Statistisk Modellering ST523, ST813
Exercise Session 2

This exercise sheet has two separate sections.

The first section contains exercises that will be addressed during the exercise session. All those exercises should be prepared BEFORE the exercise session.

The second section contains self study exercises which will not be covered. But make sure you know how to solve these.

2 Exercises

Please prepare all the exercises of this section BEFORE the exercise session.

Exercise 2.1

(Poisson distribution, parametrization and identifiability)

Assume that the number of eggs that an insect lays follows a Poisson-distribution with parameter $\lambda > 0$. From each egg, a new insect hatches with probability $p \in (0, 1)$. The hatching of one egg is considered independent from all other eggs. A biologist observes $N \in \mathbb{N}$ insects and notes for each insect the total number of eggs lain as well as the number of hatched eggs. Assume both λ and p are unknown.

a) Find a parametrisation, i.e. determine the distribution of the data depending on the parameters as well as the corresponding parameter space Θ.

b) Now consider the situation, where we only observe how many eggs hatched for each insect. Show that the model above is now unidentifiable.

Exercise 2.2

(Identifiability in linear models)

Consider the following linear model for random observations $Y_1, \ldots, Y_n$:

$$Y_i = \sum_{i=1}^{p} x_{ij}\beta_j + \epsilon_i, \quad i = 1, \ldots, n.$$

We are assuming that (x_{ij}) are known constants for $j = 1, \ldots, p$ and $i = 1, \ldots, n$, and that the residuals $(\varepsilon_i)_{i=1}^n$ are i.i.d. with common distribution $N(0, 1)$.

a) Show that the model (with parameters $(\beta_1, \ldots, \beta_p)$) is identifiable if and only if the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_p$ are linearly independent. Here, $\mathbf{x}_j = (x_{1j}, \ldots, x_{nj})^T$ for $j = 1, \ldots, p$.

b) Argue why the model is not identifiable when $n < p$.

Exercise 2.3

Simple linear regression:

Let $\mathbf{Y} \in \mathbb{R}^n$ be a random (response) vector and consider a linear model for $\mathbf{Y}$ with an intercept and a single predictor x. The entries in $\mathbf{Y}$ can be expressed as

$$Y_i = \beta_1 + \beta_2 x + \varepsilon_i, \text{ for } i = 1, \ldots, n$$

where $\varepsilon_1, \ldots, \varepsilon_n$ are independent error terms with $\mathrm{E}(\varepsilon_i) = 0$ and $\mathrm{Var}(\varepsilon_i) = \sigma^2 > 0$ for $i = 1, \ldots, n$. Equivalently we have

$$\mathbf{Y} = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \varepsilon,$$

with $\mathrm{E}(\varepsilon) = 0$ and $\mathrm{Var}(\varepsilon) = \sigma^2 \mathbf{I}$.

The aim of this exercise is to derive an explicit formula for the least-square estimator $(\hat{\beta}_1, \hat{\beta}_2)^T$ and to discuss its precision.

a) State and solve the normal equations, and show that $(\hat{\beta}_1, \hat{\beta}_2)^T$ is given by

$$\hat{\beta}_1 = \bar{\mathbf{Y}} - \bar{x}\hat{\beta}_2 \quad \text{and} \quad \hat{\beta}_2 = \frac{S_{xy}}{S_x^2}.$$

b) What is the variance matrix of $(\hat{\beta}_1, \hat{\beta}_2)^T$ and what does the precision of the estimated slope $\hat{\beta}_2$ depend on?

Exercise 2.4

Expectation and variance of random vectors

Let $\mathbf{Y} = (Y_1, \ldots, Y_m)^T$ random vector and $\mathbf{A}$ $(n \times m)$-matrix, $\mathbf{b} \in \mathbb{R}^n$, $\mathbf{A}, \mathbf{b}$ non-random. Show that

- $E(\mathbf{AY} + \mathbf{b}) = \mathbf{A}E(\mathbf{Y}) + \mathbf{b}$
- $Var(\mathbf{AY} + \mathbf{b}) = \mathbf{A}Var(\mathbf{Y})\mathbf{A}^T$

Assume that all considered moments exist.

Hint: You can further consider as known that $E(aY + b) = aE(Y) + b$ and $Var(aY + b) = a^2 Var(Y)$ hold for (univariate) random variables Y and $a, b \in \mathbb{R}$.

Exercise 2.5

Orthogonal subspaces, L and $L^\perp$

Let L be a linear subspace of $\mathbb{R}^n$ and $L^\perp$ its orthogonal complement given by

$$L^\perp = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x}^T \mathbf{y} = 0 \text{ for all } \mathbf{y} \in L\}$$

Show that:

- L^T is again a linear subspace.
- $L \cap L^\perp = \{\mathbf{0}\}$
- The decomposition $\mathbf{x} = \mathbf{u} + \mathbf{v}$ with $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in L$, $\mathbf{v} \in L^\perp$ is unique.

Make a plot that shows an example of a linear subspace L of $\mathbb{R}^2$ together with its orthogonal complement $L^\perp$.

3 Self-study at home

These basic exercises will not be covered during the exercise section but make sure you know how to solve these.

Exercise 3.1

(matrix calculus in R)

Enter the below matrices in **R**

$$A = \begin{pmatrix} -1 & 3 \\ 4 & 2 \end{pmatrix}, B = \begin{pmatrix} 4 & -3 \\ 1 & -2 \\ -2 & 0 \end{pmatrix}, \text{ and } C = \begin{pmatrix} 5 \\ -4 \\ 2 \end{pmatrix}$$

and perform the following computations in **R** : The functions `solve(), t(), matrix()` or `as.matrix()` should be useful as well as the matrix multiplication operator `%*%`.

a) $5A$ b) BA c) $A^T B^T$

d) $C^T B$ e) $(B^T B)^{-1} B^T C$

Exercise 3.2

(matrix calculus in R - part II)

Check in **R** that the matrix

$$Q = \begin{pmatrix} 2/3 & -2/3 & 1/3 \\ 1/3 & 2/3 & 2/3 \\ 2/3 & 1/3 & -2/3 \end{pmatrix}$$

is orthogonal, i.e. verify that $Q^{-1} = Q^T$, and compute the determinant of Q.

Exercise 3.3

(Gram-Schmidt)

Find the orthonormalisation of the vectors:

$$x_1 = \begin{pmatrix} 4 \\ 0 \\ 0 \\ 1 \end{pmatrix} \text{ and } x_2 = \begin{pmatrix} 3 \\ 1 \\ 0 \\ -1 \end{pmatrix}.$$

Do the calculations by hand and verify with **R** .

(*Hint:* The `gramSchmidt()` method from the pracma-package can be useful for verification in **R** here.)

Use the orthonormalisation to find the projection of

$$y = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

onto the plane spanned by x_1 and x_2. (*Hint:* find $P_L(y)$ using the matrix $E = [e_1, e_2]$, where e_1 and e_2 are the orthonormal basis vectors of $span\{x_1, x_2\}$)

Exercise 3.4

(spectral decomposition of symmetric matrices)

Calculate the spectral decomposition of the matrix Q by hand

$$Q = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

i.e. represent Q as BDB^T where D is a diagonal matrix containing the eigenvalues of Q on the diagonal and B is an orthogonal matrix containing the corresponding eigenvectors of Q. Check your calculations in **R** .

Hint: In **R** you can use the function `eigen(Q)`.

Exercise 3.5

(Variance and covariance)

Let X and Y be two continuous random variables with probability density functions f_X and f_Y. The *variance* of X and the *covariance* of X and Y are defined as

$$\begin{aligned} \mathrm{Var}(X) &= \sigma_X^2 = \mathrm{E}[(X - \mathrm{E}[X])^2] \qquad \text{and} \\ \mathrm{Cov}(X, Y) &= \mathrm{E}[(X - \mathrm{E}[X])(Y - \mathrm{E}[Y])] \end{aligned}$$

Assume in what follows that all considered moments exist, e.g. $E[X^2] < \infty, |E[XY]| < \infty$ etc.

a) Use linearity of the expectation to show $\mathrm{Var}(X) = \mathrm{E}[X^2] - \mathrm{E}[X]^2$.

b) Show that $E[XY] = E[X]E[Y]$ if X and Y are independent.

   Use the fact that the common probability density of the random vector (X, Y) satisfies $f_{X,Y} = f_X f_Y$ in this case.

c) Prove that $\mathrm{Cov}(X, Y) = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]$, and show that independence of X and Y implies $\mathrm{Cov}(X, Y) = 0$.

d) Do the above results also hold for discrete random variables?