# Estimation theory – Report 2

Marta Frankowska, 208581
Agnieszka Szkutek, 208619

November 16, 2017

## Contents



# 1 Exercise 1

We have data file with 100 rows and 4 columns. We take the first column as a column vector $y$ and the remaining 3 columns as a matrix $X$, where $y$ depends on $X$. We assume that the model for our data is as follows

$$y = X\alpha + u. \tag{1}$$

We will use regression model function in R to compute the parameters and compare them with the results we obtain manually.

```
data <- read.table('data_lab_2.csv', sep = ",", dec = ",", header = FALSE)
attach(data)

N <- 100
K <- 3
```

```
X <- as.matrix(data[, -1])
y <- as.matrix(data[, 1])
# linear regression model using lm()
model <- lm(V1 ~ . - 1, data)

summary(model)

##
## Call:
## lm(formula = V1 ~ . - 1, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.68919 -0.47894  0.08483  0.47957  2.06775
##
## Coefficients:
##    Estimate Std. Error t value Pr(>|t|)
## V2  2.05860    0.06862   30.00   <2e-16 ***
## V3  1.07476    0.06720   15.99   <2e-16 ***
## V4  0.88974    0.07506   11.85   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8956 on 97 degrees of freedom
## Multiple R-squared:  0.9951,Adjusted R-squared:  0.9949
## F-statistic:  6517 on 3 and 97 DF,  p-value: < 2.2e-16
```

## 1.1 Part 1

We want to express the loss function

$$L = \sum_{n=1}^{N} u_n^2$$

as a function of $y$, $X$ and $\alpha$. From (1) we have $u = y - X\alpha$. Then

$$\begin{aligned} L = \sum_{n=1}^{N} u_n^2 = u'u = (y - X\alpha)'(y - X\alpha) = (y' - \alpha'X')(y - X\alpha) = \\ = y'y - y'X\alpha - \alpha'X'y + \alpha'X'X\alpha = y'y - 2y'X\alpha + \alpha'X'X\alpha. \end{aligned}$$

## 1.2 Part 2

Next we will use the following equalities

$$\frac{\partial A\beta}{\partial \beta'} = A \qquad \text{and} \qquad \frac{\partial \beta' A \beta}{\partial \beta'} = \beta'(A + A')$$

to calculate the first derivative of $L$ with respect to $\alpha'$.

$$\frac{\partial L(\alpha)}{\partial \alpha'} = 0 - 2y'X + \alpha'(X'X + X'X) = -2y'X + 2\alpha' X'X.$$

### 1.3 Part 3

Now, to minimize the $L$ function, we will solve the first order condition equation $\frac{\partial L(\alpha)}{\partial \alpha'} = 0$.

$$\begin{aligned}
\frac{\partial L(\alpha)}{\partial \alpha'} &= 0 \\
-2y'X + 2\alpha' X'X &= 0 \\
\alpha' X'X &= y'X \quad / \cdot (X'X)^{-1} \\
\alpha' &= y'X(X'X)^{-1} \\
\alpha &= (X'X)^{-1}X'y
\end{aligned}$$

So $\hat{\alpha} = (X'X)^{-1}X'y$ is the LS estimator of the model parameter.

The first vector is the theoretical estimator of $\alpha$ and the second is the estimator obtained with R's linear regression model:

| | theoretical | using lm() |
|---|---|---|
| alpha_1 | 2.05859906656869 | 2.05859906656869 |
| alpha_2 | 1.07475575539487 | 1.07475575539487 |
| alpha_3 | 0.889740635967315 | 0.889740635967315 |

Table 1: Estimator of alpha

### 1.4 Part 4

We are using unbiased estimator for the variance of residuals

$$\hat{\sigma}^2 = \frac{u'u}{N-K},$$

where in our case $N = 100$ and $K = 3$.

| theoretical | using lm() |
|---|---|
| 0.80202 | 0.80210 |

Table 2: Variance of residuals

The first row is the theoretical estimator of $\sigma^2$ and the second is the squared residual standard error obtained with R's linear regression model function.

### 1.5 Part 5

We assume that the residuals are uncorrelated and homoscedastic. The variance-covariance matrix of LS estimator is

$$\hat{\Sigma}_{\hat{\alpha}} = \hat{\sigma}^2(X'X)^{-1},$$

where $\hat{\sigma}^2$ is the estimator of variance of the residuals.

We can calculate the variance-covariance matrix using R

```
##              V2           V3           V4
## V2  0.004708172 -0.002019469 -0.002544603
## V3 -0.002019469  0.004516371 -0.002430459
## V4 -0.002544603 -0.002430459  0.005633808
##              V2           V3           V4
## V2  0.004708172 -0.002019469 -0.002544603
## V3 -0.002019469  0.004516371 -0.002430459
## V4 -0.002544603 -0.002430459  0.005633808
```

The first result is the theoretical estimator of $\Sigma$ and the second is the one obtained with R's linear regression model function.

Variance-covariance matrix for $\sqrt{N}\hat{\alpha}$ is equal to

$$\hat{\Sigma}_{\sqrt{N}\hat{\alpha}} = N\hat{\sigma}^2(X'X)^{-1},$$

and we can calculate it in R

```
##            V2         V3         V4
## V2  0.4708172 -0.2019469 -0.2544603
## V3 -0.2019469  0.4516371 -0.2430459
## V4 -0.2544603 -0.2430459  0.5633808
```

### 1.6 Part 6

The $t$-statistic tests the hypothesis $H_0: \ \alpha_i = 0$, $H_1: \ \alpha_i \neq 0$. The $t$-ratio is the ratio of the sample regression coefficient to its standard error. So

$$t_{\hat{\alpha}_i} = \frac{\hat{\alpha}_i}{\sqrt{Var\hat{\alpha}_i}} \quad \text{and} \quad t_{\hat{\alpha}_i} \sim t(N-K) = t(100-3) = t(97).$$

| | theoretical | using lm() |
|---|---|---|
| t.alpha1.hat | 30.0016825108455 | 30.0016825108456 |
| t.alpha2.hat | 15.9924488168616 | 15.9924488168616 |
| t.alpha3.hat | 11.8539317172235 | 11.8539317172235 |

Table 3: t-ratios of the parameters

## 2 Exercise 2

In this exercise we assume that $\alpha_1 + \alpha_2 + \alpha_3 = 0$ and $\alpha_2 - \alpha_3 = 0$.

### 2.1 Part 1

We know that the restriction matrix $R$ satisfies equation $R\alpha = r$. In this case

$$R \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad R = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

Rank of the restriction matrix is $rank(R) = 2$, because there are two linearly independent vectors in matrix $R$, $[1, 1, 1]$ and $[0, 1, -1]$.

## 2.2 Part 2

To express vector $\alpha$ and the loss function $L$ as functions of $\alpha_3$, we will first express $\alpha_1$ and $\alpha_2$ as functions of $\alpha_3$.

$$\begin{cases} \alpha_1 + \alpha_2 + \alpha_3 = 0 \\ \alpha_2 - \alpha_3 = 0 \end{cases} \Leftrightarrow \begin{cases} \alpha_1 = -\alpha_2 - \alpha_3 \\ \alpha_2 = \alpha_3 \end{cases} \Leftrightarrow \begin{cases} \alpha_1 = -2\alpha_3 \\ \alpha_2 = \alpha_3 \end{cases}$$

So

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} -2\alpha_3 \\ \alpha_3 \\ \alpha_3 \end{bmatrix} = \alpha_3 \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

and

$$L(\alpha_3) = y'y - 2y'X\alpha + \alpha'X'X\alpha = y'y - 2y'X\alpha_3 \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} + \alpha_3 \begin{bmatrix} -2 & 1 & 1 \end{bmatrix} X'X\alpha_3 \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} =$$

$$= y'y - 2\alpha_3 y'X \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} + \alpha_3^2 \begin{bmatrix} -2 & 1 & 1 \end{bmatrix} X'X \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

## 2.3 Part 3

Estimator of $\alpha_3$ is equal to

$$\hat{\alpha}_3 = \arg\min_{\alpha_3} L(\alpha_3)$$

We can calculate

$$\frac{\partial L(\alpha_3)}{\partial \alpha_3} = 0 - 2y'X \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} + 2\alpha_3 \begin{bmatrix} -2 & 1 & 1 \end{bmatrix} X'X \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

From F.O.C and the above formula we obtain

$$\hat{\alpha}_3 \begin{bmatrix} -2 & 1 & 1 \end{bmatrix} X'X \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = y'X \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

$$\hat{\alpha}_3 = y'X \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \left( \begin{bmatrix} -2 & 1 & 1 \end{bmatrix} X'X \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \right)^{-1}$$

For data from exercise 1, this estimator is equal to

```
## [1] -0.6334175
```

Now we can calculate variance of estimator $\hat{\alpha}_3$. Transforming the model as follows, we obtain $u$.

$$
\begin{aligned}
y &= X\hat{\alpha} + u \\
y &= \hat{\alpha}_3 X \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} + u \\
u &= y - \hat{\alpha}_3 X \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}
\end{aligned}
$$

Let $\tilde{X} = X \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$. Using $\hat{\sigma}^2 = \frac{u'u}{N-K}$ we can calculate the variance-covariance matrix of $\hat{\alpha}_3$, namely $\hat{\Sigma}_{\hat{\alpha}_3} = \hat{\sigma}^2(\tilde{X}'\tilde{X})^{-1}$. The exact value of $\hat{\Sigma}_{\hat{\alpha}_3}$ for our data is equal to

```
##         [,1]
## [1,] 0.23171
```

## 2.4 Part 4

We know that

$$
\begin{cases} \hat{\alpha}_1 = -2\hat{\alpha}_3 \\ \hat{\alpha}_2 = \hat{\alpha}_3 \end{cases} \quad \Rightarrow \quad \hat{\alpha} = \begin{bmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ \hat{\alpha}_3 \end{bmatrix} = \begin{bmatrix} -2\hat{\alpha}_3 \\ \hat{\alpha}_3 \\ \hat{\alpha}_3 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \hat{\alpha}_3
$$

and

$$
\begin{gathered}
Var(\hat{\alpha}_1) = Var(-2\hat{\alpha}_3) = 4Var(\hat{\alpha}_3) \\
Var(\hat{\alpha}_2) = Var(\hat{\alpha}_3) \\
Cov(\hat{\alpha}_2, \hat{\alpha}_1) = Cov(\hat{\alpha}_1, \hat{\alpha}_2) = Cov(-2\hat{\alpha}_3, \hat{\alpha}_3) = -2Cov(\hat{\alpha}_3, \hat{\alpha}_3)) = -2Var(\hat{\alpha}_3) \\
Cov(\hat{\alpha}_3, \hat{\alpha}_1) = Cov(\hat{\alpha}_1, \hat{\alpha}_3) = Cov(-2\hat{\alpha}_3, \hat{\alpha}_3) = -2Cov(\hat{\alpha}_3, \hat{\alpha}_3)) = -2Var(\hat{\alpha}_3) \\
Cov(\hat{\alpha}_2, \hat{\alpha}_3) = Cov(\hat{\alpha}_3, \hat{\alpha}_2) = Cov(\hat{\alpha}_3, \hat{\alpha}_3) = Var(\hat{\alpha}_3)
\end{gathered}
$$

So the variance-covariance matrix of $\alpha$ is as follows

$$
\begin{gathered}
\hat{\Sigma}_{\hat{\alpha}} = \begin{bmatrix} Var(\hat{\alpha}_1) & Cov(\hat{\alpha}_1, \hat{\alpha}_2) & Cov(\hat{\alpha}_1, \hat{\alpha}_3) \\ Cov(\hat{\alpha}_2, \hat{\alpha}_1) & Var(\hat{\alpha}_2) & Cov(\hat{\alpha}_2, \hat{\alpha}_3) \\ Cov(\hat{\alpha}_3, \hat{\alpha}_1) & Cov(\hat{\alpha}_3, \hat{\alpha}_2) & Var(\hat{\alpha}_3) \end{bmatrix} = \\
= Var(\hat{\alpha}_3) \begin{bmatrix} 4 & -2 & -2 \\ -2 & 1 & 1 \\ -2 & 1 & 1 \end{bmatrix} = 0.23171 \begin{bmatrix} 4 & -2 & -2 \\ -2 & 1 & 1 \\ -2 & 1 & 1 \end{bmatrix}
\end{gathered}
$$

Now we want to calculate the $t$-ratios for $\hat{\alpha}_i$ $(i = 1, 2, 3)$

$$
t_{\hat{\alpha}_i} = \frac{\hat{\alpha}_i}{\sqrt{Var\hat{\alpha}_i}},
$$

where $t_{\hat{\alpha}_i} \sim t(N-K) = t(100-3) = t(97)$.

$$t_{\hat{\alpha}_2} = t_{\hat{\alpha}_3} = \frac{\hat{\alpha}_3}{\sqrt{Var(\hat{\alpha}_3)}}$$

$$t_{\hat{\alpha}_1} = \frac{-2\hat{\alpha}_3}{\sqrt{Var(-2\hat{\alpha}_3)}} = \frac{-2\hat{\alpha}_3}{\sqrt{4Var(\hat{\alpha}_3)}} = \frac{-2\hat{\alpha}_3}{2\sqrt{Var(\hat{\alpha}_3)}} = -t_{\hat{\alpha}_3}$$

| | |
|---|---|
| t.alpha1.hat | 1.31588409531758 |
| t.alpha2.hat | -1.31588409531758 |
| t.alpha3.hat | -1.31588409531758 |

Table 4: t-ratios of alpha1.hat, alpha2.hat and alpha1.hat

# 3 Exercise 3

Next, we consider the following model

$$y_n = \alpha_1 X_{1n} + \alpha_2 X_{2n} + \varepsilon_n$$

with $\alpha_1 = \alpha_2 = 1$ and $\varepsilon_n \sim N(0,1)$.

## 3.1 Part 1

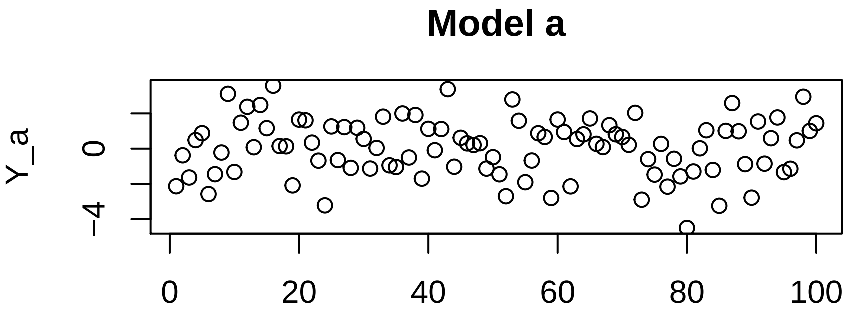
We will generate samples from above model in two cases:

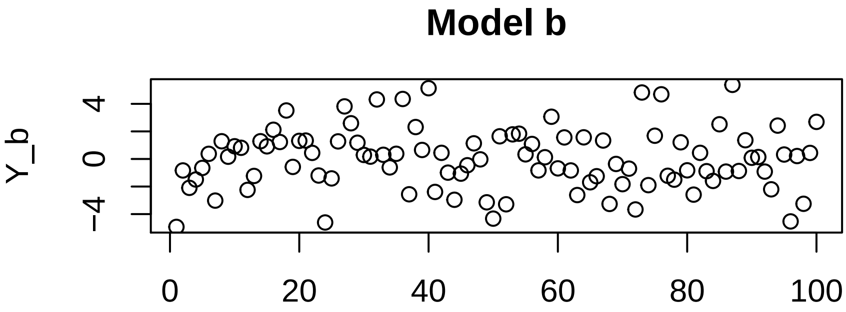
a) $X_1$ and $X_2$ are independent and $X \sim N(0, I_2)$,

b) $X_1$ and $X_2$ are dependent and $X \sim N(0, \Sigma)$.

We assume that in this exercise

$$\Sigma = \begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix}.$$




## 3.2 Part 2 and 3

We have the following models:

$$y_n = \alpha_1 X_{1n} + \alpha_2 X_{2n} + \varepsilon_n$$

$$y_n = \beta X_{1n} + u_n$$

We will use the formula from subsection 1.3 and Monte Carlo method with 1000 repetitions to obtain $\alpha_1$ and $\beta$ in those models.

Below we compare the results for different sample sizes.

| N | theoretical_alpha_1 | alpha_1_est | beta_est |
|---:|---:|---:|---:|
| 10.00000 | 1.00000 | 0.99509 | 0.99579 |
| 100.00000 | 1.00000 | 0.99811 | 1.00073 |
| 1000.00000 | 1.00000 | 1.00025 | 0.99946 |

Table 5: Estimator of alpha1 and beta for model a

| N | theoretical_alpha_1 | alpha_1_est | beta_est |
|---:|---:|---:|---:|
| 10.00000 | 1.00000 | 0.98538 | 1.79253 |
| 100.00000 | 1.00000 | 0.99330 | 1.79472 |
| 1000.00000 | 1.00000 | 1.00282 | 1.80132 |

Table 6: Estimator of alpha1 and beta for model b

### 3.3 Part 4

Estimator is consistent if

$$\hat{\theta} \to^P \theta.$$

From assumptions of the Least Squares method we know that

$$\hat{\alpha_1} \to^P \alpha_1.$$

But what about $\beta$? From *Econometric analysis* by W.H. Greene for

$$y = X_1\alpha_1 + X_2\alpha_2 + \varepsilon$$

we have the following normal equations

$$\begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} X_1'y \\ X_2'y \end{bmatrix}$$

By manipulating the first equation we get

$$\begin{aligned} \alpha_1 X_1'X_1 + X_1'X_2\alpha_2 = X_1'y \quad \Rightarrow \quad \alpha_1 &= \underbrace{(X_1'X_1)^{-1}X_1'y}_{\hat{\beta}} - (X_1'X_1)^{-1}X_1'X_2\alpha_2 \\ \alpha_1 &= \hat{\beta} - (X_1'X_1)^{-1}X_1'X_2\alpha_2 \end{aligned}$$

If $(X_1'X_1)^{-1}X_1'X_2\alpha_2$ is convergent to 0 (with probability), then $\hat{\beta} \to^P \alpha_1$. If $X_1'$ and $X_2$ are orthogonal, then $X_1'X_2 = 0$, and then $\hat{\beta} \to^P \alpha_1$.

In case (a) $X_1$ and $X_2$ are independent, so they are orthogonal, so $\hat{\beta}$ is consistent. In case (b) $X_1$ and $X_2$ are dependent, so they are not orthogonal, so $\hat{\beta}$ is not consistent.