Teventev Vlad Gr. 4.

Problem 1

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

$$L(\hat{\beta}) = \left(y - X(\hat{\beta}_1+\hat{\beta}_2)\right)^T\left(y - X(\hat{\beta}_1+\hat{\beta}_2)\right) + \lambda(\hat{\beta}_1^2+\hat{\beta}_2^2) \longrightarrow \min$$

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

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

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

Symmetric system:

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

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

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

b) if $\lambda \to \infty$

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

c) $$\lim_{\lambda\to 0}\frac{2\sum X_i y_i}{2\sum X_i^2 + \lambda} = \frac{\sum X_i y_i}{\sum X_i^2} = \frac{X y^T}{X X^T}$$

Problem 2

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

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

$$E(\hat\beta|X) = E\left[(X^TX)^{-1}X^Ty|X\right] = E\left[(X^TX)^{-1}X^T\cdot[X\beta+u]|X\right] =$$

$$= E\left[\left[(X^TX)^{-1}X^TX\beta + (X^TX)^{-1}X^Tu\right]|X\right] = E\left((X^TX)^{-1}X^TX\beta\,|X\right) +$$

$$+ E\left((X^TX)^{-1}X^Tu|X\right) = \cancel{(X^TX)^{-1}X^TX}\beta + (X^TX)^{-1}X^T \cancelto{0}{E(u|X)} = \beta$$

$$E(\hat\beta) = E\left[E(\hat\beta|X)\right] = E(\beta) = \beta$$

b) $V(\hat\beta|X) = V\left[(X^TX)^{-1}X^TX\beta + (X^TX)^{-1}X^Tu\right] =$

$$= \cancelto{0}{V(\beta)} + V\left((X^TX)^{-1}X^Tu\right) = (X^TX)^{-1}X^T\,V(u)\left((X^TX)^{-1}X^T\right)^T =$$

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

c) Since, $W \neq I \Rightarrow V(\hat\beta|X) \neq \sigma^2(X^TX)^{-1} \Rightarrow$ standart confidence intervals for β are gonna be invalid. This can be due to presence of heteroscedasticity $V(u|X) = \sigma_i^2$ or autocorrelation. Moreover, nothing is said about distribution of error $\rightarrow$ to perform tests errors must be distributed normally.

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

$$= Cov\left(X\beta+u,\ \beta+(X^TX)^{-1}X^Tu\,|X\right) = Cov(X\beta,\beta|X) +$$

$$+ \cancelto{0}{Cov(u,\beta|X)} + Cov\left(X\beta,\ (X^TX)^{-1}X^Tu|X\right) + Cov\left(u,\ (X^TX)^{-1}X^Tu\,|X\right) =$$

$$= X\cancelto{0}{V(\beta)} + X\cancelto{0}{Cov(\beta, u|X)}\,X(X^TX)^{-1} + V(u)\cdot X(X^TX)^{-1} =$$

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

Problem 3

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

a) $$X^T \cdot X = \underset{3\times 2}{\begin{pmatrix} 2 & -1 & 1 \\ 1 & 2 & 1 \end{pmatrix}} \underset{2\times 3}{\begin{pmatrix} 2 & 1 \\ -1 & 2 \\ 1 & 1 \end{pmatrix}} = \underset{2\times 2}{\begin{pmatrix} 6 & 1 \\ 1 & 6 \end{pmatrix}}$$

Diagonalisation: eigenvalues + eigenvectors

$$X^TX - \lambda = \begin{pmatrix} 6-\lambda & 1 \\ 1 & 6-\lambda \end{pmatrix}$$

$$\det(X^TX - \lambda) = (6-\lambda)^2 - 1 = 35 - 12\lambda + \lambda^2$$

Eigenvalues:

$$\lambda^2 - 12\lambda + 35 = 0 \Rightarrow \lambda_{1,2} = \begin{cases} \frac{12+2}{2} = 7 \\ \frac{12-2}{2} = 5 \end{cases}$$

$$\Delta = 144 - 140 = 2^2$$

Eigenvectors:

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

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

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

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

b) $$\|v_1\| = \sqrt{2} \qquad \|v_2\| = \sqrt{2}$$

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

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

$$XX^T = \begin{pmatrix} 5 & 0 & 3 \\ 0 & 5 & 1 \\ 3 & 1 & 2 \end{pmatrix} \qquad \Downarrow \quad X V \Sigma^{-1} = U$$

From characteristic equation:

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

$$\hookrightarrow v_1 = \begin{pmatrix} 3/2 \\ 1/2 \\ 1 \end{pmatrix}; \quad v_2 = \begin{pmatrix} -1/3 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -3/5 \\ -1/5 \\ 0 \end{pmatrix},$$

Columns of U are normalized eigenvectors:

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

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

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

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

Problem 4

column vectors

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

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

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

e) FOC:

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

$d\mathcal{L} = d(w^T X^T X w) - \lambda\, d(w^T w) - \lambda\, \cancel{d(1)} \to 0$

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

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

$(w^T X^T X - \lambda w^T)\, dw = 0 \to$ must hold $\forall\, dw$

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

$\Rightarrow$ FOC: $X^T X w = \lambda w$

$\lambda \to$ corresponding eigenvalue

$w \to$ eigenvector of $X^T X$

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

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

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

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

$\to$ columns of V are eigenvectors of $X^T X$

$\to$ diagonal values d_{ii}^2 - eigenvalues of $X^T X$

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

$\Rightarrow \Sigma^2 \underbrace{V^T w}_{A} = \lambda \underbrace{V^T w}_{A} \Rightarrow d_{ii}^2 A_i = \lambda A_i \Rightarrow$ we must select eigenvector corresponding to largest eigenvalue $d_{11}^2 \Rightarrow w = v_1$