Data science in Economics. HA2.
Shumakov Anton, group 5

№1

$\hat{y} = \hat{\beta}_1 x + \hat{\beta}_2 x$ $\qquad$ $loss(\hat{\beta}) = (y-\hat{y})^T(y-\hat{y}) + \lambda \hat{\beta}^T\hat{\beta}$

a) $loss(\hat{\beta}_1) = (y - (\hat{\beta}_1 x + \hat{\beta}_2 x))^T (y - \hat{\beta}_1 x - \hat{\beta}_2 x) + \lambda \hat{\beta}_1^T \hat{\beta}_1.$

F.O.C. $\frac{\partial loss}{\partial \hat{\beta}_1} = 0.$

$\frac{\partial (y^Ty - y^T\hat{\beta}_1 x - y^T\hat{\beta}_2 x + (\hat{\beta}_1 x)^T(\hat{\beta}_1 x) + (\hat{\beta}_1 x)^T(\hat{\beta}_2 x) + (\hat{\beta}_2 x)^T(\hat{\beta}_2 x) - \lambda \hat{\beta}_1^T\hat{\beta}_1)}{\partial \hat{\beta}_1} = -x^Ty +$

$+ 2(x^Tx)\hat{\beta}_1 + x^T\hat{\beta}_2 x - 2\lambda\hat{\beta}_1 = 0 \Rightarrow \quad \hat{\beta}_1(2(x^Tx) - 2\lambda) = x^Ty - x^T\hat{\beta}_2 x$

$\hat{\beta}_1 = (2((x^Tx) - \lambda))^{-1}(x^Ty - x^T\hat{\beta}_2 x).$

Analogically for $\hat{\beta}_2$: $\qquad \hat{\beta}_2 = (2((x^Tx) - \lambda))^{-1}(x^Ty - x\hat{\beta}_1 x^T)$

b) If $\lambda \to \infty$, then $(2((x^Tx) - \lambda)) \to -\infty \Rightarrow (2((x^Tx) - \lambda))^{-1} \to 0. \Rightarrow \hat{\beta}_1, \hat{\beta}_2 \to 0.$

c) $\hat{\beta}_1 + \hat{\beta}_2 = (2((x^Tx) - \lambda))^{-1}(x^Ty - x^T\hat{\beta}_2 x + x^Ty - x\hat{\beta}_1 x^T) = (2((x^Tx) - \lambda))^{-1}(2x^Ty - x^T\hat{\beta}_2 x - x\hat{\beta}_1 x^T)$

Since the first part is the same $(2((x^Tx) - \lambda))^{-1}$, then as $\lambda \to \infty$, $\hat{\beta}_1 + \hat{\beta}_2 \to 0$

№2.

$y = X\beta + u,$

$E(u|X) = 0$

$Var(u|X) = \sigma^2 W$ $\quad$, $W \neq I$

$\hat{\beta}$ - OLS estimator of β. $\Rightarrow \hat{\beta} = (X^TX)^{-1}X^Ty$

a) $E(\hat{\beta}|X) = E((X^TX)^{-1}X^Ty|X) = (X^TX)^{-1}X^TE(X\beta + u|X) = (X^TX)^{-1}X^TX\beta + 0 = \beta$

$E(\hat{\beta}) = E((X^TX)^{-1}X^Ty) = (X^TX)^{-1}X^TE(X\beta + u) = \beta + (X^TX)^{-1}X^TE(u)$

b) $Var(\hat{\beta}|X) = Var((X^TX)^{-1}X^Ty|X) = (X^TX)^{-1}X^T\, Var(y|X) \cdot [(X^TX^{-1})X^T]^T =$

$Var(y|X) = Var(X\beta + u|X) = Var(u|X)$

$= \sigma^2 W(X^TX^{-1}) \cancel{X^TX(X^TX)} = \sigma^2 W(X^TX)^{-1}$

c) The standard confidence interval will be invalid because $W \neq I$.

d) $Cov(y, \hat{\beta}|X) = Cov(X\beta + u, \hat{\beta}|X) = Cov(X\beta, (X^TX)^{-1}X^Ty|X) + Cov(u, (X^TX)^{-1}X^Ty|X) =$

$= \beta Cov(X, \cancel{(X^TX)^{-1}X^TX}\beta + u|X) + Cov(u, \cancel{(X^TX)^{-1}X^TX}\beta + u|X) = \beta Cov(X, \overset{=0}{\beta}|X) + \beta Cov(X, u|X) +$

$Var(u|X) = \beta(E(\overset{=0}{\cancel{uX}}|X) - E\overset{=0}{(u|X)}\overset{=0}{E(X|X)}) + \sigma^2 W = \sigma^2 W.$

13.

$$X = \begin{pmatrix} 2 & 1 \\ -1 & 2 \\ 1 & 1 \end{pmatrix} \qquad X^T = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 2 & 1 \end{pmatrix}$$

$$X^TX = \begin{pmatrix} 2 & 1 \\ -1 & 2 \\ 1 & 1 \end{pmatrix} \times \begin{pmatrix} 2 & -1 & 1 \\ 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 0 & 3 \\ 0 & 5 & 1 \\ 3 & 1 & 2 \end{pmatrix}$$

$$\det(X^TX - \lambda I) = \begin{vmatrix} 5-\lambda & 0 & 3 \\ 0 & 5-\lambda & 1 \\ 3 & 1 & 2-\lambda \end{vmatrix} = (5-\lambda)(5-\lambda)(2-\lambda) + 0 + 0 - 9(5-\lambda) - (5-\lambda) - 0 =$$

$$= -\lambda^3 + 12\lambda^2 - 35\lambda = -\lambda(\lambda^2 - 12\lambda + 35) = -\lambda(\lambda - 5)(\lambda - 7) = 0$$

$$D = \begin{vmatrix} 0 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 7 \end{vmatrix}$$

b) SVD.

$X = \begin{pmatrix} 2 & 1 \\ -1 & 2 \\ 1 & 1 \end{pmatrix} = U \cdot D \cdot V^T$, where $U^TU = I\ [n \times n]$

$V^TV = I\ [k \times k]$

D - diagonal $[n \times k]$

$$X = \begin{pmatrix} 2 & 1 \\ -1 & 2 \\ 1 & 1 \end{pmatrix} = \underbrace{\overset{3\times 3}{\begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{pmatrix}}}_{U} \times \overset{3\times 2}{\begin{pmatrix} d_{11} & 0 \\ 0 & d_{22} \\ 0 & 0 \end{pmatrix}} \times \underbrace{\overset{2\times 2}{\begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}}}_{V^T}$$

"Some magic in python"

$$U = \begin{pmatrix} -0,8 & 0,32 & -0,51 \\ -0,26 & -0,95 & -0,17 \\ -0,53 & 1,66 & 0,84 \end{pmatrix}$$

Code:

```
import numpy as np
import numpy.linalg as lt
X = np.array([[2, 1],
              [-1, 2],
              [1, 1]])
u, D, V = np.linalg.svd(X, full_\
                        matrices = True)
```

$$V = \begin{pmatrix} -0,7 & -0,7 \\ 0,7 & -0,7 \end{pmatrix}$$

$$D = \begin{pmatrix} 2,64 & 0 \\ 0 & 2,24 \end{pmatrix}$$