Mathematical Statistics, Winter semester 2021/22

Solutions to Problem sheet 1

1) Suppose that

$$Y = X\theta + \varepsilon$$

holds for some $\theta \in \mathbb{R}^k$, where $E\varepsilon = 0_n$, $\mathrm{Cov}(\varepsilon) = \Sigma$, Σ being a regular matrix. Suppose further that the matrix X has full column rank k.

(i) Show that $X^T\Sigma^{-1}X$ is a regular matrix and that

$$\widehat{\theta} = (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}Y$$

is an unbiased estimator of θ. Compute $E_\theta\left[(\widehat{\theta}-\theta)(\widehat{\theta}-\theta)^T\right]$.

Solution

We show that $a^TX^T\Sigma^{-1}Xa \neq 0$ holds for all $a \in \mathbb{R}^k$ such that $a \neq 0_k$.

Let $a \in \mathbb{R}^k$ be arbitrary, $a \neq 0_k$. Recall that X is a matrix with k columns and n rows. Since $\mathrm{rank}(X) = k$ it follows that X has full column rank, i.e. the columns of X are linearly independent. Therefore $Xa \neq 0_n$. Since Σ^{-1} is positive definite we obtain that

$$a^TX^T\Sigma^{-1}Xa = (Xa)^T\Sigma^{-1}(Xa) > 0.$$

Hence, $X^T\Sigma^{-1}X$ is a regular $(k \times k)$-matrix.

Unbiasedness of $\widehat{\theta}$:

$$\begin{aligned} E_\theta\widehat{\theta} &= E_\theta\left[(X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}(X\theta+\varepsilon)\right] \\ &= (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}E_\theta(X\theta+\varepsilon) = \theta \qquad \forall \beta \in \mathbb{R}^k \end{aligned}$$

Matrix risk of $\widehat{\theta}$:

$$\begin{aligned} E_\theta\left[(\widehat{\theta}-\theta)(\widehat{\theta}-\theta)^T\right] &= \mathrm{Cov}\,(\widehat{\theta}-\theta) \\ &= \mathrm{Cov}\left((X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}Y\right) \\ &= (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}\underbrace{\mathrm{Cov}(Y)}_{=\Sigma}\Sigma^{-1}X(X^T\Sigma^{-1}X)^{-1} \\ &= (X^T\Sigma^{-1}X)^{-1} \end{aligned}$$

(ii) Let $\widetilde{\theta} = LY$ be any arbitrary unbiased estimator of θ.

Show that

$$E_\theta\left[(\widetilde{\theta}-\theta)(\widetilde{\theta}-\theta)^T\right] \;-\; E_\theta\left[(\widehat{\theta}-\theta)(\widehat{\theta}-\theta)^T\right]$$

is non-negative definite.

Hint: A symmetric and positive definite $(n \times n)$-matrix M can be represented as $M = \sum_{i=1}^n \lambda_i e_i e_i^T$, where $\lambda_1, \ldots, \lambda_n$ are the (positive) eigenvalues and $e_1, \ldots, e_n$ are corresponding eigenvectors with $e_i^T e_j = 0$ for $i \neq j$. Then $M^{1/2} := \sum_{i=1}^n \sqrt{\lambda_i} e_i e_i^T$ and $M^{-1/2} := \sum_{i=1}^n (1/\sqrt{\lambda_i}) e_i e_i^T$.

To prove (ii), use the fact that

$$\Big(L\Sigma^{1/2} \,-\, (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1/2}\Big)\Big(L\Sigma^{1/2} \,-\, (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1/2}\Big)^T$$

is non-negative definite.

Solution:

Matrix risk of $\widetilde{\theta}$:

$$E_\theta\left[(\widetilde{\theta}-\theta)(\widetilde{\theta}-\theta)^T\right] \;=\; \mathrm{Cov}\,(LY) \;=\; L\,\mathrm{Cov}(Y)L^T \;=\; L\Sigma L^T$$

$\widetilde{\theta} = LY$ unbiased $\Longrightarrow$

$$E_\theta[L(X\theta+\varepsilon)] \;=\; LX\theta \;=\; \theta$$

has to be fulfilled for all $\theta \in \mathbb{R}^k$ $\quad\Longrightarrow LX = I_k$

Comparison of $\mathrm{Cov}(\widehat{\theta}) = \sigma^2(X^T\Sigma^{-1}X)^{-1}$ and $\mathrm{Cov}(\widetilde{\theta}) = L\Sigma L^T$:
(We use the fact that, for an arbitrary matrix M, MM^T is positive semidefinite.)

$$\begin{aligned}
0_{k\times k} \;&\preceq\; \Big(L\Sigma^{1/2} \,-\, (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1/2}\Big)\Big(L\Sigma^{1/2} \,-\, (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1/2}\Big)^T \\
&=\; L\Sigma L^T \,-\, (X^T\Sigma^{-1}X)^{-1}\underbrace{X^T\Sigma^{-1/2}\;\Sigma^{1/2}L^T}_{=I_k} \\
&\qquad -\, \underbrace{L\Sigma^{1/2}\;\Sigma^{-1/2}X}_{=I_k}(X^T\Sigma^{-1}X)^{-1} \,+\, (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1/2}\;\Sigma^{-1/2}(X^T\Sigma^{-1}X)^{-1} \\
&=\; L\Sigma L^T \quad-\quad (X^T\Sigma^{-1}X)^{-1}
\end{aligned}$$

2) (i) Let

$$X = \begin{pmatrix} 1 & v_1 & v_1^2 & \cdots & v_1^k \\ 1 & v_2 & v_2^2 & \cdots & v_2^k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & v_n & v_n^2 & \cdots & v_n^k \end{pmatrix}.$$

Prove that X^TX is regular if the set $\{v_1, \ldots, v_n\}$ contains at least $k+1$ different values.

Hint: Choose $c = (c_1, \ldots, c_{k+1})^T \neq 0_{k+1} := (0, \ldots, 0)^T$ *and compute* c^TX^TXc.

Solution:

We show that $c^TX^TXc \neq 0$ for all $c \in \mathbb{R}^{k+1}$ such that $c \neq 0_{k+1}$

Let $c \in \mathbb{R}^{k+1}$ be arbitrary, $c \neq 0_{k+1}$. Then

$$Xc = \begin{pmatrix} 1 & v_1 & v_1^2 & \cdots & v_1^k \\ 1 & v_2 & v_2^2 & \cdots & v_2^k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & v_n & v_n^2 & \cdots & v_n^k \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_{k+1} \end{pmatrix} = \begin{pmatrix} \sum_{j=0}^k c_{j+1} v_1^j \\ \vdots \\ \sum_{j=0}^k c_{j+1} v_n^j \end{pmatrix},$$

which yields that

$$c^TX^TXc = \sum_{i=1}^n \Big(\underbrace{\sum_{j=0}^k c_{j+1} v_i^j}_{= f_c(v_i)} \Big)^2$$

f_c is a nonzero polynomial of degree $\leq k$

$\Longrightarrow$ f_c has at most k zeroes

$\Longrightarrow$ If $\#\{v_1, \ldots, v_n\} \geq k+1$, then $f_c(v_1) = \ldots = f_c(v_n) = 0$ is impossible

$\Longrightarrow$ $c^TX^TXc \neq 0$

(ii) Let

$$X = \begin{pmatrix} v_1 & v_1^2 & \cdots & v_1^k \\ v_2 & v_2^2 & \cdots & v_2^k \\ \vdots & \vdots & \ddots & \vdots \\ v_n & v_n^2 & \cdots & v_n^k \end{pmatrix}.$$

Prove that $X^T X$ is regular if the set $\{v_1, \dots, v_n\}$ contains at least k different non-zero values.

Hint: Consider the matrix

$$\widetilde{X} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & v_1 & v_1^2 & \cdots & v_1^k \\ 1 & v_2 & v_2^2 & \cdots & v_2^k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & v_n & v_n^2 & \cdots & v_n^k \end{pmatrix}.$$

Solution:

Suppose that $\{v_1, \dots, v_n\}$ contains at least k different non-zero values.
$\Longrightarrow \#\{0, v_1, \dots, v_n\} \geq k+1$
$\Longrightarrow \widetilde{X}$ has full column rank $k+1$
Delete the first column of $\widetilde{X}$: $\quad \Longrightarrow \begin{pmatrix} 0 & \cdots & 0 \\ & X & \end{pmatrix}$ has rank k
$\Longrightarrow X$ has rank k

3) Consider the linear regression model $Y_i = \theta_1 + x_i\theta_2 + \varepsilon_i$, $i = 1, \ldots, n$, where $\varepsilon_1, \ldots, \varepsilon_n$ are i.i.d. with $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$. Let $\widehat{\theta}$ be the least squares estimator of θ.

(i) Suppose that $x_i \neq x_j$, for some (i, j).
Compute $E[(\widehat{\theta}_i - \theta_i)^2]$, for $i = 1, 2$.
Hint: The inverse of a regular matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *is given by* $\frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

(ii) Suppose that $x_1, \ldots, x_n$ can be chosen by an experimenter, where $x_i \in [-1, 1]$ and $n \geq 2$ is even.
Which choice of $x_1, \ldots, x_n$ minimizes $E[(\widehat{\theta}_i - \theta_i)^2]$? (Take into account that $x_1, \ldots, x_n$ have to be chosen such that $x_i \neq x_j$, for some (i, j); otherwise the least squares estimator is not uniquely defined.)

Solution: We write the regression model in matrix/vector form:

$$Y = X \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \varepsilon,$$

where $X = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$

$$\Longrightarrow X^T X = \begin{pmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix}$$

$$\Longrightarrow (X^T X)^{-1} = \frac{1}{n \sum x_i^2 - (\sum x_i)^2} \begin{pmatrix} \sum x_i^2 & -\sum x_i \\ -\sum x_i & n \end{pmatrix}$$

We obtain

$$E\left[(\widehat{\theta}_1 - \theta_1)^2\right] = \sigma^2 \frac{\sum x_i^2}{n \sum x_i^2 - (\sum x_i)^2} = \sigma^2 \frac{1}{n - (\sum x_i)^2 / (\sum x_i^2)}$$

$$E\left[(\widehat{\theta}_2 - \theta_2)^2\right] = \sigma^2 \frac{n}{n \sum x_i^2 - (\sum x_i)^2}$$

These risks are obviously minimized if and only if $\sum x_i = 0$ and $\sum x_i^2 = n$
$\Longrightarrow \#\{i\colon\ x_i = 1\} = \#\{i\colon\ x_i = -1\} = n/2$