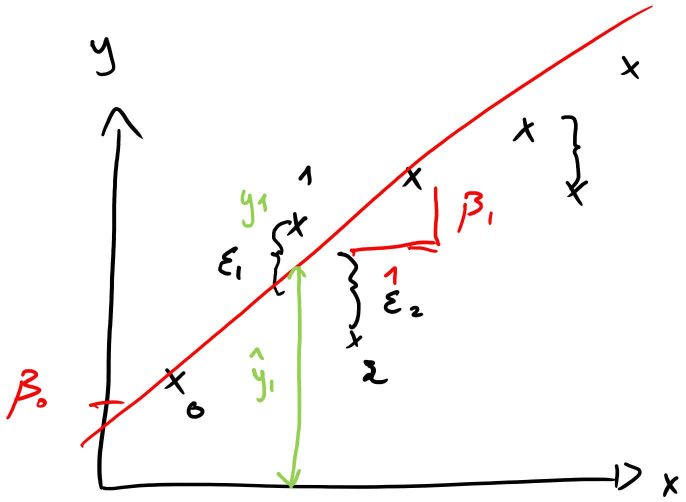
Linear regresjon

Data : $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$



Regresjonsmodell : $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad , \quad i = 0, \ldots, n$

↳ Må estimere β_0 og β_1 $\rightarrow$ $\hat{\beta}_0$ og $\hat{\beta}_1$

Finn linja v. å minimalisere $\{\varepsilon_i\}$ SSE = Sum of sqares of errors

$$L = SSE = \sum_{i=0}^{n} (y_i - \hat{y}_i)^2 \quad , \quad \hat{y}_i = \beta_0 + \beta_1 x_i$$

$$\bar{x} = \sum_{i=0}^{n} \frac{x_i}{n}$$

Minimalisering $\frac{\partial L}{\partial \beta_0} = 0 \quad , \quad \frac{\partial L}{\partial \beta_1} = 0 \longrightarrow$

$$\hat{\beta}_1 = \frac{\overline{xy} - \bar{x}\,\bar{y}}{\overline{x^2} - \bar{x}^2}$$

$$\overline{xy} = \sum_{i=1}^{n} \frac{x_i y_i}{n}$$

$$\overline{x^2} = \sum \frac{x_i^2}{n}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

Multivariat lineær regresjon

Data : $(\underline{x}^0, y^0), (\underline{x}^1, y^1), \ldots, (\underline{x}^{N-1}, y^{N-1})$, N samples

der $\underline{x}^t = [x_1, x_2, x_3, \ldots, x_M]^t$, t er sample-nr.

Regresjonsmodell : $y = w_0 + w_1 x_1 + w_2 x_2 + \ldots + w_M x_M$

Sample-matrise : $\underline{\underline{X}} = \begin{bmatrix} 1 & x_1^0 & x_2^0 & x_3^0 & \cdots & x_M^0 \\ 1 & x_1^1 & x_2^1 & x_3^1 & \cdots & x_M^1 \\ \vdots & & & & & \vdots \\ \vdots & x_1^t & x_2^t & \cdots & & x_M^t \\ \vdots & \vdots & & & & \vdots \\ 1 & x_1^{N-1} & \cdots & & \cdots & x_M^{N-1} \end{bmatrix}$,

Vekter : $\underline{w} = [w_0, w_1, \ldots, w_M]$

y-verdier : $\underline{y} = [y^0, \ldots, y^{N-1}]$

$\Rightarrow \underline{y} = \underline{\underline{X}} \cdot \underline{w}^T \longrightarrow$ Ønsker å estimere $\underline{w}$ (kall gjerne $\hat{\underline{w}}$)

↳ Kan utlede normalligningen: $\underline{\hat{w}} = (\underline{\underline{X}}^T \cdot \underline{\underline{X}})^{-1} \cdot \underline{\underline{X}}^T \cdot \underline{y}$