Plan

1 Optimering av kvadratiske funksjoner på matriseform
2 Lineær regresjon

① Optimering av kvadratiske funksjoner

Eks: a) $f(x,y) = \underbrace{x^2 - 4xy + 3y^2}_{\substack{\text{grad 2}\\ \text{kvadratisk}\\ \text{form}}} \underbrace{+7x - 8y}_{\substack{\text{grad 1}\\ \text{lineær}\\ \text{form}}} \underbrace{+5}_{\text{grad}}$

Stasjonære pkt.:

$$\begin{cases} f'_x = 2x - 4y + 7 = 0 \\ f'_y = -4x + 6y - 8 = 0 \end{cases}$$

$$\begin{aligned} 2x - 4y &= -7 \\ -4x + 6y &= 8 \end{aligned} \quad \cdot 2$$

$$\begin{aligned} 2x - 4y &= -7 \\ -2y &= -6 \end{aligned}$$

$y = 3 \quad x = 5/2$

Stasjonære pkt: $(5/2, 3)$

Hesse-matrisen

$$H(f) = \begin{pmatrix} f''_{xx} & f''_{xy} \\ f''_{xy} & f''_{yy} \end{pmatrix} = \begin{pmatrix} 2 & -4 \\ -4 & 6 \end{pmatrix}$$

$$2A = 2 \cdot \begin{pmatrix} 1 & -2 \\ -2 & 3 \end{pmatrix} =$$

$$\begin{pmatrix} A & B \\ B & C \end{pmatrix}$$

$A = f''_{xx}(5/2, 3) = 2$

$B = f''_{xy}(5/2, 3) = -4$

$C = f''_{yy}(5/2, 3) = 6$

$\det H(f) = AC - B^2 = 2 \cdot 6 - (-4)^2 = 12 - 16 = -4 < 0$

$\Rightarrow (5/2, 3)$ sadelpkt (hverken maks eller min)

$$f(x,y) = \underbrace{x^2 - 4xy + 3y^2}_{} + \underbrace{7x - 8y}_{} + 5$$

$$= (x\ \ y)\underbrace{\begin{pmatrix} 1 & -2 \\ -2 & 3 \end{pmatrix}}_{A}\begin{pmatrix} x \\ y \end{pmatrix} + \underbrace{(7\ \ -8)}_{B}\cdot\begin{pmatrix} x \\ y \end{pmatrix} + \underbrace{(5)}_{C}$$

$$f(\underline{x}) = \underline{x}^T A \underline{x} + B\underline{x} + C$$

matriseformen til en kvadratisk funksjon

Stasjonære pkt:

$$f'(\underline{x}) = \begin{pmatrix} f'_x \\ f'_y \end{pmatrix} = 2A\cdot\underline{x} + B^T = \underline{0}$$

$$2A\underline{x} = -B^T \quad \Big|\ \tfrac{1}{2}\cdot$$

$$A\underline{x} = -\tfrac{1}{2}B^T$$

$$\begin{pmatrix} 1 & -2 \\ -2 & 3 \end{pmatrix}\cdot\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -7/2 \\ 4 \end{pmatrix}$$

$|A| \neq 0$ en løsn.

$|A| = 0$ ingen løsn. / uendelig mange løsn.

$$f = B\underline{x} = (7\ \ -8)\cdot\begin{pmatrix} x \\ y \end{pmatrix} = 7x - 8y$$

$$f'(\underline{x}) = B^T = \begin{pmatrix} 7 \\ -8 \end{pmatrix}$$

$$f = \underline{x}^T A \underline{x} = (x\ \ y)\begin{pmatrix} 1 & -2 \\ -2 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$= 1\cdot x^2 - 2xy \quad 3\cdot y^2$$

$$f'(\underline{x}) = \begin{pmatrix} \overset{2\cdot 1}{2} & -2\cdot 2 \\ -2\cdot 2 & \underset{2\cdot 3}{6} \end{pmatrix}\cdot\begin{pmatrix} x \\ y \end{pmatrix}$$

$$= 2A\cdot\underline{x}$$

Oppsummering:

$f(\underline{x}) = \underline{x}^T A \underline{x} + B\underline{x} + C$ kvadratisk fn. i n variable

$\underline{x} = (x_1, \dots, x_n)$ som kolonnevektor

A: symm. $n\times n$-matrise

B: radvektor ($1\times n$-matrise)

Da har vi:

i) $f'(\underline{x}) = 2A\cdot\underline{x} + B^T$

$$f'(\underline{x}) = \begin{pmatrix} f'_{x_1} \\ f'_{x_2} \\ \vdots \\ f'_{x_n} \end{pmatrix}$$

ii) Stasjonære pkt: $2A\underline{x} + B^T = 0$

$$A\underline{x} = -\tfrac{1}{2}\cdot B^T$$

$|A| \neq 0$: ett stasj. pkt $\underline{x} = -\frac{1}{2}A^{-1}\cdot B^T$

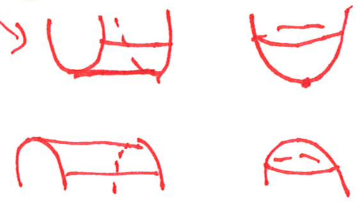
$|A| = 0$: ingen stasj. pkt eller uendelig mange stasj. pkt.

iii) $H(f) = 2A$

$$H(f) = \begin{pmatrix} f''_{x_1x_1} & f''_{x_1x_2} & \cdots \\ \vdots & & \end{pmatrix}$$

symm. $n\times n$-matrise

$\boxed{\begin{array}{lll} A \text{ pos. semidefin} & \Rightarrow & \text{alle stasj. pkt er minimum} \\ \text{f konveks} & & \\ A \text{ neg. semidefin} & \Rightarrow & \text{alle stasj pkt er maksimum} \\ \text{f konkav} & & \\ A \text{ indefinit} & \Rightarrow & \text{alle stasj. pkt sadelpkt} \end{array}}$



Eks: $f(x,y,z) = \underline{2x^2 - 6xy + 3y^2 + 4yz + 3z^2} + \underline{4x - y - z} + \underline{5}$

$f(\underline{x}) = \underline{x}^T A \underline{x} + B\underline{x} + C$

$$A = \begin{array}{c c} \begin{pmatrix} 2 & -3 & 0 \\ -3 & 3 & 2 \\ 0 & 2 & 3 \end{pmatrix} & \begin{array}{c} x \\ y \\ z \end{array} \\ \begin{array}{ccc} x & y & z \end{array} & \end{array}$$

$B = (4 \quad -1 \quad -1) \qquad C = (5)$

Stasj. pkt: $f'(\underline{x}) = 2A\underline{x} + B^T = \underline{0}$

$$A\underline{x} = -\tfrac{1}{2} B^T$$

$$\begin{pmatrix} 2 & -3 & 0 \\ -3 & 3 & 2 \\ 0 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2 \\ 1/2 \\ 1/2 \end{pmatrix}$$

$|A| = 3\cdot(6-9) - 2\cdot 4 = -9 - 8 = -17$

ett stasj. pkt.

$$\left(\begin{array}{ccc|c} 2 & -3 & 0 & -2 \\ -3 & 3 & 2 & 1/2 \\ 0 & 2 & 3 & 1/2 \end{array}\right) \begin{array}{c} \\ \cdot 2 \\ \cdot 2 \end{array} \to \left(\begin{array}{ccc|c} 2 & -3 & 0 & -2 \\ -6 & 6 & 4 & 1 \\ 0 & 4 & 6 & 1 \end{array}\right)]3 \to \left(\begin{array}{ccc|c} 2 & -3 & 0 & -2 \\ 0 & -3 & 4 & -5 \\ 0 & 4 & 6 & 1 \end{array}\right)]4/3 \to \left(\begin{array}{ccc|c} 2 & -3 & 0 & -2 \\ 0 & -3 & 4 & -5 \\ 0 & 0 & 34/3 & -17/3 \end{array}\right)$$

$\frac{34}{3} z = -\frac{17}{3} \Rightarrow z = \underline{-1/2} \quad -3y - 2 = -5 \Rightarrow y = 1 \quad 2x - 3\cdot 1 = -2 \Rightarrow x = 1/2$

Stasj. pkt: $(x,y,z) = (\underline{1/2, 1, -1/2})$

Definithet:

Må finne egenverdiene til A: $\lambda_1, \lambda_2, \lambda_3$

$\lambda_1 \cdot \lambda_2 \cdot \lambda_3 = \det(A) = -17$

$\lambda_1 + \lambda_2 + \lambda_3 = \operatorname{tr}(A) = 8$

A pos semidefin.: $\lambda_1, \lambda_2, \lambda_3 \geq 0$ Nei

A neg. —"— : $\lambda_1, \lambda_2, \lambda_3 \leq 0$ Nei

$\underline{A \text{ indefinit}} \Rightarrow$ det stasjonære pkt er et sadelpkt.

Bruker at en symm. $n\times n$-matrise A m/egenverdier $\lambda_1, \dots, \lambda_n$ har:

$\det(A) = \lambda_1 \cdot \lambda_2 \cdot \ldots \cdot \lambda_n$

$\operatorname{tr}(A) = \lambda_1 + \lambda_2 + \ldots + \lambda_n \qquad \operatorname{tr}(A) =$ summen av tallene på diagonalen

Eks: $n=2$ (to variabler)

$$H(f) = \begin{pmatrix} f''_{xx} & f''_{xy} \\ f''_{xy} & f''_{yy} \end{pmatrix} = \begin{pmatrix} A & B \\ B & C \end{pmatrix}$$

Definithet til $\begin{pmatrix} A & B \\ B & C \end{pmatrix}$:

$AC - B^2 = \det = \lambda_1 \cdot \lambda_2$

$AC - B^2 > 0,\ A > 0$: minimum

$AC - B^2 > 0,\ A < 0$: maksimum

$AC - B^2 < 0$: sadelpkt

$AC - B^2 < 0$: $\lambda_1 \cdot \lambda_2 < 0$

en pos., en neg. egenverdi

$\Rightarrow$ indefinit

$AC - B^2 > 0$: $\lambda_1 \cdot \lambda_2 > 0$

pos. defn.

$\lambda_1, \lambda_2 > 0$

$\lambda_1 + \lambda_2 = A + C > 0$

$A, C > 0$

$\lambda_1, \lambda_2 < 0$

$\lambda_1 + \lambda_2 = \mathrm{tr}(A) = A + C < 0$

$A, C < 0$

neg. defn.

Men: $AC - B^2 > 0$

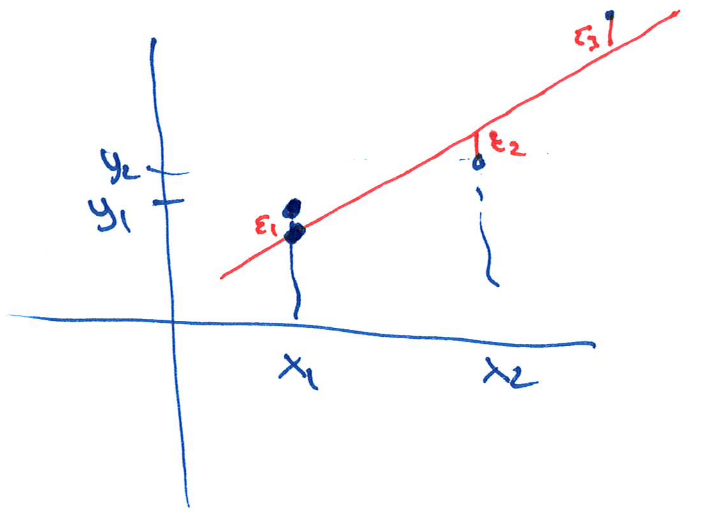
$AC > B^2 \geq 0$

$AC > 0$

$A, C > 0$ eller $A, C < 0$

(2) Lineær regresjon : Modell $y = \beta_0 + \beta_1 x + \varepsilon_1$

x_1	x_2	$\cdots\ x_n$	y
x_{11}	x_{21}		y_1
x_{12}	x_{22}		y_2
			$\vdots$
x_{1N}	x_{2N}		y_N



Idé: Vi velger β_0, β_1 slik at $f(\beta_0,\beta_1) = \varepsilon_1^2 + \varepsilon_2^2 + \dots + \varepsilon_N^2$ er minst mulig

Modell:

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

$$f(\beta_0, \beta_1, \dots, \beta_n) = \left[y_1 - (\beta_0 + \beta_1 x_{11} + \dots + \beta_n x_{1n})\right]^2 + \dots$$

Kvadratisk fn. i $\beta_0, \dots, \beta_n$ som vi skal minimere.

$$\underline{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} \quad X = \begin{pmatrix} 1 & x_{11} & \dots & x_{n1} \\ 1 & x_{12} & \dots & x_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1N} & \dots & x_{nN} \end{pmatrix} \quad \underline{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \quad \underline{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{pmatrix}$$

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

$$\underline{y} - X \cdot \underline{\beta} = \underline{\varepsilon} : \quad \underset{\varepsilon_1^2 + \varepsilon_2^2 + \dots + \varepsilon_N^2}{\underline{\varepsilon}^T \cdot \underline{\varepsilon}} = (\underline{y} - X\underline{\beta})^T \cdot (\underline{y} - X\underline{\beta})$$

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

$$= \underbrace{\underline{y}^T \cdot \underline{y}}_{\text{konst.}} - \underbrace{\underline{\beta}^T X^T \cdot \underline{y} - \underline{y}^T X \underline{\beta}}_{\text{lineære}} + \underbrace{\underline{\beta}^T X^T X \underline{\beta}}_{\text{kvadr.}}$$

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

Problem:

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

Stasjonære pkt:

$$f'(\beta) = 2(X^TX)\beta - 2(y^TX)^T = 0$$

$$2(X^TX)\beta = 2(X^Ty)$$

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

N×N lin. sys

konkret N×N-matrise

konkret N-vektor

Resultat: X^TX er positiv semidefin $\Rightarrow$ evt. stasjonære pkt er min.

Se opg 5.8 i [DA].

I de fleste tilfellene: $|X^TX| \neq 0$

$$\boxed{\beta = (X^TX)^{-1}(X^Ty)}$$

regresjonslinjen

Merk: $|X^TX| = 0 \iff \lambda_1, \lambda_2, \dots, \lambda_n \geq 0$ og minst en egenverdi $\lambda_i = 0$

$$\Updownarrow$$

$\beta^T(X^TX)\beta = 0$ for minst en $\beta \neq 0$

Men $\beta^T(X^TX)\beta = (X\beta)^T\cdot(X\beta) = \|X\beta\|^2$,

så dette betyr at $X\beta \neq 0$ for en $\beta \neq 0$, dvs

$X\beta = 0$ har ikke-trivielle løsn

Derfor: $|X^TX| = 0 \iff \operatorname{rk} X < n+1$

(de $n+1$ kolonnene i X er lineært avhengige)