Stavanger, January 2, 2020

Theoretical exercise 1

ELE520 Machine learning

A PDF version of the student solution of the exercise shall be submitted on CANVAS.

Problem 1

A similar problem can be found in the example problems set. The main difference lies in that while one-dimensional feature vectors are used in the example, you will have to solve the same kind of problem with two-dimensional feature vectors in this problem.

Given a random vector $\boldsymbol{x} = [x_1, x_2]^T$ characterised by the following two-dimensional probability density function:

$$p(\boldsymbol{x}) = \begin{cases} c & \text{if } a_1 < x_1 < b_1 \text{ og } a_2 < x_2 < b_2 \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

The probability density function is also shown in figure 1.

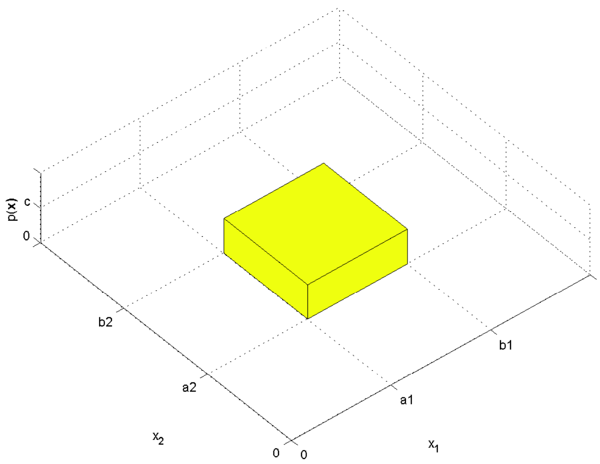

Figure 1: Uniform 2-dimensional probability density function.

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

a) Find c.

b) Find the expected value of $\boldsymbol{x}$ and make an illustration of $p(\boldsymbol{x})$ showing the value.

c) Find the covariance matrix for $\boldsymbol{x}$. The following equation might be useful.

$$E[(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})^T] = E[\boldsymbol{x}\boldsymbol{x}^T] - \boldsymbol{\mu}\boldsymbol{\mu}^T \tag{2}$$

d) Explain the significance of a covariance matrix having identical elements on the diagonal.

e) Explain the significance of a covariance matrix being diagonal.

Problem 2

A similar problem can be found in the example problems set. This problem is given as an opportunity to freshen up your acquaintance with eigen vector analysis. This will later help recognise the significance of diagonalising the covariance matrix which is key to understanding how the data generated from a probability function is oriented in feature space where the principal directions are given by the eigen vectors and their magnitude by the eigen values.

A probability density function , $p(\boldsymbol{x})$, $\boldsymbol{x} = (x_1\ x_2)^t$, has a gaussian distribution around the expected value $\boldsymbol{\mu} = (1\ 1)^t$ with the covariance matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}. \tag{1}$$

a) Find the principal axes of the probability density function through decomposing $\boldsymbol{\Sigma}$ according to

$$\boldsymbol{\Sigma} = \boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{\Phi}^t \tag{2}$$

der $\boldsymbol{\Lambda}$ og $\boldsymbol{\Phi}$ er eigenverdi- og eigenvektormatriser:

$$\begin{aligned} \boldsymbol{\Phi} &= \begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix} \\ \boldsymbol{\Lambda} &= \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \end{aligned} \tag{3}$$

Arrange the eigen vectors according to the eigen values in descending order. Furthermore, the eigen vectors should have unity length ($\mathbf{e}_i^t\mathbf{e}_i = 1$). Principal axis number i might then be found by scaling the eigen vectors according to $\sqrt{\lambda_i}\mathbf{e}_i$.

b) Sketch the contour lines of $p(\boldsymbol{x})$ in the feature space spanned by $\boldsymbol{x}$. Indicate the principal axes and the expected value.