Akhmatukaev Dokka, gr 6

Q 1

$x(\hat\beta_1 + \hat\beta_2)$

a) $loss(\hat\beta) = (y-\hat y)^T (y-\hat y) + \lambda \hat\beta^T \hat\beta$, $\hat y = \hat\beta_1 x + \hat\beta_2 x$, $\hat\beta = \begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix}$,

let $\beta = \hat\beta_1 + \hat\beta_2$ then: $\hat\beta^T \hat\beta = \hat\beta_1^2 + \hat\beta_2^2$

$$loss(\hat\beta) = (y-\beta x)^T (y - \beta x) + \lambda(\hat\beta_1^2 + \hat\beta_2^2) =$$

$$= y^T y - 2\beta y^T x + \beta^2 x^T x$$

$$loss(\hat\beta) = A - 2\beta \underbrace{Cov(y, x)}_{\gamma} + \beta^2 I + \lambda(\hat\beta_1^2 + \hat\beta_2^2)$$

minimizing $loss(\hat\beta)$ by $\hat\beta_1$ and $\hat\beta_2$ will give us:

$$\frac{\partial loss(\hat\beta)}{\partial \hat\beta_1} = 2\hat\beta_1 \lambda = 0 \qquad \frac{\partial loss(\hat\beta)}{\partial \hat\beta_2} = 2\hat\beta_2 \lambda = 0$$

$$\hat\beta_1 = 0 \qquad \hat\beta_2 = 0$$

$$\hat\beta_1 = \hat\beta_2$$

minimizing results in equal coefficients and since:

$\hat\beta_1 + \hat\beta_2 = \beta$ and $\hat\beta_1 = \hat\beta_2 \Rightarrow \hat\beta_1 = \hat\beta_2 = \frac{\beta}{2}$ now plugging into loss:

$\Rightarrow \hat\beta_1^2 + \hat\beta_1^2 = 2\frac{\beta^2}{2}$

$$loss(\beta) = A - 2\beta\gamma + \beta^2 I + \lambda\frac{\beta^2}{2} = A - 2\beta\gamma + \beta^2\left(I + \frac{\lambda}{2}\right)$$

$$\frac{\partial loss(\beta)}{\partial \beta} = -2\gamma + 2\beta\left(I + \frac{\lambda}{2}\right) = 0$$

$$\beta = \frac{\gamma}{I + \frac{\lambda}{2}}$$

$$\hat\beta_1 = \hat\beta_2 = \frac{\beta}{2} = \frac{1}{2} \cdot \frac{\gamma}{I + \frac{\lambda}{2}} = \frac{\gamma}{I + \lambda}$$

$$\hat\beta_1 = \hat\beta_2 = \underline{\frac{\gamma}{I + \lambda}}$$

b) $\lim\limits_{\lambda \to \infty} \frac{\gamma}{I + \lambda_{\to\infty}} = 0 \Rightarrow \hat\beta_1 \to 0$ and $\hat\beta_2 \to 0$

c) $\lim\limits_{\lambda \to 0} \frac{\gamma}{I + \frac{\lambda}{2}_{\to 0}} = \frac{\gamma}{I} \Rightarrow \hat\beta_1 + \hat\beta_2 \to \frac{\gamma}{I}$

Q2

a) $y = X\beta + u$

$E(u|X) = 0$

$Var(u|X) = \sigma^2 W$

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

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

$= (X^T X)^{-1} X^T X \beta + 0 = \underline{\beta}$

$E(\hat{\beta}) = E(E(\hat{\beta}|X)) = \underline{\beta}$

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

$= (X^T X)^{-1} X \, Var(u|X) \cdot X (X^T X)^{-1} = (X^T X)^{-1} X \sigma^2 W X (X^T X)^{-1} =$

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

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

c) Since $W \neq I$ the error term is heteroskedastic which means that the standard OLS formula $\sigma^2 (X^T X)^{-1}$ is wrong and confidence intervals will generally be invalid.

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

$= Cov(X\beta + u, (X^T X)^{-1} X^T (X\beta + u) | X)$

~~Xβ is~~ $Cov(X\beta, (X^T X)^{-1} X^T X\beta | X) = 0$

$Cov(u, (X^T X)^{-1} X^T u | X) = (X^T X)^{-1} X^T \, Var(u|X) =$

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

Q3

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

$\det \begin{pmatrix} 6-\lambda & 1 \\ 1 & 6-\lambda \end{pmatrix} = \lambda^2 - 12 + 35 = 0$

$\lambda_1 = 7 \quad \lambda_2 = 5$

$\lambda_1 = 7$

$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \quad \begin{matrix} -x_1 + x_2 = 0 \\ x_1 = x_2 \end{matrix} \Rightarrow \begin{pmatrix} x_2 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - v_1$

$\lambda_2 = 5$

$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \begin{matrix} x_1 + x_2 = 0 \\ x_1 = -x_2 \end{matrix} \Rightarrow \begin{pmatrix} -x_2 \\ x_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} - v_2$

eigenvector matrix: $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} - V$ inverse: $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$

$A = V \Lambda V^{-1} \Rightarrow \begin{pmatrix} 6 & 1 \\ 1 & 6 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 7 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$

B) $SVD = U D V^T \quad \sigma_1 = \sqrt{\lambda_1} = \sqrt{7} \quad \sigma_2 = \sqrt{\lambda_2} = \sqrt{5}$

$X \vec{v_1} = \sigma_1 \vec{u_1} \rightarrow \vec{u_1} = \frac{1}{\sigma_1} X \vec{v_1} = \frac{1}{\sqrt{7}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{3\sqrt{14}}{14} \\ \frac{\sqrt{14}}{14} \\ \frac{\sqrt{14}}{7} \end{pmatrix}$

$X \vec{v_2} = \sigma_2 \vec{u_2} \rightarrow \vec{u_2} = \frac{1}{\sigma_2} X \vec{v_2} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} -\frac{\sqrt{10}}{10} \\ \frac{3\sqrt{10}}{10} \\ 0 \end{pmatrix}$

$SVD = \begin{pmatrix} \frac{3\sqrt{14}}{14} & -\frac{\sqrt{10}}{10} \\ \frac{\sqrt{14}}{14} & \frac{3\sqrt{10}}{10} \\ \frac{\sqrt{14}}{7} & 0 \end{pmatrix} \begin{pmatrix} \sqrt{7} & 0 \\ 0 & \sqrt{5} \\ \cancel{0} & \cancel{0} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} =$

$= \cancel{\begin{pmatrix} \frac{43}{2} & \frac{41}{2} \\ -1 & 2 \\ 1 & 1 \end{pmatrix}} \quad \underline{\begin{pmatrix} 2 & 1 \\ -1 & 2 \\ 1 & 1 \end{pmatrix}}$

[illegible]

c) $X_1 = \sigma_1 u_1 v_1^T = \sqrt{7} \begin{pmatrix} \frac{3\sqrt{14}}{14} \\ \frac{\sqrt{14}}{14} \\ \frac{\sqrt{14}}{7} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} =$

$= \underline{\begin{pmatrix} \frac{3}{2} & \frac{3}{2} \\ \frac{1}{2} & \frac{1}{2} \\ 1 & 1 \end{pmatrix}}$

Q4

c) $||Xw||^2 = w^T X^T X w$

s.t. $||w^2|| = w^T w = 1$

$L = w^T X^T X w - \lambda (w^T w - 1)$

e) F.O.C: $w^T w = \sum w_i^2 \quad \frac{d}{dw} [w^T w] = 2w$

$L'_w = 2 w X^T X - 2\lambda w = 0$

$X^T X w = \lambda w$ => w is eigenvector of $X^T X$ and λ is eigenvalue

$L'_\lambda = -(w^T w - 1) = 0$ => $||w||^2 = 1$

f) SVD of $X^T X = V D^T D V^T$ $\quad D = \begin{pmatrix} d_{11}^2 & & & \\ & d_{22}^2 & & \\ & & d_{33}^2 & \\ & & & \ddots \end{pmatrix}$

eigenvectors are colums of V

the largest eigenvalue is ~~the~~ d_{11}^2 and from e we know that w is an eigenvector => we need eigenvector corresponding to the largest eigenvalue => optimal w is the first column of V => $w = v_1$