} u_p$$

med ortogonal bas.

* Gram-Schmidt.

Viktig olikhet:

Sats 16 (Cauchy-Schwarz olikhet)

$$|\langle u, v \rangle| \leq \|u\| \|v\|$$

(absolutbelopp)

Beweis: Om $u = 0$ så är det trivialt.

Om $u \neq 0$, bildar vi $W = \text{span}\{u\}$.

$\forall v$ får:

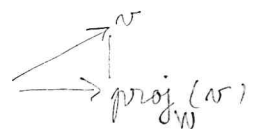
$$\|\text{proj}_W(v)\| = \left\| \frac{\langle v, u \rangle}{\langle u, u \rangle} u \right\| =$$

$$= \left| \frac{\langle v, u \rangle}{\langle u, u \rangle} \right| \|u\| = \frac{|\langle u, v \rangle|}{\|u\|^2} \|u\| =$$

$$= \frac{|\langle u, v \rangle|}{\|u\|}. \text{ Alltså:}$$

$$\frac{|\langle u, v \rangle|}{\|u\|} = \|\text{proj}_W(v)\| \stackrel{\text{Pythagoras}}{\leq} \|v\|$$

$$\text{dvs } |\langle u, v \rangle| \leq \|u\| \|v\|. \quad \square$$

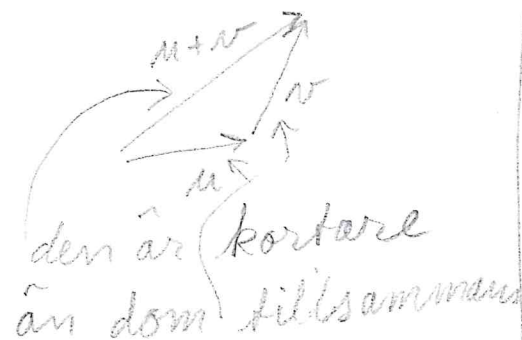


En olikhet till:

Sats 17 (Triangelolikheten)

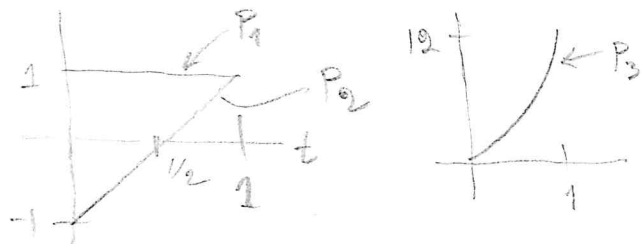
$$\|m+n\| \leq \|m\| + \|n\|$$

Läs beviset.



Ex $V = \mathcal{C}([0,1])$ med $\langle f, g \rangle = \int_0^1 f g dx$.

$P_1(t) = 1$, $P_2(t) = 2t - 1$, $P_3(t) = 12t^2$
är linjärt oberoende.



Ortogonalisera med Gram-Schmidt:

$$1. q_1 = P_1, W_1 = \text{span}\{q_1\} = P_0 = \{\text{poly. grad } 0\}$$

$$2. q_2 = P_2 - \frac{\langle P_2, q_1 \rangle}{\langle q_1, q_1 \rangle} q_1 = P_2 - \text{proj}_{W_1}(P_2)$$

$$\text{där } \langle P_2, q_1 \rangle = \langle P_2, P_1 \rangle = \int_0^1 (2t-1) \cdot 1 dt = 0$$

dvs P_2, P_1 redan ortogonala.

Alltså:

$$q_2 = P_2 \text{ och } W_2 = \text{span}\{q_1, q_2\} = P_1$$

$$3. q_3 = P_3 - \frac{\langle P_3, q_1 \rangle}{\langle q_1, q_1 \rangle} q_1 - \frac{\langle P_3, q_2 \rangle}{\langle q_2, q_2 \rangle} q_2$$

$$= P_3 - \text{proj}_{W_2}(P_3) \text{ där}$$

$$\langle P_3, q_1 \rangle = \int_0^1 12t^2 \cdot 1 dt = 4$$

$$\langle q_1, q_1 \rangle = \int_0^1 1^2 dt = 1$$

$$\langle P_3, q_2 \rangle = \int_0^1 12t^2 \cdot (2t-1) dt = \dots = 2$$

$$\langle q_2, q_2 \rangle = \int_0^1 (2t-1)^2 dt = \dots = \frac{1}{3}$$

$$\text{Alltså: } \text{proj}_{W_2}(P_3) = \frac{4}{1} q_1 + \frac{2}{\sqrt{3}} q_2 =$$

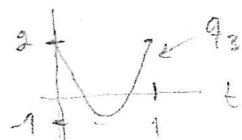
$$= 4q_1 + 6q_2$$

och

$$q_3 = P_3 - 4q_1 - 6q_2, \quad W_3 = \text{span}\{q_1, q_2, q_3\} =$$

$$= P_2$$

$$\text{dvs } \begin{cases} q_3(t) = 12t^2 - 4 \cdot 1 - 6(2t-1) \\ \quad = 12t^2 - 12t + 2 = 12t(t-1) + 2 \\ q_2(t) = 2t-1 \\ q_1(t) = 1 \end{cases}$$



är en ortogonalbas för underrummet

$$P_2 = \{\text{polynom av grad } \leq 2\}.$$

Finite elementmetoden bygger
approximation med polynom.
Kommer i Flervariabel i lp. 4.

Vad är bästa approximationen av

$$f(t) = \begin{cases} 1, & t \in (0, \frac{1}{2}) \\ 0, & t \in (\frac{1}{2}, 1) \end{cases}$$


med polynom av
grad ≤ 1 ? Vi beräknar

$$P = \text{proj}_{P_2}(f) = \text{proj}_{W_2}(f) = \frac{\langle f, q_1 \rangle}{\langle q_1, q_1 \rangle} q_1 + \frac{\langle f, q_2 \rangle}{\langle q_2, q_2 \rangle} q_2$$

$$\text{Här är } \langle f, q_1 \rangle = \int_0^1 f(t) \cdot 1 dt = \int_0^{1/2} 1 dt = \frac{1}{2}$$

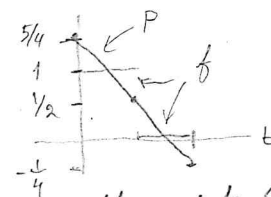
$$\langle f, q_2 \rangle = \int_0^1 f(t) (2t-1) dt = \int_0^{1/2} (2t-1) dt =$$

$$= \left[(t^2 - t) \right]_0^{1/2} = \frac{1}{4} - \frac{1}{2} = -\frac{1}{4}$$

så att

$$P = \frac{\frac{1}{2}}{1} q_1 + \frac{-\frac{1}{4}}{\frac{1}{3}} q_2 = \frac{1}{2} q_1 - \frac{3}{4} q_2$$

$$\text{dvs } p(t) = \frac{1}{2} - \frac{3}{4}(2t-1) = -\frac{3}{2}t + \frac{5}{4}$$



$$\langle q_3, q_3 \rangle = \int_0^1 (12t^2 - 12t + 2)^2 dt =$$

$$= \frac{4}{5}$$

$$\langle f, q_3 \rangle = \int_0^{1/2} (12t^2 - 12t + 2) dt =$$

$$= 0$$

Wolfram Alpha.

$$\text{En term till: } \frac{\langle f, q_3 \rangle}{\langle q_3, q_3 \rangle} q_3 = 0$$

dvs ger inget mer.

%% Example 2 i 6.7

clear all

A=[1 1 0 0;1 1 0 0;1 0 1 0;1 0 1 0;1 0 0 1;1 0 0 1]

b=[-3 -1 0 2 5 1]'

B=rref([A b]) % vi ser att systemet är inkonsistent

%% normalekvationerna

C=rref([A'*A A'*b])

% vi ser att $x(4)=t$ är fri och $x=[3 \ -5 \ -2 \ 0]'$ + $t*[-1 \ 1 \ 1 \ 1]'$

t=0; % välj t

% t=1

x=[3 -5 -2 0]' + t*[-1 1 1 1]' % lösningsmängden till normalekvationerna =
% = mängden av minstakvadratlösningar

r=norm(A*x-b) % residualen

%% backslash varnar men ger en av minstakvadratlösningarna

xx=A\b