Stavanger, November 19, 2021

Solutions to theoretical exercise 1

ELE520 Machine learning

Problem 1

a)

$$\begin{aligned}
\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(\mathbf{x})d\mathbf{x} &= 1 \\
\int_{a_2}^{b_2}\int_{a_1}^{b_1} c dx_1 dx_2 &= 1 \\
c\int_{a_2}^{b_2} [x_1]_{a_1}^{b_1} dx_2 &= 1 \\
c\int_{a_2}^{b_2} (a_1 - b_1) dx_2 &= 1 \\
c(b_1 - a_1)[x_2]_{a_2}^{b_2} &= 1 \\
c(b_1 - a_1)(b_2 - a_2) &= 1
\end{aligned} \tag{1}$$

i.e.

$$c = \frac{1}{(b_1 - a_1)(b_2 - a_2)}.$$

b)

$$\begin{aligned}
\boldsymbol{\mu} &= \mathbb{E}[\boldsymbol{x}] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \boldsymbol{x} p(\boldsymbol{x}) d\boldsymbol{x} \\
&= \int_{a_2}^{b_2}\int_{a_1}^{b_1} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \frac{1}{(b_1 - a_1)(b_2 - a_2)} dx_1 dx_2 \\
&= \begin{bmatrix} \frac{1}{(b_1-a_1)(b_2-a_2)} [\frac{x_1^2}{2}]_{a_1}^{b_1} [x_2]_{a_2}^{b_2} \\ \frac{1}{(b_1-a_1)(b_2-a_2)} [\frac{x_2^2}{2}]_{a_2}^{b_2} [x_1]_{a_1}^{b_1} \end{bmatrix} \\
&= \begin{bmatrix} \frac{a_1+b_1}{2} \\ \frac{a_2+b_2}{2} \end{bmatrix}.
\end{aligned}$$

This is the center of the rectangle where $p(\boldsymbol{x}) \neq 0$ as illustrated in figure 1.

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

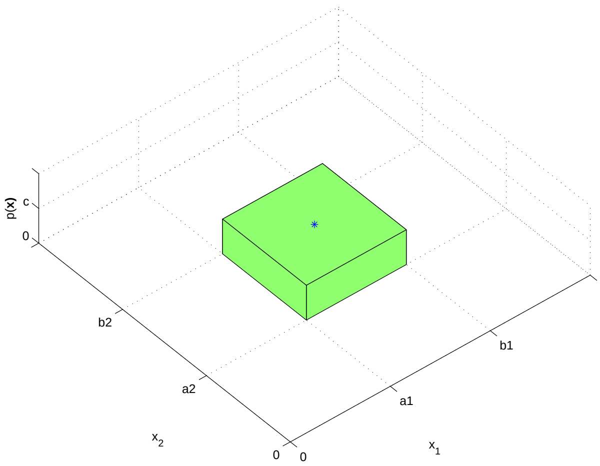

Figure 1: Uniform 2-dimensional probability density function with center of gravity shown raised to level c.

c)

$$\begin{aligned}
\boldsymbol{\Sigma} &= \mathbb{E}[(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})^T] = E\left[\begin{array}{cc} (x_1-\mu_1)^2 & (x_1-\mu_1)(x_2-\mu_2) \\ (x_1-\mu_1)(x_2-\mu_2) & (x_2-\mu_2)^2 \end{array}\right] \\
&= \left[\begin{array}{cc} \mathbb{E}[x_1^2] & \mathbb{E}[x_1x_2] \\ \mathbb{E}[x_1x_2] & \mathbb{E}[x_2^2] \end{array}\right] - \left[\begin{array}{cc} \mu_1^2 & \mu_1\mu_2 \\ \mu_1\mu_2 & \mu_2^2 \end{array}\right], \qquad (2)
\end{aligned}$$

where

$$\begin{aligned}
\mathbb{E}[x_1^2] &= \int_{a_2}^{b_2}\int_{a_1}^{b_1} x_1^2 \frac{1}{(b_1-a_1)(b_2-a_2)} dx_1 dx_2 \\
&= \frac{1}{(b_1-a_1)(b_2-a_2)} \frac{1}{3} [x_1^3]_{a_1}^{b_1} [x_2]_{a_2}^{b_2} \\
&= \frac{(b_2-a_2)(b_1^3-a_1^3)}{3(b_1-a_1)(b_2-a_2)} \\
&= \frac{(b_1-a_1)(b_1^2+a_1^2+a_1b_1)}{3(b_1-a_1)} \\
&= \frac{b_1^2+a_1^2+a_1b_1}{3}, \\
\mathbb{E}[x_2^2] &= \frac{b_2^2+a_2^2+a_2b_2}{3} \qquad (3)
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}[x_1x_2] &= \frac{1}{(b_1-a_1)(b_2-a_2)} \int_{a_2}^{b_2} x_2 dx_2 \int_{a_1}^{b_1} x_1 dx_1 \\
&= \frac{1}{(b_1-a_1)(b_2-a_2)(2)(2)} [x_2^2]_{a_2}^{b_2} [x_1^2]_{a_1}^{b_1} \\
&= \frac{(b_1^2-a_1^2)(b_2^2-a_2^2)}{4(b_1-a_1)(b_2-a_2)} \\
&= \frac{(a_1+b_1)(a_2+b_2)}{4}. \qquad (4)
\end{aligned}$$

Substituting this into (2) we get

$$\boldsymbol{\Sigma} = \begin{bmatrix} \frac{(a_1-b_1)^2}{12} & 0 \\ 0 & \frac{(a_2-b_2)^2}{12}. \end{bmatrix} \tag{5}$$

d) All variables in $\boldsymbol{x}$ have the same variance.

e) The variables in $\boldsymbol{x}$ are uncorrelated.

Problem 2

a) We first find the eigenvalues by solving the characteristic equation given as

$$\begin{aligned} |(\boldsymbol{M} - \lambda \boldsymbol{I})| &= 0 \\ \begin{vmatrix} 5-\lambda & 3 \\ 3 & 5-\lambda \end{vmatrix} &= 0 \\ (5-\lambda)^2 - 9 &= 0 \\ &\Downarrow \\ \lambda_1 = 8 \qquad & \lambda_2 = 2. \end{aligned} \tag{1}$$

Then we find the corresponding eigenvectors by substituting the eigenvalues we found into $\boldsymbol{Mx} = \lambda\boldsymbol{x}$ and solve for $\boldsymbol{x}$.

Substitute $\lambda_1 = 8$:

$$\begin{aligned} \boldsymbol{Mx} &= \lambda_1 \boldsymbol{x} \\ (\boldsymbol{M} - \lambda_1 \boldsymbol{I})\boldsymbol{x} &= \boldsymbol{0} \\ \begin{pmatrix} 5-\lambda_1 & 3 \\ 3 & 5-\lambda_1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= 0 \\ \begin{pmatrix} -3 & 3 \\ 3 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= 0 \\ -3x_1 \quad 3x_2 &= 0 \\ 3x_1 \quad -3x_2 &= 0 \\ &\Downarrow \\ x_2 = x_1 \qquad & x_1 \text{ arbitrary} \\ &\Downarrow \\ x_1 = t \qquad & x_2 = t \end{aligned} \tag{2}$$

.

Correspondingly we solve for $\lambda_2 = 2$:

$$
\begin{aligned}
\boldsymbol{Mx} &= \lambda_2 \boldsymbol{x} \\
(\boldsymbol{M} - \lambda_2 \boldsymbol{I})\boldsymbol{x} &= \mathbf{0} \\
\begin{pmatrix} 5-\lambda_2 & 3 \\ 3 & 5-\lambda_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= 0 \\
\begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= 0 \\
3x_1 \quad 3x_2 &= 0 \\
3x_1 \quad 3x_2 &= 0 \\
&\Downarrow \\
x_2 = -x_1 \qquad & \qquad x_1 \text{ arbitrary} \\
&\Downarrow \\
x_1 = t \qquad & \qquad x_2 = -t
\end{aligned}
\tag{3}
$$

.

Thus we get the eigenvectors $\mathbf{e}_1 = \begin{pmatrix} t & t \end{pmatrix}^t$ and $\mathbf{e}_2 = \begin{pmatrix} t & -t \end{pmatrix}^t$ for λ_1 and λ_2 respectively.

We can substitute arbitrary variables, e.g. $\mathbf{e}_1 = 1/\sqrt{2}\begin{pmatrix} 1 & 1 \end{pmatrix}^t$ og $\mathbf{e}_2 = 1/\sqrt{2}\begin{pmatrix} -1 & 1 \end{pmatrix}^t$ (the scaling ensures unity length).

Thus we get the eigenvector- and eigenvalue-matrices $\mathbf{\Phi}$ and $\mathbf{\Lambda}$:

$$
\begin{aligned}
\mathbf{\Phi} &= \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \\
\mathbf{\Lambda} &= \begin{pmatrix} 8 & 0 \\ 0 & 2 \end{pmatrix}
\end{aligned}
\tag{4}
$$

The principal axes are drawn as a vector between $\boldsymbol{\mu}$ and $\mathbf{p}_i = \boldsymbol{\mu} + \sqrt{\lambda_i}\mathbf{e}_i, i = 1, 2$ which are found to be $(3\ 3)^t$ and $(0\ 2)^t$ respectively.

b) The contour line of the probability density function are shown in figure 2 along with the principal axes.

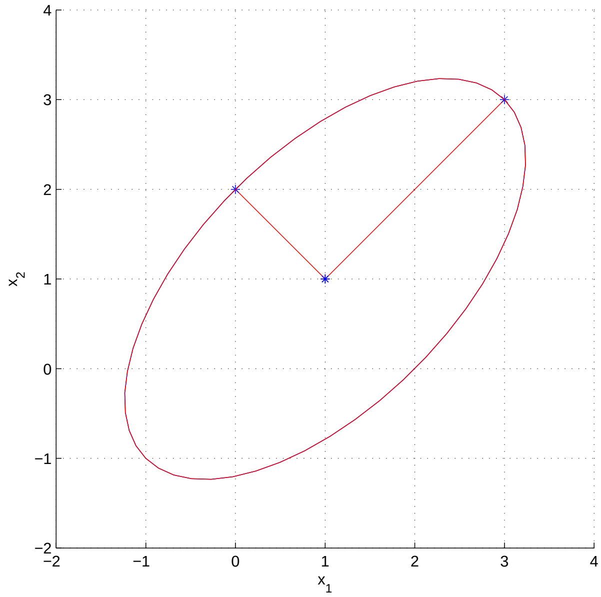

Figure 2: Contour line for 2D distribution.