Mathematical Statistics, Winter semester 2021/22
Solutions to Problem sheet 2

4) Suppose that

$$Y_i \;=\; \theta \,+\, \varepsilon_i, \qquad i = 1, \ldots, n,$$

holds for some $\theta \in \Theta := \mathbb{R}$, where $\varepsilon_1, \ldots, \varepsilon_n$ are independent random variables such that $E\varepsilon_i = 0$ and $\mathrm{var}(\varepsilon_i) = \sigma_i^2 > 0$ for $i = 1, \ldots, n$.

Compute the best linear estimator of θ.

Solution
We rewrite this regression model in vector/matrix form:

$$\underbrace{\begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}}_{=:Y} = \underbrace{\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}}_{=:X} \theta \,+\, \underbrace{\begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}}_{=:\varepsilon},$$

where $E\varepsilon = 0_n$ and $\mathrm{Cov}(\varepsilon) = \Sigma := \begin{pmatrix} \sigma_1^2 & 0 & \cdots & \cdots & 0 \\ 0 & \sigma_2^2 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \sigma_{n-1}^2 & 0 \\ 0 & \cdots & \cdots & 0 & \sigma_n^2 \end{pmatrix}.$

The inverse matrix of Σ is equal to $\Sigma^{-1} = \begin{pmatrix} 1/\sigma_1^2 & 0 & \cdots & \cdots & 0 \\ 0 & 1/\sigma_2^2 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1/\sigma_{n-1}^2 & 0 \\ 0 & \cdots & \cdots & 0 & 1/\sigma_n^2 \end{pmatrix}.$

According to Problem 1 (1st Problem sheet), the best linear unbiased estimator of θ has the form

$$\widehat{\theta} \;=\; (X^T \Sigma^{-1} X)^{-1} X^T \Sigma^{-1} Y \;=\; \frac{\sum_{i=1}^n \sigma_i^{-2} Y_i}{\sum_{i=1}^n \sigma_i^{-2}}.$$

5) Consider the linear regression model

$$\underbrace{\begin{pmatrix} Y_{11} \\ \vdots \\ Y_{1n_1} \\ \vdots \\ Y_{k1} \\ \vdots \\ Y_{kn_k} \end{pmatrix}}_{=:Y} = \underbrace{\begin{pmatrix} \mathbb{1}_{n_1} & & \\ & \ddots & \\ & & \mathbb{1}_{n_k} \end{pmatrix}}_{=:X} \underbrace{\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}}_{=:\beta} + \underbrace{\begin{pmatrix} \varepsilon_{11} \\ \vdots \\ \varepsilon_{1n_1} \\ \vdots \\ \varepsilon_{k1} \\ \vdots \\ \varepsilon_{kn_k} \end{pmatrix}}_{=:\varepsilon},$$

where $\beta_1, \ldots, \beta_k$ are unknown parameters.
Compute the least squares estimator $\widehat{\beta}$ of β and compute $X\widehat{\beta}$.

Solution
We have that

$$X^T X = \begin{pmatrix} n_1 & & \\ & \ddots & \\ & & n_k \end{pmatrix},$$

and so

$$(X^T X)^{-1} = \begin{pmatrix} 1/n_1 & & \\ & \ddots & \\ & & 1/n_k \end{pmatrix}.$$

Since

$$X^T Y = \begin{pmatrix} \sum_{j=1}^{n_1} Y_{1j} \\ \vdots \\ \sum_{j=1}^{n_k} Y_{kj} \end{pmatrix}$$

we obtain that

$$\widehat{\beta} = \begin{pmatrix} (1/n_1) \sum_{j=1}^{n_1} Y_{1j} \\ \vdots \\ (1/n_k) \sum_{j=1}^{n_k} Y_{kj} \end{pmatrix} =: \begin{pmatrix} \bar{Y}_{1\cdot} \\ \vdots \\ \bar{Y}_{k\cdot} \end{pmatrix}$$

and

$$X\widehat{\beta} = \begin{pmatrix} \bar{Y}_{1\cdot} \, \mathbb{1}_{n_1} \\ \vdots \\ \bar{Y}_{k\cdot} \, \mathbb{1}_{n_k} \end{pmatrix}.$$

6) Consider the model

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \qquad (1 \le i \le k,\ 1 \le j \le m).$$

Rewrite this model in vector/matrix form $Y = \bar{X}\theta + \varepsilon$, where $\theta = (\mu, \alpha_1, \ldots, \alpha_k)^T$ is the unknown parameter.
What is the rank of the matrix $\bar{X}$?
Compute the least squares estimator of θ under the side condition $\sum_{i=1}^k \alpha_i = 0$.
Hint: Consider the linear model from exercise 5). Since

$$\{\bar{X}\theta\colon \theta \in \mathbb{R}^{k+1}\} = \{X\beta\colon \beta \in \mathbb{R}^k\},$$

$\widehat{\theta}$ can be chosen such that $\bar{X}\widehat{\theta} = X\widehat{\beta}$.

Solution
The model can be rewritten as

$$\underbrace{\begin{pmatrix} Y_{11} \\ \vdots \\ Y_{1m} \\ \vdots \\ Y_{k1} \\ \vdots \\ Y_{km} \end{pmatrix}}_{=:Y} = \underbrace{\begin{pmatrix} & \mathbb{1}_m & & \\ \mathbb{1}_{km} & & \ddots & \\ & & & \mathbb{1}_m \end{pmatrix}}_{=:\bar{X}} \underbrace{\begin{pmatrix} \mu \\ \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix}}_{=:\theta} + \underbrace{\begin{pmatrix} \varepsilon_{11} \\ \vdots \\ \varepsilon_{1m} \\ \vdots \\ \varepsilon_{k1} \\ \vdots \\ \varepsilon_{km} \end{pmatrix}}_{=:\varepsilon}.$$

The matrix $\bar{X}$ has rank k since the last k columns of $\bar{X}$ are linearly independent and the first column can be written as the sum of the last k columns.

It follows from Exercise 5 that the least squares estimator $\widehat{\theta} = (\widehat{\mu}, \widehat{\alpha}_1, \ldots, \widehat{\alpha}_k)^T$ of θ has to satisfy

$$\widehat{\mu} + \widehat{\alpha}_1 = \bar{Y}_{1\cdot}$$

$$\widehat{\mu} + \widehat{\alpha}_k = \bar{Y}_{k\cdot}$$

The side conditions yield that

$$\widehat{\alpha}_1 = \bar{Y}_{1\cdot} - \bar{Y}_{\cdot\cdot}$$

$$\widehat{\alpha}_k = \bar{Y}_{k\cdot} - \bar{Y}_{\cdot\cdot}$$

and

$$\mu = \bar{Y}_{\cdot\cdot},$$

where $\bar{Y}_{\cdot\cdot} = (1/k)\sum_{i=1}^k \bar{Y}_{i\cdot} = \frac{1}{km}\sum_{i=1}^k \sum_{j=1}^m Y_{ij}$.