Forelesning 6 — ELE 3719 Matematikk

Eivind Eriksen — Feb 10, 2015

Plan:

1. Lineær regresjon i flere variable
2. Sannsynlighetsregning

Pensum:

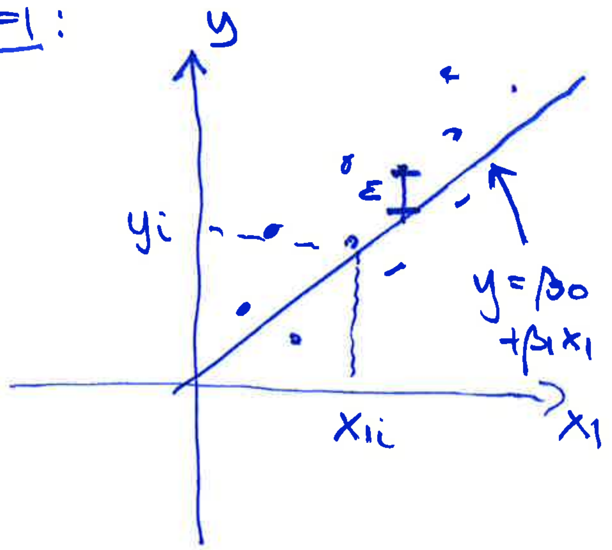
[MKF] 2.4

[S] 2.1–2.9

① Lineær regresjon

	x_1	x_2	x_3	$\dots$	x_n	y
1	x_{11}	x_{21}	x_{31}	$\dots$	x_{n1}	y_1
2	x_{12}	x_{22}	x_{32}	$\dots$	x_{n2}	y_2
3	$\vdots$					
$\vdots$						
N						

n=1:



Mål: Finne en lineær løsning

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n$$

gitt ved $\beta_0, \beta_1, \beta_2, \dots, \beta_n$, som er "best mulig" tilpasning til dataene.

Modell: $y = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n + \varepsilon$

Likningssystem:

$$y_1 = \beta_0 + \beta_1 x_{11} + \dots + \beta_n x_{n1} + \varepsilon_1$$
$$y_2 = \beta_0 + \beta_1 x_{12} + \dots + \beta_n x_{n2} + \varepsilon_2$$
$$\vdots$$
$$y_N = \beta_0 + \beta_1 x_{1N} + \dots + \beta_n x_{nN} + \varepsilon_N$$

Husk: y_i, x_{ij} gitt

β_i ukjent (variable)

$\varepsilon_i = \varepsilon_i(\beta_j)$

Likningssystem:

$$\boxed{\underline{y} = X \cdot \underline{\beta} + \underline{\varepsilon}}$$

$$\underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}}_{\underline{y}} = \underbrace{\begin{pmatrix} 1 & x_{11} & x_{21} & \cdots & x_{n1} \\ 1 & x_{12} & x_{22} & \cdots & x_{n2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1N} & x_{2N} & \cdots & x_{nN} \end{pmatrix}}_{X} \underbrace{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}}_{\underline{\beta}} + \underbrace{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_N \end{pmatrix}}_{\underline{\varepsilon}}$$

Hva betyr det at feilen er minst mulig?

$\|\underline{\varepsilon}\|$ minst mulig $\iff$ $\|\varepsilon\|^2$ minst mulig

$$\|\varepsilon\|^2 = \varepsilon_1^2 + \varepsilon_2^2 + \dots + \varepsilon_N^2$$

metode: <u>minste kvadraters metode</u>

Merk: Når $\underline{v}, \underline{w}$ er vektorer, så har vi

$$\underbrace{\underline{v} \cdot \underline{w}}_{\text{prikke-produkt}} = v_1 w_1 + v_2 \cdot w_2 + \dots + v_n \cdot w_n = \underbrace{\underline{v}^T \underline{w}}_{\text{matrise-multiplikasjon}}$$

$$(v_1\ v_2\ \dots,\ v_n) \cdot \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$$

Problem: Velg $\underline{\beta}$ slik at $\|\varepsilon\|^2$ er minst mulig når $\underline{y} = X \cdot \underline{\beta} + \underline{\varepsilon}$.

$$\|\varepsilon\|^2 = \underbrace{\underline{\varepsilon} \cdot \underline{\varepsilon}}_{\text{prikk produkt}} = \underbrace{\underline{\varepsilon}^T \underline{\varepsilon}}_{\text{matrise produkt}} = (\underline{y} - X\underline{\beta})^T \cdot (\underline{y} - X\underline{\beta})$$

$$= (\underline{y}^T - \underline{\beta}^T X^T) \cdot (\underline{y} - X\beta)$$

$$= \underbrace{\underline{y}^T \cdot \underline{y}}_{\text{konstant-ledd}} - \underbrace{\underline{\beta}^T X^T \underline{y} - \underline{y}^T X \underline{\beta}}_{\text{lineære former}} + \underbrace{\beta^T X^T X \beta}_{\text{kvadr. form i } \underline{\beta}}$$

$$= \underline{\beta}^T (X^T X) \underline{\beta} - 2(\underline{y}^T X) \cdot \underline{\beta} + (\underline{y}^T \cdot \underline{y})$$

annengradsfunksjon

$\beta^T X^T \underline{y} \leftarrow$ 1x1-matrise

$\parallel$

$(\underline{\beta}^T X^T \underline{y})^T = \underline{y}^T X \underline{\beta}$

Konklusjon så langt:

$$\|\underline{\varepsilon}\|^2 = \underline{\beta}^T \cdot (X^T X) \underline{\beta} - 2 y^T X \cdot \underline{\beta} + \underline{y}^T \cdot \underline{y}$$

annengradsfunksjon i $(\beta_0, \beta_1, \ldots, \beta_n)$

Vi vil finne minimum.

Dersom $\quad f(\underline{x}) = \underbrace{\underline{x}^T A \underline{x}}_{\substack{\text{kvadr.}\\ \text{form}\\ (A \text{ symm.})}} + \underbrace{B \cdot \underline{x}}_{\substack{\text{lineær}\\ \text{form}\\ (B \text{ rad})}} + \underbrace{C}_{\text{konstant}}$

Så har vi:

$$\frac{\partial f}{\partial \underline{x}} = \begin{pmatrix} f'_{x_1} \\ f'_{x_2} \\ \vdots \\ f'_{x_n} \end{pmatrix} = 2A \cdot \underline{x} + B^T$$

Dessuten:

i) A positiv semidefinitt $(\underline{x}^T A \underline{x} \geq 0)$ $\Leftrightarrow$ f konveks

ii) A negativ semidefinitt $(\underline{x}^T A \underline{x} \leq 0)$ $\Leftrightarrow$ f konkav

I i) er alle stasjonære pkt globale min

ii) — " — globale max

$rk(A^TA) = rk(A)$:

$A^TA \cdot \underline{x} = \underline{0}$
$n \times n$-system

$A \cdot \underline{x} = \underline{0}$
$m \times n$-system

de to systemene har de samme løsningene, og derfor like mange frihetsgrader

$$n - rk(A^TA) = n - rk(A)$$

Bevis:

Hvis $\underline{x}$ er løsn. av $A \cdot \underline{x} = \underline{0}$

$$\Downarrow$$

$$A^T \cdot A\underline{x} = A^T \cdot \underline{0}$$

$$A^TA \cdot x = \underline{0} \Rightarrow \underline{x} \text{ er løsn. av } A^TA \cdot \underline{x} = \underline{0}$$

Hvis $\underline{x}$ er løsn. av $A^TA \cdot \underline{x} = 0$

$$\Downarrow$$

$$\underline{x}^T \cdot A^TA \cdot \underline{x} = \underline{x}^T \cdot \underline{0}$$

$$\Downarrow$$

$$\|A\underline{x}\|^2 = (A \cdot \underline{x})^T \cdot (A\underline{x}) = \underline{0}$$

$$\Downarrow$$

$$\|A\underline{x}\| = 0$$

$$\Downarrow$$

$$A\underline{x} = \underline{0} \Rightarrow \underline{x} \text{ er løsn. av } A \cdot \underline{x} = \underline{0}$$

$$f(\underline{\beta}) = \|\underline{\varepsilon}\|^2 = \underline{\beta}^T(X^TX)\underline{\beta} - 2\underline{y}^T X \underline{\beta} + \underline{y}^T \cdot \underline{y}$$

$$\frac{\partial f}{\partial \underline{\beta}} = 2(X^TX)\cdot \beta + (-2y^TX)^T$$

$$= 2X^TX \cdot \underline{\beta} - 2X^T\underline{y} = \underline{0}$$

$$\cancel{2}X^TX \cdot \beta = \cancel{2}X^Ty$$

$$(X^TX)\cdot \beta = X^Ty$$

$$\boxed{\underline{\beta} = (X^TX)^{-1} \cdot X^T\underline{y}} \leftarrow \text{hvis } |X^TX| \neq 0$$

Må vise: i) X^TX positiv semidefinitt (slik at det stasjonære punktet er minimum)

ii) $|X^TX| \neq 0$

Om matriser på formen $A^T \cdot A$:

$$\underset{m\times n}{A} \rightsquigarrow \underbrace{\underset{n\times m}{A^T} \cdot \underset{m\times n}{A}}_{n\times n}$$

(kvadratisk)

$$(A^TA)^T = A^T \cdot (A^T)^T = A^T \cdot A$$

Symmetrisk

A^TA er positiv semidefinitt:

$$\underline{x}^T \cdot (A^TA)\cdot \underline{x} = \underline{x}^TA^TA\underline{x} = (A\underline{x})^T \cdot A\underline{x} \qquad y = A \cdot \underline{x}$$

$$= \underline{y}^T \cdot \underline{y} = \|\underline{y}\|^2 \geq 0 \quad \text{for alle } \underline{x}.$$

$$\Downarrow$$

A^TA er positiv semidefinitt

Derfor har vi:

$$|X^T X| \neq 0 \iff rk(X^T X) = n+1$$

$\uparrow$

~~[illegible]~~

$(n+1)\times(n+1)$

$$\Updownarrow$$

$$rk(X) = n+1$$

$\uparrow$

$N\times(n+1)$-matrise

$$\Updownarrow$$

pivot-posisjon i hver kolonne i X

$$\Updownarrow$$

kolonnevektoren i X er lineært uavhengige

$$X = \begin{pmatrix} 1 & x_{11} & x_{21} & \cdots & x_{n1} \\ 1 & x_{12} & x_{22} & & x_{n2} \\ 1 & x_{13} & x_{23} & & x_{n3} \\ \vdots & \vdots & | & & | \\ 1 & & & & \end{pmatrix}$$

__Eks:__ $x_3 = 7 + x_1$ i alle rader er en typisk lineær relasjon

Konklusjon:
Man bør alltid sørge for at kolonnene i X er lineært uavhengige.

Da er $|X^T X| \neq 0$, og

$$\underline{\beta} = (X^T X)^{-1} \cdot (X^T \underline{y})$$

er beste tilpasning

__Eks:__

x	y
1	-1
2	0
3	4

$$y = \begin{pmatrix} -1 \\ 0 \\ 4 \end{pmatrix}$$

$X^T y =$

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix}\begin{pmatrix} -1 \\ 0 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 \\ 11 \end{pmatrix}$$

$$y = -4 + 2{,}5x$$

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

$$|X^T X| = 6 \qquad \beta = (X^T X)^{-1}(X^T y) = \frac{1}{6} \cdot \begin{pmatrix} 14 & -6 \\ -6 & 3 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 11 \end{pmatrix} = \begin{pmatrix} -24/6 \\ 15/6 \end{pmatrix} = \begin{pmatrix} -4 \\ 2.5 \end{pmatrix}$$

Eks: Vi antar at vi har en bank-konto og to aksjer som kan kjøpe ved t=0 og t=1

		B	A-1	A-2	Betinget krav:
t=0		1	5	10	
t=1	w_1	1	6	12	100
	w_2	1	6	8	0
	w_3	1	4	8	0

Kravet er repliserbar: om det fins en handelsstrategi (H_0, H_1, H_2) slik at

$$H_0 + 6H_1 + 12H_2 = 100$$
$$H_0 + 6H_1 + 8H_2 = 0$$
$$H_0 + 4H_1 + 8H_2 = 0$$

$$\begin{pmatrix} 100 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 6 & 12 \\ 1 & 6 & 8 \\ 1 & 4 & 8 \end{pmatrix} \cdot \begin{pmatrix} H_0 \\ H_1 \\ H_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}$$

$$\begin{pmatrix} H_0 \\ H_1 \\ H_2 \end{pmatrix} = \left(\begin{pmatrix} 1 & 6 & 12 \\ 1 & 6 & 8 \\ 1 & 4 & 8 \end{pmatrix}^T \cdot \begin{pmatrix} 1 & 6 & 12 \\ 1 & 6 & 8 \\ 1 & 4 & 8 \end{pmatrix} \right)^{-1} \cdot \begin{pmatrix} 1 & 6 & 12 \\ 1 & 6 & 8 \\ 1 & 4 & 8 \end{pmatrix}^T \cdot \begin{pmatrix} 100 \\ 0 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} -200 \\ 0 \\ 25 \end{pmatrix}$$

$H_0 = -200$ $\varepsilon_1 = 0$
$H_1 = 0$ $\varepsilon_2 = 0$
$H_2 = 25$ $\varepsilon_3 = 0$

Kravet er repliserbart siden $\underline{\varepsilon} = \underline{0}$.

$\underline{\text{Pris} = 50} = 200 - 10 \cdot 25$

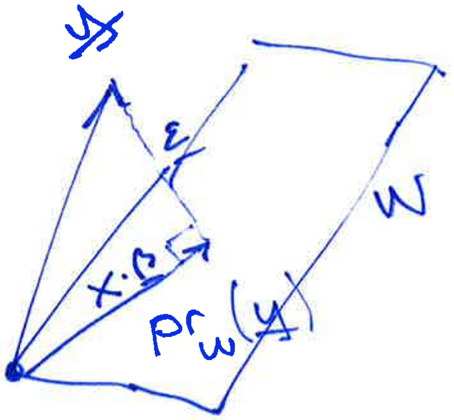
(fair pris for kontrakten)

Geometrisk bilde v/ projeksjon:

$$\underline{y} = X \cdot \underline{\beta} + \underline{\varepsilon}$$

$$X = \begin{pmatrix} 1 & x_{11} & \cdots \\ 1 & x_{12} & \\ \vdots & \vdots & \\ 1 & & \end{pmatrix}$$

$$\underline{1} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \quad \underline{x}_1 = \begin{pmatrix} x_{11} \\ x_{12} \\ \vdots \end{pmatrix} \quad \cdots$$



$$W = \text{span}\{\underline{1}, \underline{x}_1, \dots\}$$

Merk:

$$y = X \cdot \beta$$
$$= \beta_0 \cdot \underline{1} + \beta_1 \cdot \underline{x}_1 + \dots$$
$$\Updownarrow$$
$$y \in W$$

For å løse problemet, kunne vi ha regnet ut

$$\underline{\beta} = \text{proj}_W(\underline{y})$$

(2) Sannsynlighetsregning:

[Stirzaker] = [S]

Kap. 1 : (kan leses)

kap 2-4: Går gjennom utvalgte emner

kap 5-6: Pensum har hovedvekt på disse kapitlene, særlig Kap. 6.

Viktige definisjoner

Et **stokastisk forsøk** er en prosess med klart definerte mulige utfall, men der det faktiske utfallet er uvisst, stokastisk.

Stokastisk = tilfeldig

Utfallsrom: $\Omega = \{\omega_1, \omega_2, \omega_3, \ldots, \omega_k\}$ der $\omega_1, \omega_2, \ldots, \omega_k$ er de k mulige utfallene

Kan ha uendelige utfallsrom:

$$\Omega = \{1, 2, 3, 4, 5, \ldots\}$$

diskret

kontinuerlig

$$\Omega = [1, 3]$$

Hendelse, begivenhet er en samling av mulige utfall (en delmengde av Ω)

Vi har en sannsynlighet $p(E)$ knyttet til hver hendelse slik at

krav → til alle sannsynlighets-fordelinger:

i) $0 \leq p(E) \leq 1$

ii) $p(\Omega) = 1$

iii) $p(E_1 \cup E_2 \cup \dots) = p(E_1) + p(E_2) + \dots$
når $E_i \cap E_j = \emptyset$ (hendelsene er parvis disjunkte)

Eks: Vi kaster to terninger.

E = vi får minst én sekser F = summen er minst 10

$\Omega = \{ (1,1), (1,2), (1,3), (1,4), (1,5), (1,6),$
$(2,1), \dots$
$\vdots$
$(6,1), (6,2), (6,3), (6,4), (6,5), (6,6) \}$

36 utfall

$E = \{ (1,6), (2,6), (3,6), (4,6), (5,6), (6,6)$
$(6,1), (6,2), (6,3), (6,4), (6,5) \}$

11 utfall

$F = \{ (6,4), (5,5), (4,6), (6,5), (5,6), (6,6) \}$

6 utfall

$E \cup F$ = "alle i E og $(5,5)$"

$\cup$: union
alle utfall i E eller F (eller begge)

$E \cap F = \{(4,6),(5,6),(6,6),$
$(6,4),(6,5)\}$

$\cap$: snitt
utfall som er
både i E og
F.

E^c = alle utfall som
ikke er i E.

c: komplement

$p(E) = \frac{11}{36} = \frac{\text{antall gunstige}}{\text{antall mulige}}$ ← gjelder om alle utfall er like sannsynlig

Konsekvenser: av i) - iii)

i) $p(E^c) = 1 - p(E)$

ii) $p(E \cup F) = p(E) + p(F) - p(E \cap F)$

iii) $p(E) = p(\omega_1) + p(\omega_2) + \dots + p(\omega_i)$
når $E = \{\omega_1, \dots, \omega_i\}$

To hendelser E og F er uavhengige
hvis $p(E \cap F) = p(E) \cap p(F)$. Sannsynligheten
for E, gitt at F har inntruffet, kalles
en betinget sannsynlighet $p(E|F)$. Vi har

$$p(E|F) = \frac{p(E \cap F)}{p(F)}$$