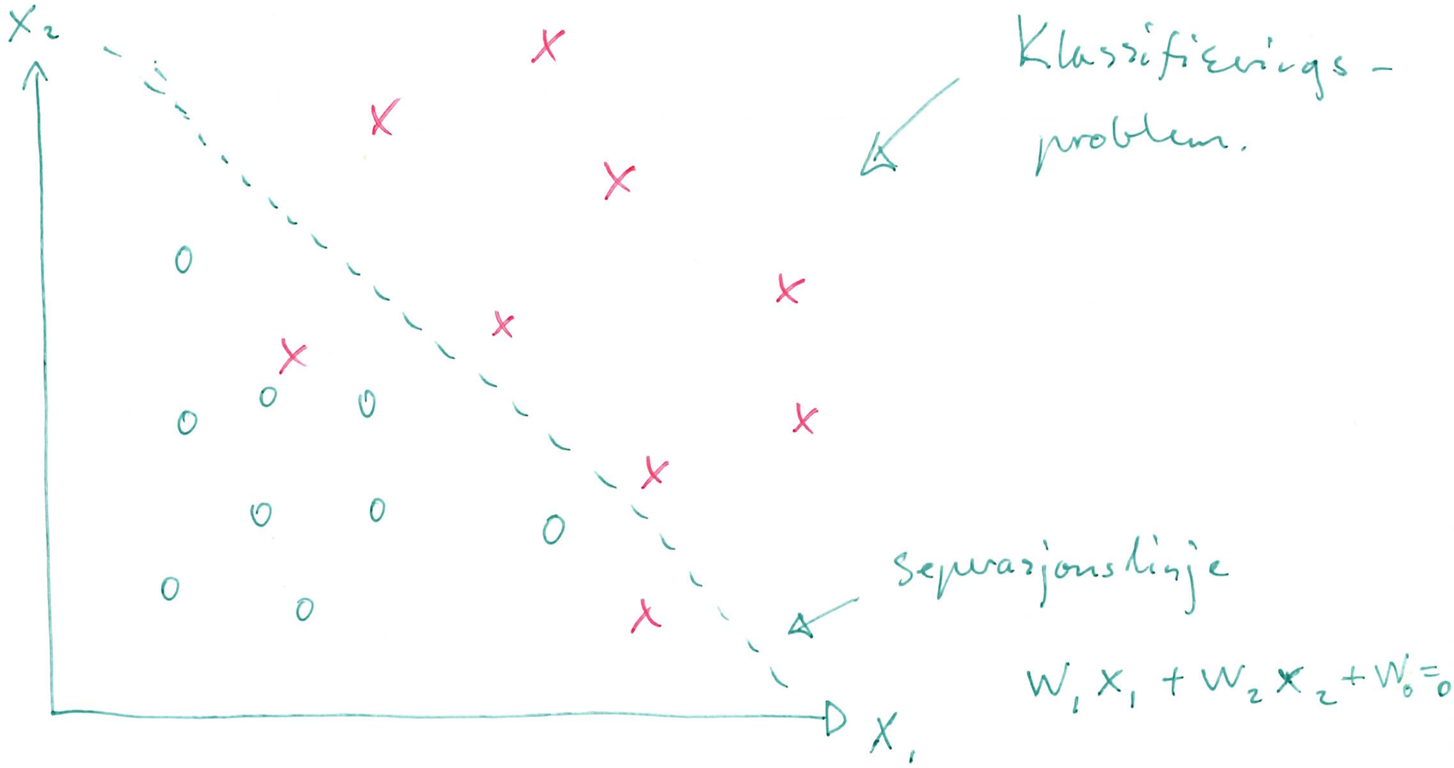
FYS-2021 ; Sett 3 : Logistisk regresjon



$$w_1 x_1 + w_2 x_2 + w_0 = 0$$

Linja kan også skrives som $\underline{x} \cdot \underline{w}^T = 0$, der vektene $\underline{w} = [w_2, w_1, w_0]$ og $\underline{x} = [x_2, x_1, 1]$

↳ Ønsker å finne alg. for å bestemme best mulig $\underline{w}$ gitt et trenings-sett

$$\underline{\underline{X}} = \begin{bmatrix} x_1^{(1)} & x_2^{(1)} & \dots & x_m^{(1)} \\ \vdots & & & \\ x_1^{(N)} & x_2^{(N)} & \dots & x_m^{(N)} \end{bmatrix}$$

← Samples

↓ Features

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

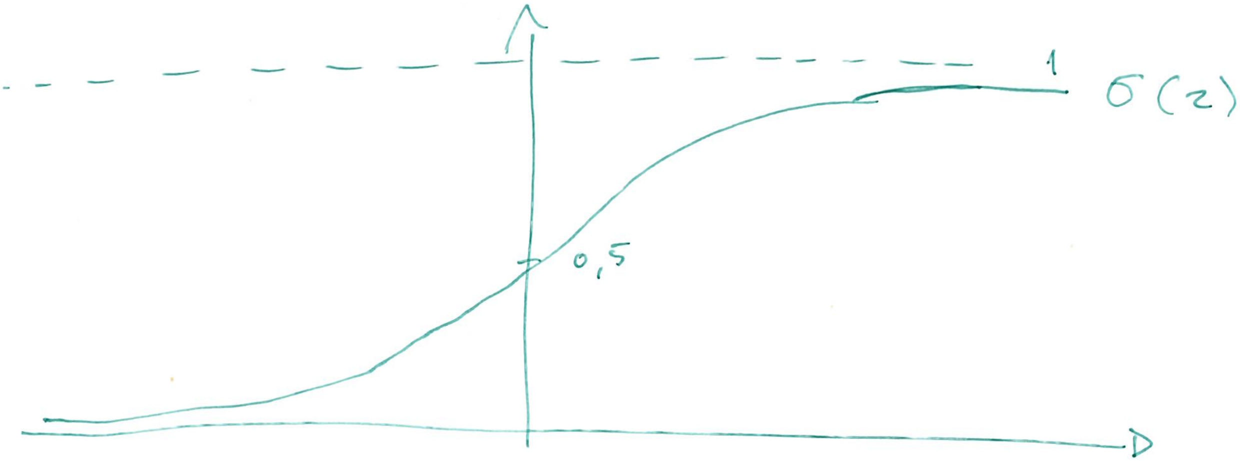
Modell : $h_w(\underline{x}) = \sigma(\underline{w}^T \cdot \underline{x})$

Et tall etter prikk-produkt.

$$\sigma(z) = \frac{1}{1+e^{-z}}$$

↳ Sigmoid / logistiske funksjon

↳ Separerer samplene et sted mellom 0 og 1.



Kostfunksjonen : Et tall som angir "feilen" mellom modell og virkelighet

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

der $\hat{y}^{(i)} = \beta_0 + \underline{x}^{(i)}\beta_1$

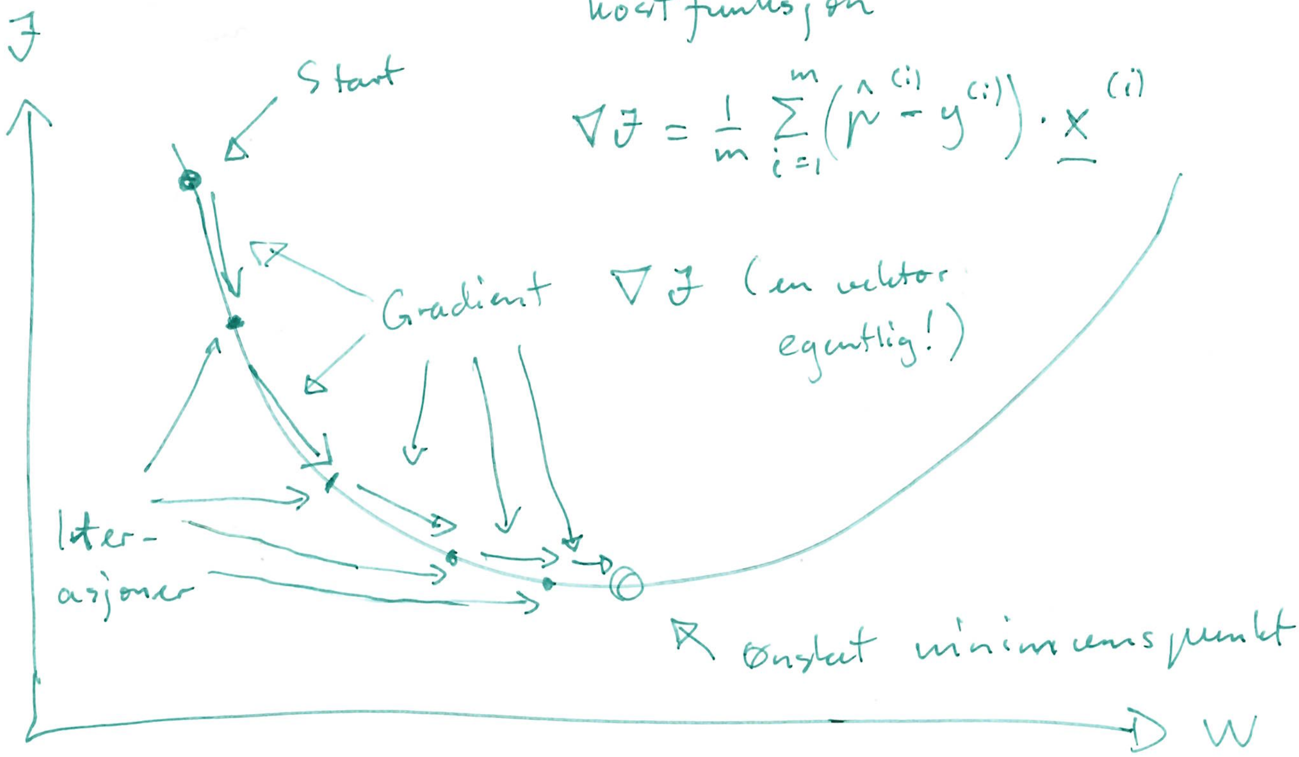
Logistisk regresjon : Bruker $\hat{p}^{(i)} = \sigma(\underline{w}^T \underline{x}^{(i)})$

Kostfunksjon : $J(\underline{w}) = -\frac{1}{m}\sum y^{(i)}\log(\hat{p}^{(i)})$

$+ (1-y^{(i)})\log(1-\hat{p}^{(i)})$

Gradient descent : Bevegelse mot min. av kostfunksjon

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



epochs = 100 (or so) ; learn_rate = 0,01

```
for epoch in range(epochs):
    tot_error = 0
    for i in range(m):                       <- Stocastic grad. descent.
        pick random index r
        x = X[r,:]
        y = y[r]                             <- Enten 0 eller 1
        p_hat = sigmoid(x · w)
        gradient = (p_hat - y) · x
        w = w - learn_rate · gradient
        tot_error += (p_hat - y)^2

    tot_error = tot_error / m                <- Samle i array og plot etterpå
```