Estimation theory – Report 3

Marta Frankowska, 208581
Agnieszka Szkutek, 208619

November 23, 2017

Contents

1 Model

In both exercises we will be using the Factor model

$$Y_{T\times N} = F_{T\times K} \cdot \lambda_{K\times N} + e_{T\times N},$$

where

- $Y_{T\times N}$ panel of observations
- $F_{T\times K}$ matrix of common (latent) factors
- $\lambda_{K\times N}$ matrix of loadings
- $e_{T\times N}$ panel of specific components

To calculate F and λ we use the following formulas:

$$\hat{F} = \sqrt{T} V_{1:K} \quad \text{and} \quad \hat{\lambda} = \frac{\hat{F}'Y}{T},$$

where $V_{1:K}$ are eigen vectors of YY' corresponding to the K largest eigenvalues.

1.1 Selecting optimal number of factors

Notation:

- $K = 1, 2, \ldots, K_{\max}$ - the number of factors,
- $e^{(K)}$ - the individual components for K factors,
- $V(K) = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}\left(e_{ti}^{(K)}\right)^2$,
- $\hat{\sigma}^2 = V(K_{\max})$ - consistent estimator of variance.

Information criteria:

- $PC_1(K) = V(K) + K\hat{\sigma}^2\frac{N+T}{NT}\ln\frac{NT}{N+T}$,
- $IPC_1(K) = \log V(K) + K\frac{N+T}{NT}\ln\frac{NT}{N+T}$.

Algorithm:

1. Set $K_{\max}$;
2. Compute $IC(K)$ for $K = 1, \ldots, K_{\max}$;
3. Choose $\hat{K}$ such that $IC(\hat{K}) = \min_{1\leq K\leq K_{\max}} IC(K)$.

```
factor.model.est <- function(Y,K_max) #function returning which K we should choose
  {
  T <- nrow(Y)
  N <- ncol(Y)

  eigen.decomp <- eigen(Y %*% t(Y)) #
  eigen.values <- eigen.decomp$values
  eigen.vectors <- eigen.decomp$vector

  # we calculate F, lambda and e for K_max
  F <- sqrt(T)*eigen.vectors[,1:K_max]
  lambda <- t(F)%*%Y/T
  e <- Y - F%*%lambda
  sigma2.hat <- sum(e^2)/(N*T)

  PC1 <- 1:K_max
  IPC1 <- 1:K_max
  for (K in 1:K_max)
  {
   # we calculate F, lambda and e for K
   F <- sqrt(T)*eigen.vectors[,1:K]
   lambda <- t(F)%*%Y/T
   e <- Y - F%*%lambda
   V <- sum(e^2)/(N*T)
   # we calculate PC1 and IPC1 for K
```

```
   PC1[K] <- V + K*sigma2.hat*((N+T)/(N*T))*log(N*T/(N+T))
   IPC1[K] <- log(V) + K*((N+T)/(N*T))*log(N*T/(N+T))
  }
  #we choose minimum for PC1 and IPC1
  min_PC1 <- min(PC1)
  min_IPC1 <- min(IPC1)
  # and looking for corresponding K
  PC1_K <- which(PC1 == min_PC1)
  IPC1_K <- which(IPC1 == min_IPC1)
  return (list(PC1_K,IPC1_K))
}
```

2 Exercise 1

We will be using data from file *dataLab3.xlsx*, where Y size is $T \times N = 100 \times 100$. To calculate the number of factors K we will use function *factor.model.est*.

2.1 Part 1

First, we calculate the number of factors for the whole sample. The function returns $\hat{K} = 3$ for both PC_1 and IPC_1. Share of explained variance

$$\frac{\sum_{i=1}^{K} \gamma_i}{\sum_{i=1}^{T} \gamma_i},$$

where γ_i are the eigenvalues of YY'. It can be used to choose the number of factors.

The plot shows the variability of the factors

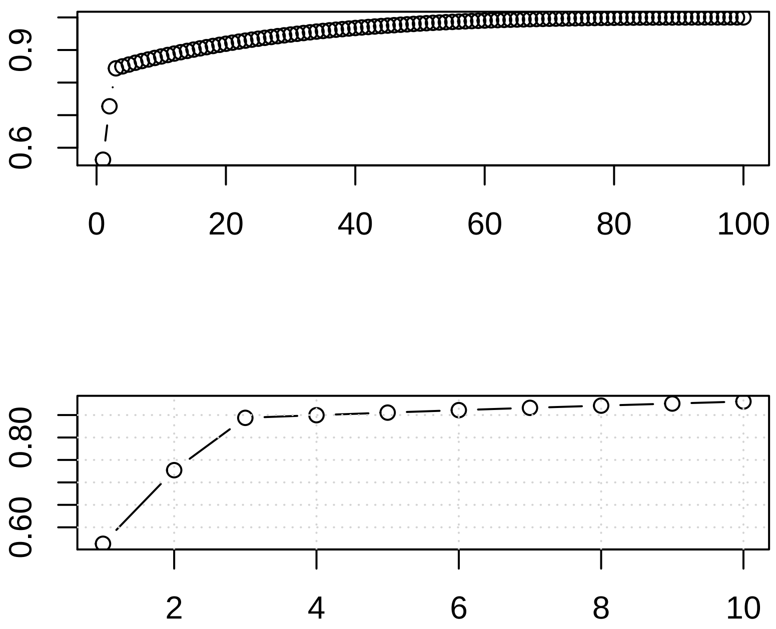

Figure 1: Share of explained variance depending on K

We can observe that for K we calculated from Information Criteria, the share of explained variance is more than 0.8.

Now we will compare estimated number of factors for the whoe sample, the first 20 columns and the first 20 rows.

	PC1	IPC1
whole sample	3.00	3.00
first 20 columns	9.00	3.00
first 20 rows	9.00	3.00

Table 1: Comparison of results

We can observe that ...

3 Exercise 2

In this exercise we will be working with data representing electricity prices from the balancing market. Each row describes the day, whereas the column describes the hour.

3.1 Part 1

We transform the data into logarithms and calculate mean for each column. Then we subtract the mean from each column.

```
data2 <- read_excel('RB.xlsx', col_names = FALSE)
Y <- as.matrix(data2)
#1
N <- ncol(Y)
T <- nrow(Y)
log_Y <- log(Y)
new_Y <- sweep(log_Y, 2, colMeans(log_Y))
```

3.2 Part 2

The plot shows the variability of the factors

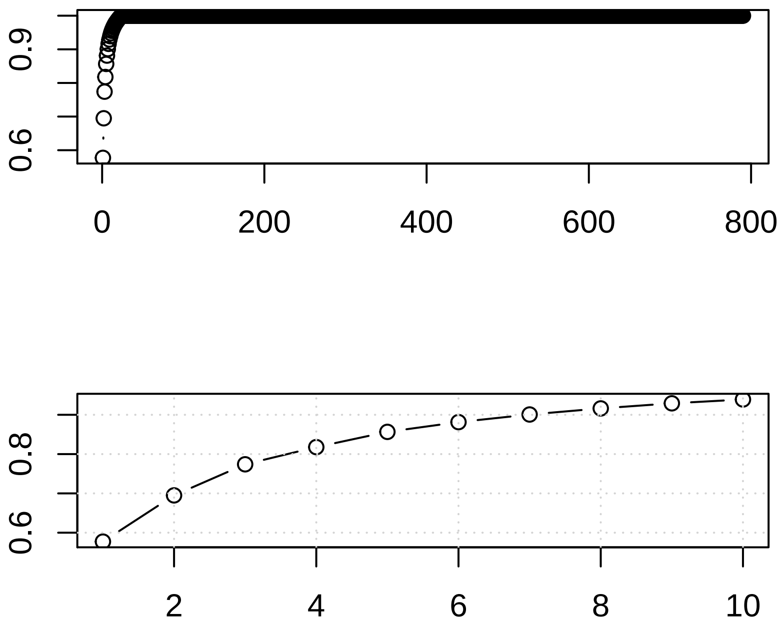

Figure 2: Share of explained variance

We can observe that if we want to have the Factor model which explains at least 80% of panel variability, we have to choose $K = 4$.

3.3 Part 3

We want to compute the information criteria with $K_{\max} = 8$. They suggest the following number of factors:

	Suggested no. of factors
PC1	8.00
IPC1	8.00

Table 2: Suggested number of factors