Assignment sheet 9

November 14, 2018

Assignment 1. (a) Let $\Omega_{n\times n}$ be a strictly positive definite matrix, and let $A_{n\times p}$ be of rank $p \leq n$ (linearly indpendent columns). Show that $B_{p\times p} = A^T\Omega A$ is strictly positive definite, hence invertible. Deduce that A^TA is strictly positive definite and invertible.
(b) Show by a counter example that if Ω is only assumed symmetric and invertible, then B is not necessarily invertible. *Hint : the simplest example for $\Omega_{2\times 2}$ will work here.*

Assignment 2. Let $X = (X_1, \dots, X_k)$ be a random vector with finite variance and independent coordinates X_i. We know that

$$X_i \sim N(0, \sigma^2) \text{ for all } i \Longrightarrow c^TX \text{ has the same distribution for any } c \in \mathbb{R}^k \text{ such that } \|c\| = 1$$

We shall see in this assignement that the converse is true for $k = 2$. The proof for any $k \geq 1$ is very similar.
Let $V = (X, Y)^T$ be a random vector in $\mathbb{R}^2$, such that

(i). X and Y are independent ;
(ii). $\mathbb{E}X^2 = \sigma^2 < \infty$,
(iii). c^TV has the same distribution for all $c \in \mathbb{S}^1$.

In the following $X_1, X_2, \dots$ represent independent copies of X.
By a judicious choice of $c \in \mathbb{S}^1$,

(i) show that $c^TV \sim X$, for all $c \in \mathbb{S}^1$.
(ii) show that X and Y have the same distribution and find the expectation of X.
(iii) Show that $X \sim \frac{1}{\sqrt{2}}(X_1 + X_2)$.
(iv) By induction show that the distribution of $n^{-1/2}\sum_{i\leq n} X_i$ is the same for all n.
(v) Use the central limit theorem to conclude that $X, Y \sim N(0, \sigma^2)$.
(vi) Show that the above result does not hold if if X and Y are not assumed to be independent. *Hint : For $U \sim$ Unif $(0, 2\pi)$ the distribution of e^{iU} is same as that of $e^{i(U+\phi)}$ for any $\phi \in \mathbb{R}$*

Assignment 3. We are given the weights of two groups of rats at the beginning and at the end of a 15-days long experiment. During these 15 days, a group is fed normally, the other with some growth inibitors. Let x the weight at the beginning of the experiment and y the weight at the end. Assume the following linear model for the weights :

$$y_{jg} = \alpha_g + \beta_g x_{jg} + \varepsilon_{jg}; \qquad j = 1, 2, 3; \quad g = 1, 2.$$

a) Write the design matrix and the vectors of parameters.
b) We are interested in the following models : (i) $\beta_1 = \beta_2$ (if plotting, we will get two parallel lines), (ii) $\alpha_1 = \alpha_2$ (if plotting, we would get the same intercept at the origin) et (iii) $\alpha_1 = \alpha_2$ et $\beta_1 = \beta_2$ (same line for the two groups).
Find a model that admits (i), (ii), (iii) as sub-models (a sub-model is obtained by a full model by fixing some parameters, in this case to zero). Give the design matrix of this model as well as the parameter vectors. Indicate as well which columns of the design matrix one needs to suppress in order to retain each one of the sub-models.
Hint : To write the model, some parameters needs to be set equal to 0. Write the model by using $\alpha_2 - \alpha_1$ and $\beta_2 - \beta_1$ as parameters.

Assignment 4. Let $X = QR = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix} \begin{pmatrix} R_1 \\ 0 \end{pmatrix}$ the QR decomposition for a full-rank matrix $X_{n\times p}$ where $n \geq p$. Assume that y is a vector of n observations sampled from $Y \sim N_n(X\beta, \sigma^2 I)$. Denote the orthogonal projection of y onto the range of X by $\hat{y}$.

a) Write $Z = Q^{\mathrm{T}} Y$. Show that $Z \sim N_n(R\beta, \sigma^2 I)$.

b) Let $u = Q^{\mathrm{T}} \hat{y}$ and $v = Q^{\mathrm{T}}(y - \hat{y})$. Show that $u = \begin{pmatrix} Q_1^{\mathrm{T}} y \\ 0 \end{pmatrix}$ and $v = \begin{pmatrix} 0 \\ Q_2^{\mathrm{T}} y \end{pmatrix}$. *Hint :* $\hat{y} = X(X^T X)^{-1} X^T y$

c) Show that $U = Q^T \hat{Y}$ and $V = Q^T(Y - \hat{Y})$ are independent. *Hint : How is Z distributed ?*

d) Show that $\hat{\beta}$ and S^2 are independent. Recall that $\hat{\beta} = (X^T X)^{-1} X^T Y$ and $S^2 = \frac{1}{n-p}(Y - X\hat{\beta})^T (Y - X\hat{\beta})$

QR **decomposition :** All real matrices $X \in \mathbb{R}^{n\times p}$, with $n \geq p$, admit a decomposition

$$X_{n\times p} = Q_{n\times n} R_{n\times p},$$

where Q is an orthogonal matrix and R is an upper triangular matrix ($R_{ij} = 0$ si $i > j$). Sometimes it is useful to write

$$X = QR = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix} \begin{pmatrix} R_1 \\ 0 \end{pmatrix} = Q_1 R_1,$$

where Q_1 is $n \times p$, and R_1 is $p \times p$. One can show that for each matrix X of full rank it is possible to chose $R_{1,ii} > 0$, $i = 1, \ldots, p$, and that the decomposition $X = Q_1 R_1$ is unique.

Assignment 5. In `R` we can write a linear model through the following command

`reponse~expression,`

where `reponse` might sometimes be absent and `expression` is a collection of terms joined by operators, all normally assembled into an arithmetic expression. For example, suppose that

$$y = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \\ 1 \\ 5 \end{pmatrix}, \quad X = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 2 & 1 & 3 \\ 0 & 2 & 1 \\ 1 & 2 & 2 \\ 0 & 2 & 3 \end{pmatrix},$$

and let `x`, `a`, `b` the columns of $X = [x, a, b]$.
Here some `R` commands and the model that they represent :

Command `R`	Model
`y~x`	$y_j = \beta_0 + \beta_1 x_j + \varepsilon_j$
`y~x-1`	$y_j = \beta_1 x_j + \varepsilon_j$
`y~x+a`	$y_j = \beta_0 + \beta_1 x_j + \beta_2 a_j + \varepsilon_j$.

Write down the design matrix corresponding to the formula : (i) `y~a-1`, (ii) `y~a+b`.