University of Stavanger

Stavanger, November 24, 2022

Theoretical exercise 2

ELE520 Machine Learning

Problem 1

Assume that the underlying a priori probabilities and class conditional probability density functions from problem 2, exercise 2 is unknown. However, we have access to measurements so that we have the following samples from the two categories (also illustrated in figure 1).

$$\mathcal{X}_1 = \left\{ \begin{pmatrix} 2 \\ 6 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 3 \\ 8 \end{pmatrix}, \begin{pmatrix} 4 \\ 6 \end{pmatrix} \right\} \tag{1}$$

and

$$\mathcal{X}_2 = \left\{ \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 2.7 \\ -4 \end{pmatrix}, \begin{pmatrix} 3.3 \\ 0 \end{pmatrix}, \begin{pmatrix} 5 \\ -2 \end{pmatrix} \right\} \tag{2}$$

a) Assuming a gaussian distribution, you are supposed to use a parametric approach for formulating the Bayes classifier from problem 2 in exercise 2, based on the two data sets. Apply the maximum-likelihood (ML) method to estimate the required functions. (The expression will look ugly, so do not exhaust yoursel trying to simplify it.)

b) Compare the decision border with the one computed in problem 2 in exercise 2. How are the two estimated density functions oriented in relation to each other and in relation to the true density functions? (Do not perform eigenanalysis, base your answer on observing the nature of the expression for the decision boundary.)

c) How can you make the decision border correspond better to the one found in problem 2, exercise 2.

Trygve Eftestøl, Professor
Faculty of Science and Technology
Department of Electrical and Computer Engineering
Kjølv Egelands hus

University of Stavanger
N-4036 Stavanger.

Telephone: +47 51 83 10 00.
Telefax: +47 51 83 17 50.
E-mail: trygve.eftestol@uis.no
www.ux.his.no/~trygve-e

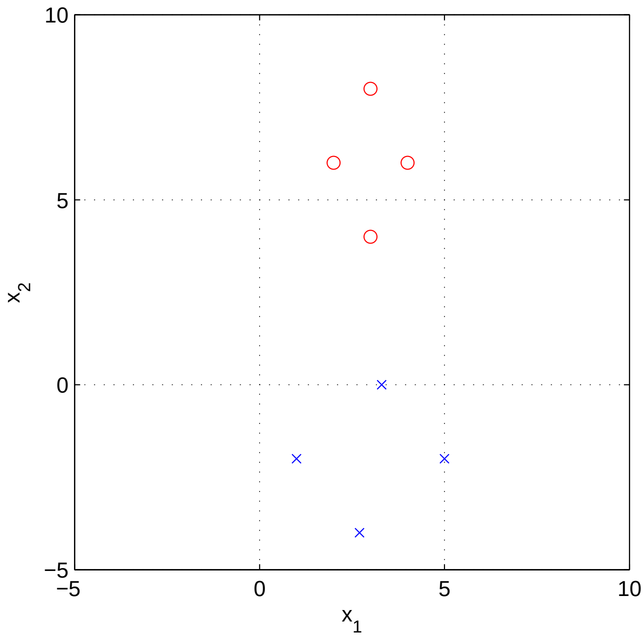

Figure 1: Samples from class ω_1 ('o') and class ω_2 ('x').

Problem 2

Using the data set from the previous problem, you are supposed to classify the feature vector $\boldsymbol{x} = (2.5\ 2.0)^T$. Use the following classifiers:

a) The Bayes classifier from the previous problem.

b) A Parzen-window classifier. Use a gaussian window function so that

$$p_n(\boldsymbol{x}) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{V_N}\phi(\boldsymbol{u}), \tag{1}$$

where $V_N = h_N^l$,

$$\phi(\boldsymbol{u}) = \frac{1}{(2\pi)^{\frac{d}{2}}|\boldsymbol{I}|^{\frac{1}{2}}} \cdot e^{-\frac{1}{2}(\boldsymbol{u})^T\boldsymbol{I}^{-1}(\boldsymbol{u})}, \tag{2}$$

with $\boldsymbol{u} = \frac{\boldsymbol{x}-\boldsymbol{x}_i}{h_N}$, $h_N = h_1/\sqrt{N}$ og $h_1 = 0.5$.

c) A Parzen-window classifier as in the previous subtask, but this time let $h_1 = 5$. Compare with the results from the previous subtask and explain what has happened.

d) A k_N-nearest neighbourhood classifier where $k_N = 1$.

e) A k_N-nearest neighbourhood classifier where $k_N = 3$.

Problem 3

Derive the maximum-likelihood-estimate for

$$\hat{\boldsymbol{\mu}} = \frac{1}{N}\sum_{k=1}^{N} \boldsymbol{x}_k \tag{1}$$

for the case where both $\boldsymbol{\mu}$ og $\boldsymbol{\Sigma}$ in the multivariate probability density function

$$p(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{l}{2}}|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \cdot e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})} \tag{2}$$

are unknown.