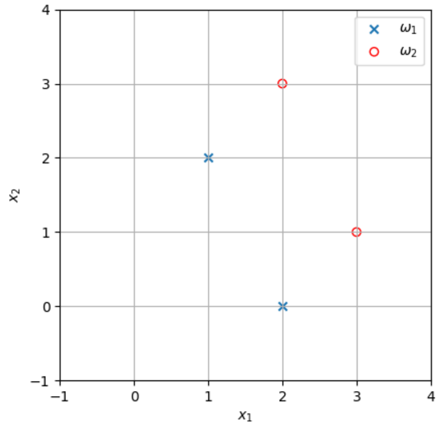
Machine Learning - Theoretical exercise 4

Téo Bouvard

February 10, 2020

Problem 1

a) We have the same number of training samples for classes ω_1 and ω_2, thus the prior probabilities are equal for both classes i.e. $P(\omega_1) = 0.5$ and $P(\omega_2) = 0.5$



b) We compute θ according to the LS-method.

$$\theta = (X^T X)^{-1} X^T y \tag{1}$$

where

$$X = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & 1 \\ 3 & 1 & 1 \\ 2 & 3 & 1 \end{bmatrix}^T \qquad y = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$$

The steps to solve this equation are shown below.

$$X^T X = \begin{bmatrix} 18 & 11 & 8 \\ 11 & 14 & 6 \\ 8 & 6 & 4 \end{bmatrix}$$

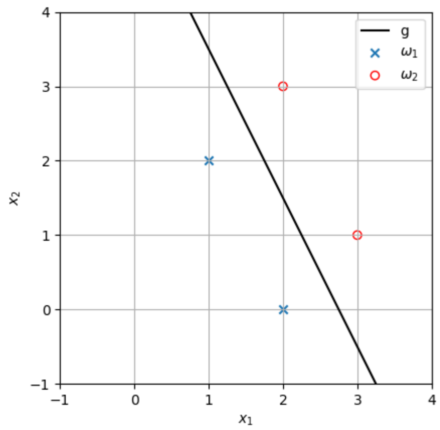
$$(X^T X)^{-1} = \begin{bmatrix} 18 & 11 & 8 \\ 11 & 14 & 6 \\ 8 & 6 & 4 \end{bmatrix}$$

$$(X^T X)^{-1} X^T = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{6} & \frac{1}{2} & \frac{1}{6} \\ 0 & -\frac{1}{3} & 0 & \frac{1}{3} \\ \frac{5}{4} & -\frac{13}{12} & -\frac{3}{4} & -\frac{7}{12} \end{bmatrix}$$

$$\theta = \begin{bmatrix} -\frac{4}{3} \\ -\frac{2}{3} \\ \frac{11}{3} \end{bmatrix}$$

To determine the decision boundary, we find the root of the discriminant function.

$$\begin{aligned} g(x) &= 0 \\ -\frac{4}{3}x_1 - \frac{2}{3}x_2 + \frac{11}{3} &= 0 \\ x_2 &= -2x_1 + \frac{11}{2} \end{aligned}$$

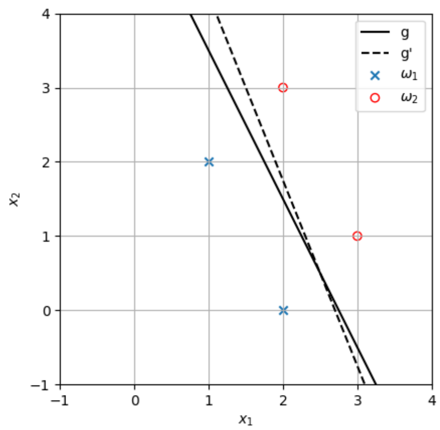


c) If we set $y_4 = -0.5$, we reduce the weight of the fourth training sample, which modifies θ.

$$\theta' = \begin{bmatrix} -\frac{5}{4} \\ -\frac{1}{2} \\ \frac{27}{8} \end{bmatrix}$$

Because $\theta' \neq \theta$, the decision boundary also changes.

$$\begin{aligned} g'(x) &= 0 \\ -\frac{5}{4}x_1 - \frac{1}{2}x_2 + \frac{27}{8} &= 0 \\ x_2 &= -\frac{5}{2}x_1 + \frac{27}{4} \end{aligned}$$



We can see that decreasing the weight of a training sample moves the decison boundary closer to it.

d) We now compute θ with the LMS-method. For the sake of clarity, we will note the threshold value τ, the learning rate μ