HA-3 Theory Mikhaylov Igor

1 a). $\hat{\beta}_1 + \hat{\beta}_2 = \alpha \Rightarrow \hat{y} = \alpha x$

$Loss(\alpha) = (y - \alpha x)^T (y - \alpha x) + \lambda(\hat{\beta}_1^2 + \hat{\beta}_2^2)$

Because $x = x$ $\hat{\beta}_1$ must be equal to $\hat{\beta}_2$ in optimal variant

$\hat{\beta}_1 = \hat{\beta}_2 = \frac{\alpha}{2}$

$Loss(\alpha) = (y - \alpha x)^T (y - \alpha x) + \frac{\lambda \alpha^2}{2}$

$\Downarrow$

$\alpha = \frac{x^T y}{x^T x + \lambda} = \frac{\hat{\beta}_1 \cdot 2}{1} = \frac{\hat{\beta}_2 \cdot 2}{1}$

b). $\lambda \to \infty \Rightarrow \lambda(\beta_1^2 + \beta_2^2)$ dominates, so

$\hat{\beta}_1, \hat{\beta}_2 \to 0 \Rightarrow \hat{y} \to 0$

c). $\lambda \to 0 \Rightarrow$ no penalty $\Rightarrow \beta_1 + \beta_2 = \alpha$

2. a).

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

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

$E(\hat{\beta}) = \beta$

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

$Var(\hat{\beta}|X) = \sigma^2 (X^T X)^{-1} X^T W X (X^T X)^{-1}$

c). It wouldn't be valid, because there are no homoscedasticity ($I \neq W$)

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

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

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

$X^T X = \begin{pmatrix} 6 & 3 \\ 3 & 6 \end{pmatrix}$

$(6-\lambda)(6-\lambda)-9=0$

$\lambda_1 = 9 \qquad \lambda_2 = 3$

$\Downarrow \qquad \Downarrow$

$\begin{pmatrix} -3 & 3 \\ 3 & -3 \end{pmatrix} v = 0 \qquad \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix} v = 0$

$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$X^T X = V \Lambda V^T$

$V = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \Lambda = \begin{pmatrix} 9 & 0 \\ 0 & 3 \end{pmatrix}$

b). $\sigma_1 = \sqrt{\lambda_1} = 3$
$\sigma_2 = \sqrt{\lambda_2} = \sqrt{3}$ $\Rightarrow \Sigma = \begin{pmatrix} 3 & 0 \\ 0 & \sqrt{3} \\ 0 & 0 \end{pmatrix}$

$X = U \Sigma V^T \quad U = X V \Sigma^{-1}$

$\cancel{\Sigma^{-1} = \begin{pmatrix} 1/\sqrt{3} \end{pmatrix}}$ $\quad \Sigma^{-1} = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/\sqrt{3} & 0 \\ 0 & 0 & 0 \end{pmatrix}$

$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & \frac{1}{\sqrt{3}} \\ \frac{1}{3} & \frac{-3}{\sqrt{3}} \\ \frac{2}{3} & 0 \end{pmatrix} \quad XV = \frac{1}{\sqrt{2}} \begin{pmatrix} 3 & 1 \\ 1 & -3 \\ 2 & 0 \end{pmatrix}$

$V = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

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

4. a). $\|Xw\|^2 = w^T X^T X w \to \max$

s.t.

$\|w\|^2 = 1$

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

b). $\frac{\partial L}{\partial w} = 2X^T X w - 2\lambda w = 0 \Rightarrow (X^T X - \lambda I)w = 0$