

Problem 2

Terentev Vlad Gr. 4.

$$L(\hat{\beta}) = (y - \hat{y})^T (y - \hat{y}) + \lambda \hat{\beta}^T \hat{\beta}, \quad \hat{y} = \hat{\beta}_1 x + \hat{\beta}_2 x$$

$$L(\hat{\beta}) = (y - x(\hat{\beta}_1 + \hat{\beta}_2))^T (y - x(\hat{\beta}_1 + \hat{\beta}_2)) + \lambda (\hat{\beta}_1^2 + \hat{\beta}_2^2) \rightarrow \min$$

$$L(\hat{\beta}) = \|y - x(\hat{\beta}_1 + \hat{\beta}_2)\|_2^2 + \lambda (\hat{\beta}_1^2 + \hat{\beta}_2^2) = \sum_{i=1}^n (y_i - x_i(\hat{\beta}_1 + \hat{\beta}_2))^2 + \lambda (\hat{\beta}_1^2 + \hat{\beta}_2^2)$$

$$\begin{cases} \frac{\partial L}{\partial \hat{\beta}_1} = \sum_{i=1}^n 2(y_i - x_i \hat{\beta}_1 - x_i \hat{\beta}_2) \cdot (-x_i) + 2\hat{\beta}_1 \lambda = 0 \Rightarrow -\sum_{i=1}^n x_i y_i + \hat{\beta}_1 \sum_{i=1}^n x_i^2 + \hat{\beta}_2 \sum_{i=1}^n x_i^2 + \hat{\beta}_1 \lambda = 0 \\ \frac{\partial L}{\partial \hat{\beta}_2} = \sum_{i=1}^n 2(y_i - x_i \hat{\beta}_1 - x_i \hat{\beta}_2) \cdot (-x_i) + 2\hat{\beta}_2 \lambda = 0 \Rightarrow -\sum_{i=1}^n x_i y_i + \hat{\beta}_1 \sum_{i=1}^n x_i^2 + \hat{\beta}_2 \sum_{i=1}^n x_i^2 + \hat{\beta}_2 \lambda = 0 \end{cases}$$

$$\Rightarrow \begin{cases} \hat{\beta}_1 \left(\sum_{i=1}^n x_i^2 + \lambda \right) = \sum_{i=1}^n x_i y_i - \hat{\beta}_2 \sum_{i=1}^n x_i^2 \\ \hat{\beta}_2 \left(\sum_{i=1}^n x_i^2 + \lambda \right) = \sum_{i=1}^n x_i y_i - \hat{\beta}_1 \sum_{i=1}^n x_i^2 \end{cases} \Rightarrow \begin{cases} \hat{\beta}_1^{OLS} = \frac{\sum x_i y_i - \hat{\beta}_2 \sum x_i^2}{\sum x_i^2 + \lambda} \\ \hat{\beta}_2^{OLS} = \frac{\sum x_i y_i - \hat{\beta}_1 \sum x_i^2}{\sum x_i^2 + \lambda} \end{cases}$$

Symmetric system:

$$\hat{\beta}_1 = \hat{\beta}_2 = k \Rightarrow k \left(\sum x_i^2 + \lambda \right) = \sum x_i y_i - k \sum x_i^2$$

$$\Rightarrow k = \frac{\sum x_i y_i}{2 \sum x_i^2 + \lambda} \Rightarrow \hat{\beta}_1^{OLS} = \hat{\beta}_2^{OLS} = \frac{\sum x_i y_i}{2 \sum x_i^2 + \lambda}$$

$$\hat{\beta}_1^{OLS} = \hat{\beta}_2^{OLS} = \frac{X Y^T}{2 X X^T + \lambda} \quad X - \text{row vector}$$

b) if $\lambda \rightarrow \infty$

$$\lim_{\lambda \rightarrow \infty} \hat{\beta}_1^{OLS} = \lim_{\lambda \rightarrow \infty} \hat{\beta}_2^{OLS} = \lim_{\lambda \rightarrow \infty} \frac{C_0}{C_1 + \lambda} = 0 \Rightarrow \text{both } \hat{\beta}^{OLS} \text{ will be equal to } 0$$

$$c) \lim_{\lambda \rightarrow 0} \frac{2 \sum x_i y_i}{2 \sum x_i^2 + \lambda} = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{X Y^T}{X X^T}$$

Problem 2

$$y = X\beta + u, \quad E(u|x) = 0 \quad \begin{matrix} X \\ n \times k \end{matrix} \quad \text{rk}(X) = k \quad V(u|x) = \sigma^2 W$$

a) OLS $\hat{\beta} = (X^T X)^{-1} X^T y$

$$E(\hat{\beta}|x) = E[(X^T X)^{-1} X^T y | x] = E[(X^T X)^{-1} X^T \cdot [X\beta + u] | x] =$$

$$= E[(X^T X)^{-1} X^T X \beta + (X^T X)^{-1} X^T u | x] = E[(X^T X)^{-1} X^T X \beta | x] + E[(X^T X)^{-1} X^T u | x] = (X^T X)^{-1} X^T X \beta + (X^T X)^{-1} X^T E(u|x) = \beta$$

$$E(\hat{\beta}) = E[E(\hat{\beta}|x)] = E(\beta) = \beta$$

b) $V(\hat{\beta}|x) = V[(X^T X)^{-1} X^T X \beta + (X^T X)^{-1} X^T u] =$
 $= V(\beta) + V[(X^T X)^{-1} X^T u] = (X^T X)^{-1} X^T V(u) (X^T X)^{-1} X^T =$
 $= (X^T X)^{-1} X^T \sigma^2 W X (X^T X)^{-1}$

c) Since, $W \neq I \Rightarrow V(\hat{\beta}|x) \neq \sigma^2 (X^T X)^{-1} \Rightarrow$

standard confidence intervals for β are gonna be invalid. This can be due to presence of heteroscedasticity $V(u|x) = \sigma^2$ or autocorrelation. Moreover, nothing is said about distribution of error \rightarrow to perform tests errors must be distributed normally.

d) $\text{Cov}(y, \hat{\beta} | x) = \text{Cov}(X\beta + u, (X^T X)^{-1} X^T y | x) =$
 $= \text{Cov}(X\beta + u, \beta + (X^T X)^{-1} X^T u | x) = \text{Cov}(X\beta, \beta | x) +$
 $+ \text{Cov}(u, \beta | x) + \text{Cov}(X\beta, (X^T X)^{-1} X^T u | x) + \text{Cov}(u, (X^T X)^{-1} X^T u | x) =$
 $= X V(\beta) + X \text{Cov}(\beta, u | x) X (X^T X)^{-1} + V(u) \cdot X (X^T X)^{-1} =$
 $= \sigma^2 W X (X^T X)^{-1}$

Problem 3

$$X = \begin{pmatrix} 2 & 1 \\ -1 & 2 \\ 1 & 1 \end{pmatrix} \quad a) \quad X^T \cdot X = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 1 \\ 1 & 6 \end{pmatrix}$$

3×2 2×3 2×2

Diagonalisation: eigenvalues + eigenvectors

$$X^T X - \lambda = \begin{pmatrix} 6-\lambda & 1 \\ 1 & 6-\lambda \end{pmatrix} \quad \det(X^T X - \lambda) =$$

$$= (6-\lambda)^2 - 1 =$$

$$= 35 - 12\lambda + \lambda^2$$

Eigenvalues:

$$\lambda^2 - 12\lambda + 35 = 0$$

$$\Delta = 144 - 140 = 2^2 \quad \Rightarrow \quad \lambda_{1,2} = \begin{cases} \frac{12+2}{2} = 7 \\ \frac{12-2}{2} = 5 \end{cases}$$

Eigenvectors:

$$\lambda_1 = 7 \quad X^T X - 7 = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \Rightarrow \begin{cases} -a + b = 0 \\ a - b = 0 \end{cases} \quad a = b = 1 \quad v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\lambda_2 = 5 \quad X^T X - 5 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \Rightarrow a + b = 0 \quad a = -b \quad a = -1 \quad v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\Rightarrow \Lambda = \begin{pmatrix} 7 & 0 \\ 0 & 5 \end{pmatrix}; \quad X^T X = Q \Lambda Q^{-1}$$

$$Q = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \quad Q^{-1} = \begin{pmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{pmatrix}; \quad X^T X = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 7 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{pmatrix}$$

$$b) \quad \|v_1\| = \sqrt{2} \quad \|v_2\| = \sqrt{2} \quad u_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \quad u_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \quad V = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

SVD of X : $X = U \Sigma U^T$

$$XX^T = \begin{pmatrix} 5 & 0 & 3 \\ 0 & 5 & 1 \\ 3 & 1 & 2 \end{pmatrix} \Rightarrow \det(XX^T - \lambda I) = -35\lambda + 12\lambda^2 - \lambda^3 = 0$$

$$-\lambda^3 + 12\lambda^2 - 35\lambda = 0$$

From characteristic equation:

$$\left. \begin{matrix} \lambda_1 = 7 \\ \lambda_2 = 5 \\ \lambda_3 = 0 \end{matrix} \right\} \text{ only non-zero eigenvalues. } \Sigma = \begin{pmatrix} \sqrt{7} & 0 & 0 \\ 0 & \sqrt{5} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\hookrightarrow U_1 = \begin{pmatrix} 3/2 \\ 1/2 \\ 1 \end{pmatrix}, U_2 = \begin{pmatrix} -1/3 \\ 1 \\ 0 \end{pmatrix}, U_3 = \begin{pmatrix} -3/5 \\ -1/5 \\ 0 \end{pmatrix}$$

Columns of U are normalized

$$\text{eigenvectors: } U = \begin{pmatrix} \frac{3\sqrt{14}}{14} & -\frac{\sqrt{10}}{10} & -\frac{3\sqrt{35}}{35} \\ \frac{\sqrt{14}}{14} & \frac{3\sqrt{10}}{10} & -\frac{\sqrt{35}}{35} \\ \frac{\sqrt{14}}{7} & 0 & -\frac{\sqrt{35}}{7} \end{pmatrix} \Rightarrow \text{SVD: } X = U \Sigma U^T$$

c) $X_1 = \sigma_1 u_1 v_1^T$ - rank 1 approximation

where σ_1 - largest singular value

u_1 - first left singular vector

v_1 - first right singular vector

$$\sigma_1 = \sqrt{7} \quad u_1 = \begin{pmatrix} 3/\sqrt{14} \\ 1/\sqrt{14} \\ 2/\sqrt{14} \end{pmatrix} \quad v_1^T = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

$$X_1 = \sqrt{7} \begin{pmatrix} 3/\sqrt{14} \\ 1/\sqrt{14} \\ 2/\sqrt{14} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 3/\sqrt{2} \\ 1/\sqrt{2} \\ 2/\sqrt{2} \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 3/2 & 3/2 \\ 1/2 & 1/2 \\ 1 & 1 \end{pmatrix}$$

Problem 4

column vectors

$$d) \begin{cases} \max \|Xw\|^2 \\ \text{s.t. } \|w\|^2 = 1 \end{cases} \Rightarrow \mathcal{L}(w, \lambda) = \|Xw\|^2 - \lambda(\|w\|^2 - 1)$$

$$= (Xw)^T (Xw) - \lambda(w^T w - 1)$$

e) FOC:

$$= \underbrace{w^T X^T X w}_{\text{symmetric}} - \lambda(\underbrace{w^T w}_{1} - 1)$$

$$d\mathcal{L} = d(w^T X^T X w) - d(\lambda w^T w) - d(\lambda)$$

$$d\mathcal{L} = d(w^T X^T X w) - \lambda d(w^T w) - d(\lambda) \xrightarrow{0}$$

$$d\mathcal{L} = 2w^T X^T X dw - 2\lambda w^T dw$$

$$d\mathcal{L} = 0 \Rightarrow w^T X^T X dw - \lambda w^T dw = 0$$

$$(w^T X^T X - \lambda w^T) dw = 0 \rightarrow \text{must hold } \forall dw$$

$$\Rightarrow w^T X^T X - w^T = 0 \Rightarrow (X^T X w)^T - \lambda w^T = 0$$

$$\Rightarrow \text{FOC: } X^T X w = \lambda w$$

\hookrightarrow corresponding eigenvalue
 $\hookrightarrow w$ eigenvector of $X^T X$

$$f) X = U \Sigma V^T$$

$$X^T X = (U \Sigma V^T)^T \cdot (U \Sigma V^T) = (V \Sigma^T U^T) \cdot (U \Sigma V^T) =$$

$$= V \underbrace{\Sigma^T U^T U}_{\text{orthogonal matrices}} \Sigma V^T = V \underbrace{\Sigma^T \Sigma}_{\text{diagonal matrices}} V^T = V \Sigma^2 V^T$$

orthogonal matrices $\Rightarrow U^T U = I$

$$X^T X = V \Sigma^2 V^T$$

\hookrightarrow columns of V are eigenvectors of $X^T X$
 \hookrightarrow diagonal values d_{ii}^2 - eigenvalues of $X^T X$

Plugging in to FOC: $V \Sigma^2 V^T w = \lambda w \Rightarrow \Sigma^2 V^T w = V^T \lambda w$

$$\Rightarrow \underbrace{\Sigma^2 V^T w}_A = \lambda \underbrace{V^T w}_A \Rightarrow d_{ii}^2 A_i = \lambda A_i \Rightarrow \text{we must select eigenvector corresponding to largest eigenvalue } d_{11}^2 \Rightarrow w = V_1$$