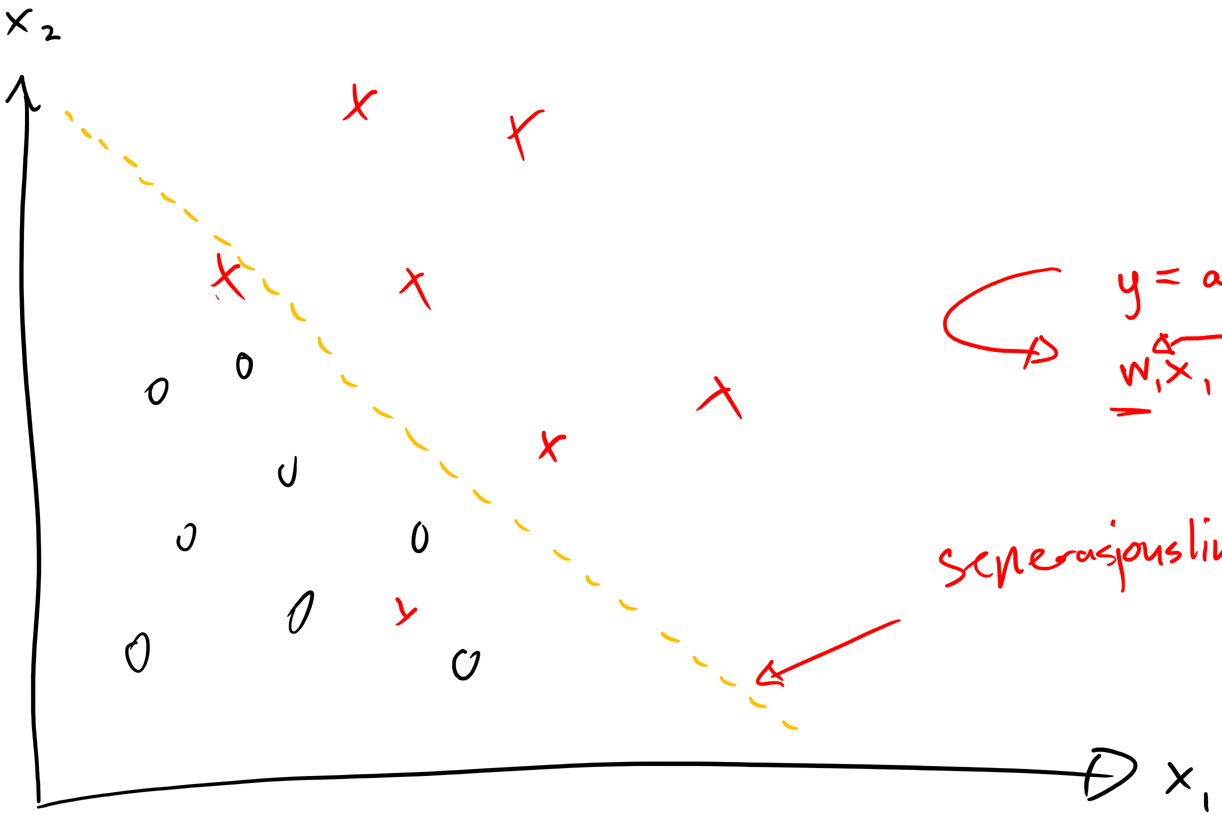


separasjonslinje

$$y = ax + b$$

$$\frac{1}{b}y - \frac{a}{b}x - 1 = 0$$

$$\underline{w_1} x_1 + \underline{w_2} x_2 + \underline{w_0} \cdot 1 = 0$$

$$\underline{W} \cdot \underline{X} = 0 \qquad \underline{X}\,\underline{W}^T$$

$$\underline{W} = [w_2, w_1, w_0]$$

$$\underline{X} = [x_2, x_1, 1]$$

$$\underline{\underline{X}} = \begin{bmatrix} 1 & x_1^{(1)} & x_2^{(1)} & x_3^{(1)} & \cdots & x_n^{(1)} \\ 1 & \vdots & & & & \\ 1 & \vdots & & & & \\ \vdots & \vdots & & & & \\ 1 & x_1^{(m)} & x_2^{(m)} & \cdots & \cdots & x_n^{(m)} \end{bmatrix}$$

← features

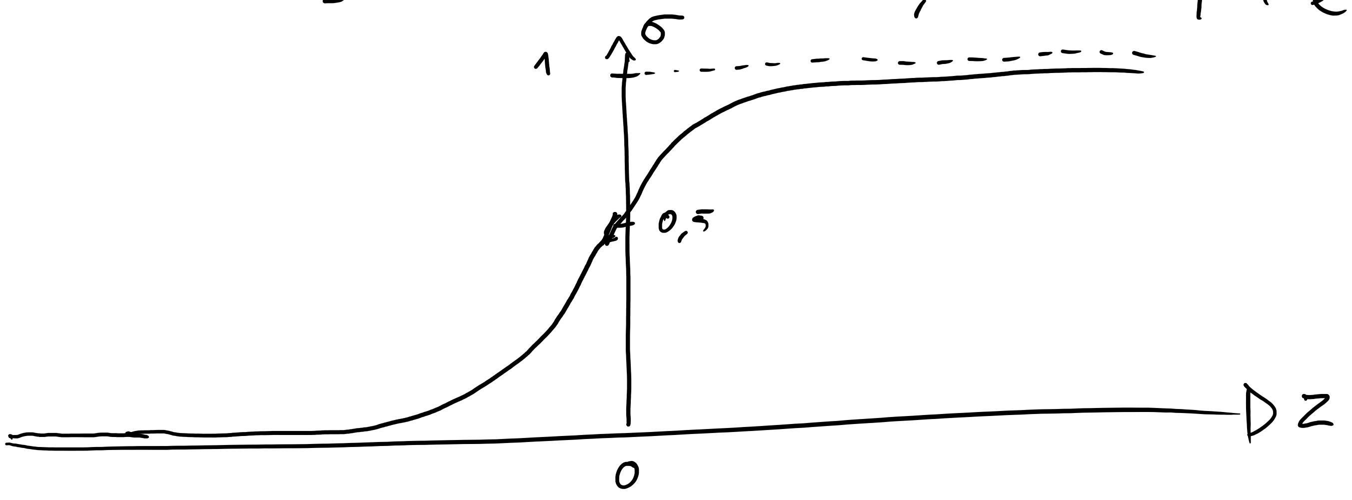
↓ samples

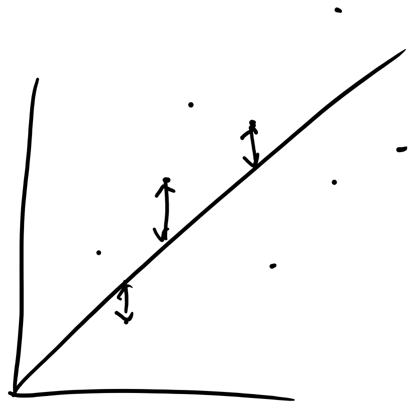
$$\begin{bmatrix} \underline{w_0} \\ \vdots \\ w_n \end{bmatrix}$$

$$\underline{X} \cdot \underline{w}$$

Binær klassifisering: Lære opp en hypotesefunksjon $h_{\underline{w}}(\underline{x})$ som skal være en sannsynlighet $\hat{p} \in [0,1]$ for enten $y=0$ eller $y=1$

$$h_{\underline{w}}(x) = \sigma(\underline{x}\,\underline{w}^T), \quad \sigma(z) = \frac{1}{1+e^{-z}}$$





Kost-funksjon: Et tall som angir feilen mellom modell og "virkelighet"

Lineær regresjon: $J(\beta_0, \beta_1) = \frac{1}{m}\sum_{i=1}^{m}(\hat{y}_i - y_i)^2$

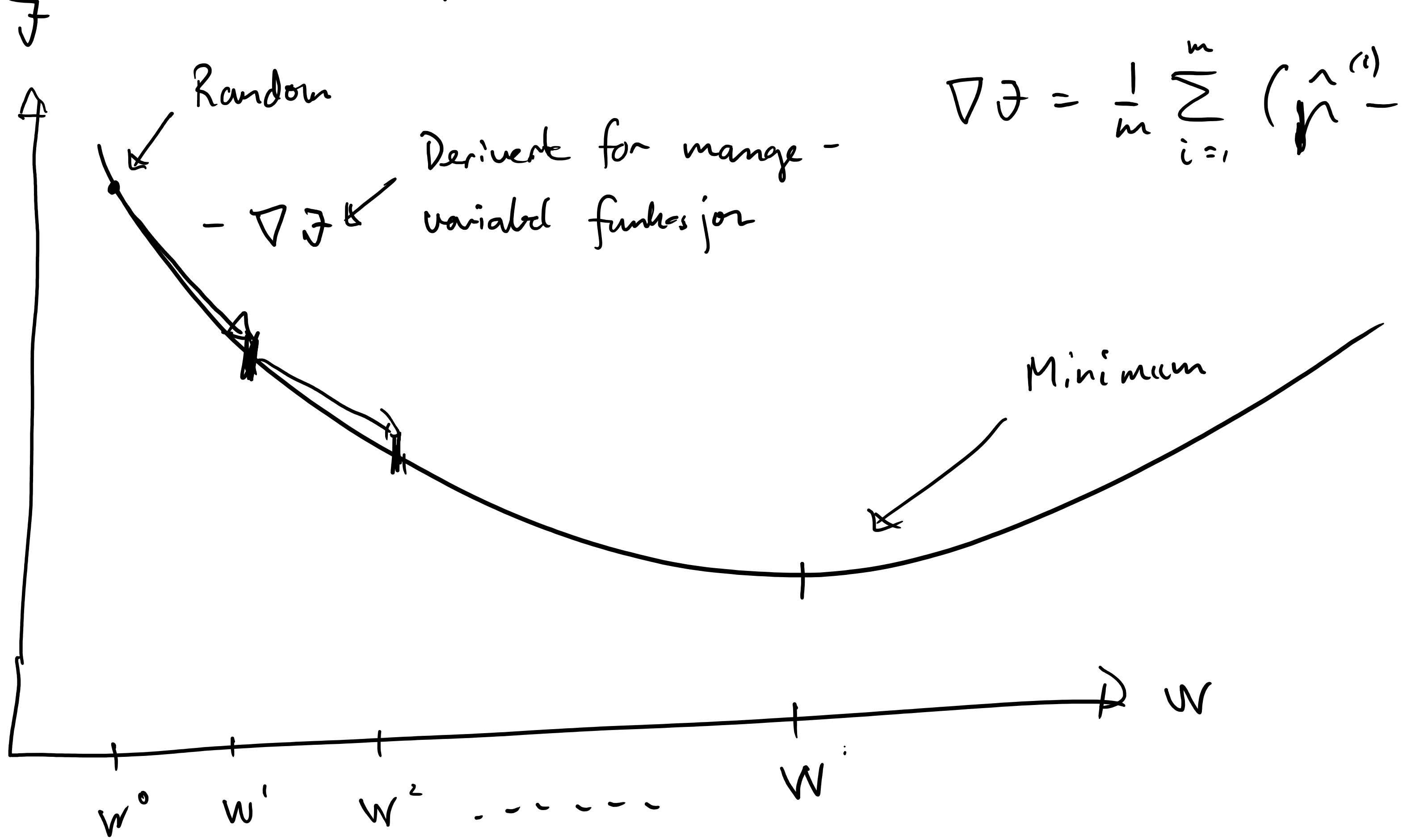
$$y = \beta_1 x + \beta_0$$
$$\hat{y}_i = \beta_1 x_i + \beta_0$$

Logistisk regresjon : Cross-entropy loss

$\hat{p}_i = \sigma(\underline{w}^T \cdot \underline{x})$

$$J(\underline{w}) = -\frac{1}{m} \sum_{i=1}^{m} y^{(i)} \log(\hat{p}_i) + (1 - y^{(i)}) \log(1 - \hat{p}^{(i)})$$

$$\nabla J = \frac{1}{m} \sum_{i=1}^{m} (\hat{p}^{(i)} - y^{(i)}) \cdot \underline{x}^{(i)}$$



epochs = 100 , learning_rate = 0,01 (α, β)

$\underline{w}$ =

for epoch in range(epochs):

tot_error = 0

for i in range(m):

pick random index r

$\underline{x} = \underline{X}[r, :]$ $y = y[r]$

p_hat = sigmoid($\underline{x} \cdot \underline{w}$)

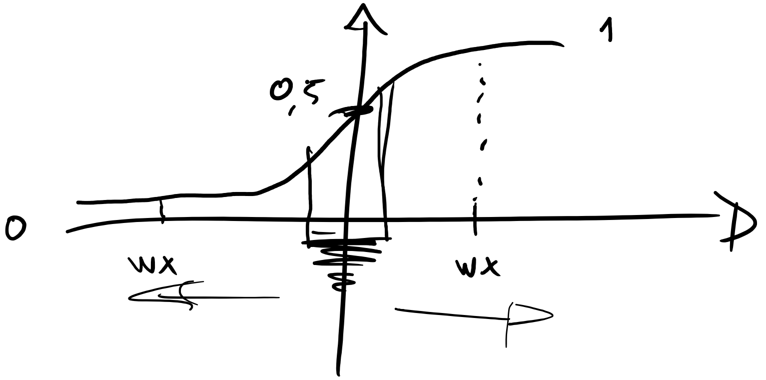
gradient = (p_hat − y) · $\underline{X}$

w = w − learning_rate · gradient

tot_error += (p_hat − y)^2

tot_error = $\frac{1}{m}\sqrt{\text{tot\_error}}$

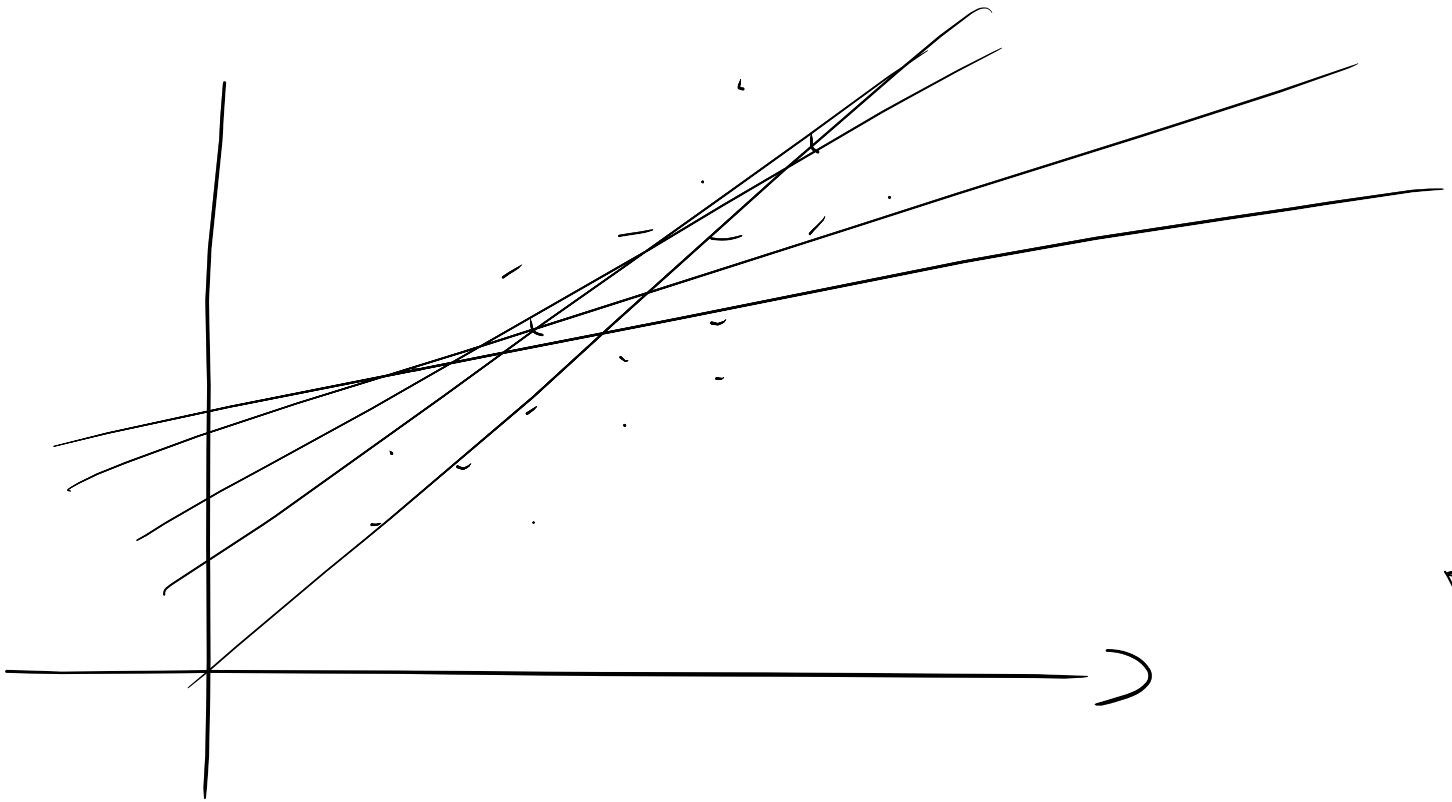
Root mean square error (RMSE)



Kan også skrives som

$w_0 + \underline{w} \cdot \underline{x}$, der $\underline{w} = [w_1, w_2]$

$x = [x_1, x_2]$



$$\hat{y} = \underline{\beta_1} x + \underline{\beta_0}$$

$$J = \frac{1}{m}\sum(\hat{y} - y_i)^2$$

$$\nabla J$$