N1

X на X

$$loss(\hat{\beta}) = (y-\hat{y})^T (y-\hat{y}) + \lambda \hat{\beta}^T \hat{\beta} \to min$$

$$\hat{y} = \hat{\beta}_1 x + \hat{\beta}_2 x = x(\hat{\beta}_1 + \hat{\beta}_2) = x\hat{\beta}$$

a) let minimise the loss function

$$Loss(\hat{\beta}) = (y-\hat{y})^T (y-\hat{y}) + \lambda \hat{\beta}^T \hat{\beta} \to min$$

FOC:

$$d(loss(\hat{\beta})) = d((y-\hat{y})^T(y-\hat{y})) + \lambda d(\hat{\beta}^T \hat{\beta}) =$$

$$= 2(y-\hat{y})^T d(y-\hat{y}) + \lambda 2 d(\hat{\beta}^*) =$$

~~$= 2(y-\hat{y})^T d(y - x\hat{\beta}) + \lambda 2 d(\hat{\beta}^*) =$~~

$$= 2(y-\hat{y})^T d(-\hat{y}) + \lambda 2 d(\hat{\beta}^*) =$$

$$= -2(y-\hat{y})^T \underset{=0}{d}(x\hat{\beta}) + \lambda 2 d(\hat{\beta}^*) =$$

$$= -2(y-\hat{y})^T \hat{\beta} \boxed{dx} - 2(y-\hat{y})^T x \, d\hat{\beta} - \lambda 2 \cdot d\hat{\beta} =$$

$$= \cancel{d} - 2 d\hat{\beta} ((y - x\hat{\beta})^T x + \lambda) \quad =$$

c/ In order to find the best approximation to X with rank = K = 1, we need to sort the values on the diagonal of D in the descending order and drop n-K = 2-1 = 1 the lowest:

$$X^TX \cancel{} = \frac{1}{-2\sqrt2}\begin{pmatrix}1 & 1\\ 1+\sqrt2 & 1-\sqrt2\end{pmatrix}\begin{pmatrix}5+\sqrt2 & 0\\ 0 & 5-\sqrt2\end{pmatrix}\begin{pmatrix}1-\sqrt2 & -1\\ -1-\sqrt2 & 1\end{pmatrix} \approx$$

$$\approx -\frac{1}{-2\sqrt2}\begin{pmatrix}1\\ 1+\sqrt2\end{pmatrix}\begin{pmatrix}5+\sqrt2\end{pmatrix}\begin{pmatrix}1-\sqrt2 & -1\end{pmatrix} =$$

$$= -\frac{1}{2\sqrt2}\begin{pmatrix}5+\sqrt2\\ 7+6\sqrt2\end{pmatrix}\begin{pmatrix}1-\sqrt2 & -1\end{pmatrix} =$$

$$= -\frac{1}{2\sqrt2}\begin{pmatrix}3-4\sqrt2 & -5-\sqrt2\\ -5-\sqrt2 & -7-6\sqrt2\end{pmatrix} = -\frac{(5+\sqrt2)}{2\sqrt2}\begin{pmatrix}\overset{23-23\sqrt2}{\cancel{7-6\sqrt2}} & \cancel{-8}\,-23\\ -23 & -23-23\sqrt2\end{pmatrix}$$

$$\Rightarrow -\frac{5-\sqrt2}{2\sqrt2}\begin{pmatrix}23-23\sqrt2 & -23\\ -23(1-\sqrt2) & -23(1+\sqrt2)(1-\sqrt2)\end{pmatrix} \to$$

$$\to -\frac{5-\sqrt2}{2\sqrt2}\begin{pmatrix}23-23\sqrt2 & -23\\ 0 & 0\end{pmatrix}$$

Other ways

rank = 1

Hence

~~FA~~ for <u>FOC:</u>

$$d\,\text{loss}(\hat{\beta}) = -2\,d\hat{\beta}\,\underline{\left(\left(yx - x^2\hat{\beta}\right)^T + \lambda\right)} = 0$$

~~let solve a quadratic equation:~~

~~$x^2\hat{\beta} - yx - \lambda = 0$~~

~~$\hat{D} = y^2 + 4\lambda\hat{\beta}$~~

~~$x_1 = \frac{y + \sqrt{y^2 + 4\lambda\hat{\beta}}}{2\hat{\beta}}$~~ ~~$\Rightarrow (x - y)^2 = \frac{y^2}{4\hat{\beta}^2} + \frac{2\lambda}{\hat{\beta}}$~~

~~$x_2 = \frac{y - \sqrt{y^2 - 4\lambda\hat{\beta}}}{2\hat{\beta}}$~~

$$(y - X\hat{\beta})^T X + \lambda = 0$$

$$\left((y - X\hat{\beta})^T X\right)^T = (-\lambda)^T$$

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

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

~~$X^T y$~~ $X^T X\hat{\beta} = X^T y + \lambda^T$

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

Hence, we can say that

$$\begin{cases} \beta_1 = (X^T X)^{-1} X^T y + (X^T X)^{-1} \cancel{R} \lambda - \beta_2 \\ \beta_2 \in R \end{cases}$$

b/ if $\lambda \to \infty$:

$$\lim_{\lambda \to \infty} \hat{\beta} = \lim_{\lambda \to \infty} \Big(\overset{\to \text{ constant}}{(X^T X)^{-1} X^T y} \Big) + \lim_{\lambda \to \infty} \underset{\to \infty}{(X^T X)^{-1} \lambda} =$$

~~[illegible]~~ $\to \infty$

c/ $$\lim_{\lambda \to 0} \beta = \lim_{\lambda \to 0} \Big((X^T X)^{-1} X^T y \Big) + \lim_{\lambda \to 0} \underset{\to 0}{\lambda (X^T X)^{-1}}$$

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

$N\ 2$

$y = X\beta + U$ $\quad$ β is non-random

$E(U|X) = 0$

$X_{n\times k}$ $\quad \eta\ \overset{k}{-}$ $\quad$ $\mathrm{rank}(X) = K$

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

$\hat{\beta} = \hat{\beta}^{ols}$

a)

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

$$= (X^T X)^{-1} X^T E(X\beta + U|X) = (X^T X)^{-1} X^T (X + 0) =$$

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

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

b/

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

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

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

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

c/ We can say that we have a case of possible heteroscedasticity here (as we do not know the values of W). Thus, the standard CI for β will be probably ~~not~~ not valid there.

d)

$\mathrm{Cov}(y, \hat{\beta} | X) =$ ~~$\mathrm{Cov}(y, v | X)$~~ $=$

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

$= \mathrm{Cov}(y, y | X) \left((X^T X)^{-1} X^T\right)^T =$

$= \mathrm{Var}(y | X) \left((X^T X)^{-1} X^T\right)^T =$

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

DSB

n3

b/

$$A^TA = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & 2 \\ 1 & 1 \end{pmatrix} --$$

$$= \begin{pmatrix} 6 & 1 \\ 1 & 6 \end{pmatrix}$$

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

$$\begin{vmatrix} 6-\lambda & 1 \\ 1 & 6-\lambda \end{vmatrix} = (6-\lambda)(6-\lambda) - 1 = \cancel{\lambda^2 - 10\lambda + 23}$$

$$D = 2^2 \cancel{100 - 4 \cdot 23} = (2\sqrt{}$$

$$\lambda_1 = \cancel{5+\sqrt{2}} \; \frac{12+2}{2} = 7$$

$$\lambda_2 = \cancel{5-\sqrt{2}} \; \frac{12-2}{2} = 5$$

~~X~~ Hence

$$D = \begin{pmatrix} \overset{7}{\cancel{5+\sqrt{2}}} & 0 \\ 0 & \overset{5}{\cancel{5-\sqrt{2}}} \end{pmatrix}$$

Let find corresponding eigen vectors

$$\lambda_1 = \overset{7}{\cancel{5+\sqrt{2}}} \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \to \begin{pmatrix} -1 & 1 \end{pmatrix} \Rightarrow v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$\lambda_2 = 5 \Rightarrow \begin{pmatrix}1&1\\1&1\end{pmatrix} \to \begin{pmatrix}1&1\end{pmatrix} \Rightarrow v_2 = \begin{pmatrix}1\\-1\end{pmatrix}$

Hence

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

$$P^{-1} = \frac{1}{-2}\begin{pmatrix}-1&-1\\-1&1\end{pmatrix}$$

b) Hence

$$X^TX = P\,D\,P^{-1} = \frac{1}{-2}\begin{pmatrix}1&1\\1&-1\end{pmatrix}\begin{pmatrix}7&0\\0&5\end{pmatrix}\begin{pmatrix}-1&-1\\-1&1\end{pmatrix}$$

~~We can say that~~

We can say that this is a representation of X^TX in SVD form

$(P = U,\ D = \Sigma,\ P^{-1} = V^T)$

c/ Here, as well as we were to find the best approximation to X with rank equal to 1; we will take the highest ~~value~~ ~~from the~~ diagonal value from the diagonal matrix and will construct an approximation using it!

$$X \approx SVD_{rank=1} = -\frac{1}{2}\begin{pmatrix} \cancel{\text{[illegible]}} & 1 \end{pmatrix}\begin{pmatrix} 7 \end{pmatrix}\begin{pmatrix} -1 & -1 \\ \cancel{\text{[illegible]}} \end{pmatrix} =$$

$$= \cancel{-\frac{7}{2}} - \frac{7}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\begin{pmatrix} -1 & -1 \end{pmatrix} =$$

$$= \frac{7}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 7/2 & 7/2 \\ 7/2 & 7/2 \end{pmatrix}$$

SVD: №4

$$X = UDV^T \qquad D = \begin{pmatrix} d_{11} & & 0 \\ \cancel{d} & d_{12} & \\ 0 & & \ddots \; 0 \end{pmatrix}$$

$$\begin{cases} \|Xw\|^2 \to max \\ s.t. \\ \|w\|^2 = 1 \end{cases}$$

a) $$\begin{cases} (Xw)^T(Xw) \to max \\ s.t. \\ w^Tw = 1 \end{cases}$$

$$L = (Xw)^T(Xw) - \lambda(w^Tw \cancel{-1}) \to max$$

b) **FOC:**

$$dL = d((Xw)^T(Xw)) - \lambda\, d(w^Tw) =$$

$$= 2(Xw)^T d(Xw) - 2\lambda\, d(w) =$$

$$= 2(Xw)^T(w\overset{=0}{dX} + X\,dw) - 2\lambda d(w) =$$

$$= 2((Xw)^T - 2\lambda X)dw = 0$$

c/ ~~$w^T = \ldots 2\lambda$~~

$(X w^T = 2\lambda X$

$X^T w^T = 2\lambda X$

$w^T = X^{T-1} 2\lambda X = (UDU^T)^{-1} \ldots 2\lambda X$

$= 2\lambda (UDU^T)^{-1} UDV^T$ ~~$= 2\lambda$~~

here

$$w^T = 2\lambda (UDU^T)^{-1} UDV^T$$

$$w = 2\lambda VDU^T (UDU^T)^{-1}$$