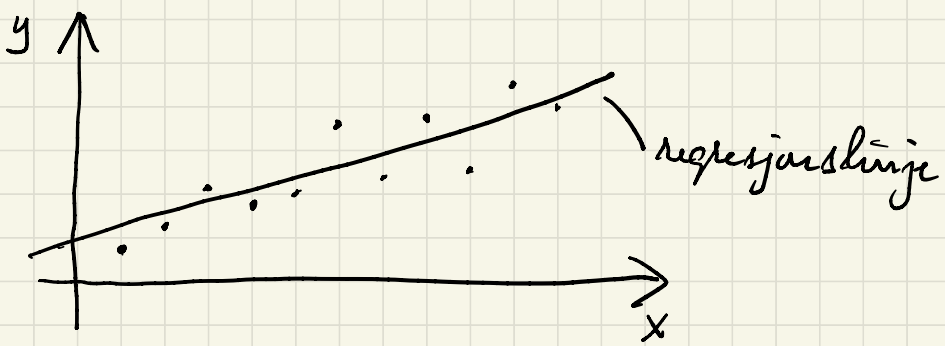
MAT 1080

3.6

3.6 Minste kvadraters metode

En svært mye brukt metode for å analysere data er regresjon, og minste kvadraters metode.

Målet er å finne en lineær sammenheng mellom to størrelser (variable) der data tyder på en slik sammenheng.



Vil finne en rett linje som "best" tilpasser data.

Her kan optimering og partiell derivasjon hjelpe oss!

Made with Goodnotes

Modell: $y = ax + b$, der a og b må bestemmes

Data: $(x_1, y_1), (x_2, y_2), \dots, (x_t, y_t)$
→ gir punktene i planet

Feil: For $x = x_i$ gir data y_i og modell gir $ax_i + b$, så absolutt feil er $|ax_i + b - y_i|$, men det er lettere å jobbe med kvadratisk feil

$$(ax_i + b - y_i)^2$$

Minimere $\sum_{i=1}^{t} (ax_i + b - y_i)^2$

gir de optimal (a, b) som tilpasser modellen best mulig til dataene.

Made with Goodnotes

La $Q(a,b) = \sum_{i=1}^{t} (ax_i + b - y_i)^2$

som er en funksjon i to variable, $Q : \mathbb{R}^2 \to \mathbb{R}$. Finner kritisk punkt ved å sette partielle deriverte lik 0:

$$\frac{\partial Q}{\partial a}(a,b) = 0 \quad , \quad \frac{\partial Q}{\partial b}(a,b) = 0$$

Regner ut:

$$\frac{\partial Q}{\partial a} = \sum_{i=1}^{t} 2(ax_i + b - y_i) \cdot x_i = 0$$

$$\Leftrightarrow \boxed{a\sum_i x_i^2 + b\sum_i x_i - \sum_i x_i y_i = 0}$$

NB: x_i, y_i er gitte konstanter, så dette er en lineær likning i to variable, a og b.

$$\frac{\partial Q}{\partial b} = \sum_i 2(ax_i + b - y_i) \cdot 1 = 0$$

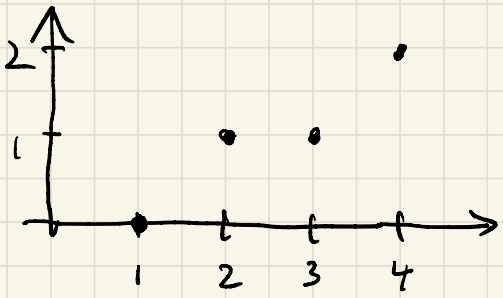
$$\Leftrightarrow \boxed{a\sum_i x_i + tb - \sum_i y_i = 0}$$

-3-

Made with Goodnotes

Så vi har to lineære likninger i to ukjente som vi kan løse og finne optimal a og b.

Eks:



$t=4$. Data:

(1, 0)
(2, 1)
(3, 1)
(4, 2)

Skriver opp likningene, eller kan løse utregningen på nytt:

$$a(1^2+2^2+3^2+4^2)+b(1+2+3+4)-(0+2+3+8)=0$$

$$a(1+2+3+4)+4b-(0+1+1+2)=0$$

$$\Updownarrow$$

$$30a+10b=14$$

$$10a+4b=4$$

Made with Goodnotes

$$30a + 10b = 13$$
$$10a + 4b = 4$$

Löser: II: $10a = 4 - 4b \Rightarrow$ I: $3(4-4b) + 10b = 13$

$\Rightarrow -2b = 1 \Rightarrow b = -\frac{1}{2} \Rightarrow 10a = 4 + 2$

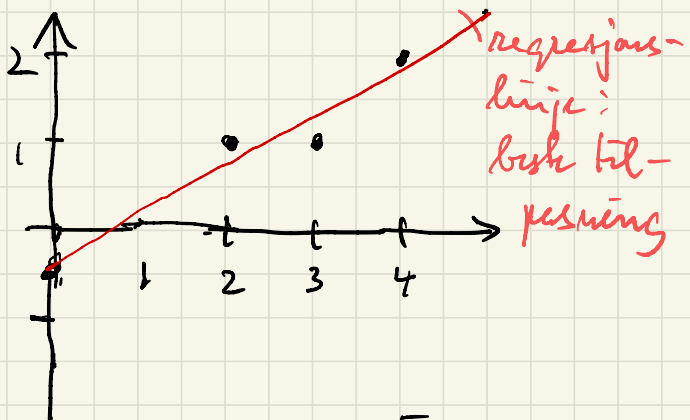
$\Rightarrow a = \frac{6}{10} = \frac{3}{5}$

Så $a = \frac{3}{5}$, $b = -\frac{1}{2}$ är optimal lösning

och därmed vill modellen

$$y = ax + b = \frac{3}{5} \cdot x - \frac{1}{2}$$

hitpassa bäst möjlig.



Made with Goodnotes