Answer sheet 9

Assignment 1. a) The assumptions imply that A is injective. If $v \in \mathbb{R}^p \setminus \{0\}$ then

$$v^T B v = v^T A^T \Omega A v = (Av)^T \Omega (Av) > 0$$

since $Av \neq 0$ and Ω is positive definite. Thus B is positive definite and in particular invertible. The special case $\Omega = I_n$ shows that $A^T A$ is strictly positive definite.

b) Choose $A = (1, 1)^T$ and

$$\Omega = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

then $A^T \Omega A = 0$.

Remark. If Ω has one positive and one negative eigenvalues, we can always find an injective A such that $A^T \Omega A = 0$.

Assignment 2. We shall use the following fact. If X_1 and X_2 are indepndent, and Y_1 and X_2 are independent, and X_1 and Y_1 have the same distribution, the for any (measurable) function g, $g(X_1, X_2)$ and $g(Y_1, X_2)$ have the same distribution.

(i) Take $c^T = (1, 0)$

(ii) Take $c^T = (0, 1)$ and use (i). And take $c^T = (-1, 0)$, to get $-X \sim X$, so that $\mathbb{E}[-X] = \mathbb{E}[X]$, and then $\mathbb{E}X = 0$.

(iii) Take $c^T = (1, 1)/\sqrt{2}$. Then

$$X \sim (X + Y)/\sqrt{2} \sim (X_1 + X_2)/\sqrt{2}.$$

(iv) We know that this is true for $n = 1, 2$. Suppose that this is true for n and write

$$(X_1 + \cdots + X_{n+1})/\sqrt{n+1} = \sqrt{n/(n+1)}[(X_1 + \cdots + X_n)/\sqrt{n}] + \sqrt{1/(n+1)} X_{n+1}.$$

This has the same distribution as $\sqrt{n/(n+1)}X + \sqrt{1/(n+1)}Y$ by the induction hypothesis. Now choose $c^T = (\sqrt{n}, 1)/\sqrt{n+1}$

(v) Since X has zero mean and finite variance σ^2, by the central limit theorem

$$(X_1 + \cdots + X_n)/\sqrt{n} \xrightarrow{d} N(0, \sigma^2).$$

By (v) this gives $X \sim N(0, \sigma^2)$, and by (ii) $Y \sim N(0, \sigma^2)$.

(vi) Let $U \sim U[0, 1]$ and

$$(X, Y) = (\cos(2\pi U), \sin(2\pi U)),$$

then by symmetry $c^T(X, Y)$ has the same distribution for all $c \in S^1$ but X and Y are not Gaussian. This is the uniform distribution on the unit circle.

Assignment 3.

a) The parameters α_1 and β_1 are only influenced by the rats in the first group, while α_2 and β_2 are only influenced by the second group. Thus

$$\begin{bmatrix} y_{11} \\ y_{21} \\ y_{31} \end{bmatrix} = \begin{bmatrix} 1 & x_{11} \\ 1 & x_{21} \\ 1 & x_{31} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{21} \\ \varepsilon_{31} \end{bmatrix}, \qquad \text{for group 1,}$$

$$\begin{bmatrix} y_{12} \\ y_{22} \\ y_{32} \end{bmatrix} = \begin{bmatrix} 1 & x_{12} \\ 1 & x_{22} \\ 1 & x_{32} \end{bmatrix} \begin{bmatrix} \alpha_2 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_{12} \\ \varepsilon_{22} \\ \varepsilon_{32} \end{bmatrix}, \qquad \text{for group 2.}$$

The model for the two groups toghether is obtained by combining the two previous models into

$$y = \begin{bmatrix} y_{11} \\ y_{21} \\ y_{31} \\ y_{12} \\ y_{22} \\ y_{32} \end{bmatrix}, \quad X = \begin{bmatrix} 1 & 0 & x_{11} & 0 \\ 1 & 0 & x_{21} & 0 \\ 1 & 0 & x_{31} & 0 \\ 0 & 1 & 0 & x_{12} \\ 0 & 1 & 0 & x_{22} \\ 0 & 1 & 0 & x_{32} \end{bmatrix}, \quad \beta = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{21} \\ \varepsilon_{31} \\ \varepsilon_{12} \\ \varepsilon_{22} \\ \varepsilon_{32} \end{bmatrix}.$$

b) The models assume that (i) $\beta_1 = \beta_2$, (ii) $\alpha_1 = \alpha_2$ et (iii) $\alpha_1 = \alpha_2$ et $\beta_1 = \beta_2$. In order to fulfill these assumptions we need to fix some parameters to 0, hence we should re-write the model using $\alpha_2 - \alpha_1$ and $\beta_2 - \beta_1$ as parameters. We can for example write the model for group 2 in terms of the difference w.r.t. group 1 :

$$\begin{bmatrix} y_{12} \\ y_{22} \\ y_{32} \end{bmatrix} = \begin{bmatrix} 1 & x_{12} \\ 1 & x_{22} \\ 1 & x_{32} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} 1 & x_{12} \\ 1 & x_{22} \\ 1 & x_{32} \end{bmatrix} \begin{bmatrix} \alpha_2 - \alpha_1 \\ \beta_2 - \beta_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_{12} \\ \varepsilon_{22} \\ \varepsilon_{32} \end{bmatrix}.$$

With this formulation, the parameters α_1 and β_1 are now common to the two groups, and the new parameters $\alpha_2 - \alpha_1$ and $\beta_2 - \beta_1$ represent the difference between the groups.
The model for the two groups combined is written as :

$$y_{jg} = \mu + \mu_d \delta_{2g} + (\gamma + \gamma_d \delta_{2g}) x_{jg} + \varepsilon_{jg}, \quad j = 1, 2, 3 \quad g = 1, 2,$$

with design matrix and parameters vector

$$X = \begin{bmatrix} 1 & 0 & x_{11} & 0 \\ 1 & 0 & x_{21} & 0 \\ 1 & 0 & x_{31} & 0 \\ 1 & 1 & x_{12} & x_{12} \\ 1 & 1 & x_{22} & x_{22} \\ 1 & 1 & x_{32} & x_{32} \end{bmatrix}, \quad \beta = \begin{bmatrix} \mu \\ \mu_d \\ \gamma \\ \gamma_d \end{bmatrix} \quad \left(= \begin{bmatrix} \alpha_1 \\ \alpha_2 - \alpha_1 \\ \beta_1 \\ \beta_2 - \beta_1 \end{bmatrix} \right).$$

We have used the indicator function δ_{2g} that takes value 1 for the group 2 and 0 otherwise. It is represented by the second column of X above.
The submodels assume (i) $\gamma_d = 0$, (ii) $\mu_d = 0$, (iii) $\mu_d = \gamma_d = 0$ and thus we suppress the following columns X : (i) 4, (ii) 2, (iii) 2 et 4.

Assignment 4.

a) If $Y \in \mathbb{R}^n$ follows a multivariate normal $N_n(X\beta, \sigma^2 I)$, then $Z = Q^{\mathrm{T}}Y$ follows a multivariate Normal distribution with expected value

$$\mathbb{E}(Z) = Q^{\mathrm{T}}X\beta = Q^{\mathrm{T}}QR\beta = R\beta$$

and covariance

$$\operatorname{Var} Z = \sigma^2 Q^{\mathrm{T}}Q = \sigma^2 I.$$

b) By direct computation and using the fact that Q is orthogonal we find

$$\begin{aligned}
u = Q^T\hat{y} &= Q^T X(X^T X)^{-1}X^T y \\
&= Q^T QR((QR)^T QR)^{-1}(QR)^T y \\
&= Q^T QR(R^T Q^T QR)^{-1}R^T Q^T y \\
&= R(R^T R)^{-1}R^T Q^T y \\
&= \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \left(\begin{bmatrix} R_1^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \right)^{-1} \begin{bmatrix} R_1^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} Q_1^{\mathrm{T}} \\ Q_2^{\mathrm{T}} \end{bmatrix} y \\
&= \begin{bmatrix} R_1 \\ 0 \end{bmatrix} (R_1^{\mathrm{T}} R_1)^{-1} R_1^{\mathrm{T}} Q_1^{\mathrm{T}} y \\
&= \begin{bmatrix} R_1 \\ 0 \end{bmatrix} R_1^{-1} R_1^{-\mathrm{T}} R_1^{\mathrm{T}} Q_1^{\mathrm{T}} y \\
&= \begin{bmatrix} R_1 R_1^{-1} Q_1^{\mathrm{T}} y \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} Q_1^{\mathrm{T}} y \\ 0 \end{bmatrix},
\end{aligned}$$

where R_1 is invertible since it's upper triangular with positive diagonal elements.
For v we write :

$$\begin{aligned}
v = Q^{\mathrm{T}}(y - \hat{y}) = Q^{\mathrm{T}}y - Q^{\mathrm{T}}\hat{y} \\
= \begin{bmatrix} Q_1^{\mathrm{T}} \\ Q_2^{\mathrm{T}} \end{bmatrix} y - \begin{bmatrix} Q_1^{\mathrm{T}} y \\ 0 \end{bmatrix} = \begin{bmatrix} Q_1^{\mathrm{T}} y \\ Q_2^{\mathrm{T}} y \end{bmatrix} - \begin{bmatrix} Q_1^{\mathrm{T}} y \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ Q_2^{\mathrm{T}} y \end{bmatrix},
\end{aligned}$$

c) Since $\operatorname{Var} Z = \sigma^2 I$ is diagonal, its diagonal entries are zero. Therefore $Z_1 = Q_1^{\mathrm{T}}Y \in \mathbb{R}^p$ and $Z_2 = Q_2^{\mathrm{T}}Y \in \mathbb{R}^{n-p}$ are independent since they are marginals of Z :

$$z = \begin{bmatrix} Q_1^{\mathrm{T}} y \\ Q_2^{\mathrm{T}} y \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} u_1 \\ v_2 \end{bmatrix},$$

where $u_1 \in \mathbb{R}^p$ and $v_2 \in \mathbb{R}^{n-p}$ are the non zero components of u and v :

$$u = \begin{bmatrix} Q_1^{\mathrm{T}} y \\ 0 \end{bmatrix} = \begin{bmatrix} u_1 \\ 0 \end{bmatrix}, \qquad v = \begin{bmatrix} 0 \\ Q_2^{\mathrm{T}} y \end{bmatrix} = \begin{bmatrix} 0 \\ v_2 \end{bmatrix}.$$

So it follows that U and V are independent.

d) We are going to show that S^2 is a function of z_2 and $\widehat{\beta}$ is a function of z_1. This will conclude the proof.

Since $v = Q^{\mathrm{T}}(y - \hat{y})$, we have $y - \hat{y} = Qv$. Then

$$\begin{aligned} S^2 &= \frac{1}{n-p}(y-\hat{y})^{\mathrm{T}}(y-\hat{y}) = \frac{1}{n-p} v^{\mathrm{T}} Q^{\mathrm{T}} Q v = \frac{1}{n-p} v^{\mathrm{T}} v \\ &= \frac{1}{n-p} \begin{bmatrix} 0 & v_2^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} 0 \\ v_2 \end{bmatrix} = \frac{1}{n-p} v_2^{\mathrm{T}} v_2 = \frac{1}{n-p} z_2^{\mathrm{T}} z_2, \end{aligned}$$

is a function of z_2. A similar computation to the one in part (b) yields that

$$\hat{\beta} = (X^{\mathrm{T}} X)^{-1} X^{\mathrm{T}} y = R_1^{-1} Q_1^{\mathrm{T}} y = R_1^{-1} u_1 = R_1^{-1} z_1,$$

is a function of z_1. The proof is done since

$$z_1 \text{ indep } z_2 \implies u_1 \text{ indep } v_2 \implies \hat{\beta} \text{ indep } S^2.$$

Indeed, measurable functions of independent variables stay independent : here $S^2 = f(v_2)$ and $\widehat{\beta} = g(u_1)$, hence for every $B, C \subseteq \mathbb{R}$ (mesurable),

$$\begin{aligned} &\mathbb{P}(S^2 \in B, \widehat{\beta} \in C) \\ =&\mathbb{P}(v_2 \in f^{-1}(B), u_1 \in g^{-1}(C)) \\ =&\mathbb{P}(v_2 \in f^{-1}(B))\mathbb{P}(u_1 \in g^{-1}(C)) \\ =&\mathbb{P}(S^2 \in B)\mathbb{P}(\widehat{\beta} \in C). \end{aligned}$$

Assignment 5.

$$X_{y\sim a-1} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2 \\ 2 \\ 2 \end{pmatrix}; \quad X_{y\sim a+b} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \\ 1 & 2 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}.$$