Subject : No. : Date : / /

Problem 2.1

$$\theta_{MSE} = \underset{\theta}{\text{argmin}} \ (y - X\theta)^T (y - X\theta)$$

$$= y^T y - 2\theta^T X^T y + \theta^T X^T X \theta$$

$$\xrightarrow{\frac{\partial}{\partial\theta}} -2X^T y + 2X^T X \theta$$

let $-2X^T y + 2X^T X\theta = 0$

$$\underline{\theta = (X^T X)^{-1} X^T y = \theta_{MSE}} \ \#$$

Problem 2.2

d is the observation number in y

$$\theta_{MLE} = \underset{\theta}{\text{argmax}} \ \frac{1}{\sqrt{(2\pi)^d |\Sigma^*|}} \times e^{-\frac{1}{2}[(y - X\theta)^T \Sigma^{*-1} (y - X\theta)]}$$

$$\xrightarrow{\text{take log}} \underset{\theta}{\text{argmax}} \ \log\left(\frac{1}{\sqrt{(2\pi)^d |\Sigma^*|}}\right) - \frac{1}{2}\left[(y - X\theta)^T \Sigma^{*-1} (y - X\theta)\right]$$

$$= \underset{\theta}{\text{argmin}} \ -\log\left(\frac{1}{\sqrt{(2\pi)^d |\Sigma|}}\right) + \frac{1}{2}\left[(y - X\theta)^T \Sigma^{*-1} (y - X\theta)\right]$$

$$= \underset{\theta}{\text{argmin}} \ (y - X\theta)^T \Sigma^{*-1} (y - X\theta)$$

$$\frac{\partial}{\partial\theta} = y^T \Sigma^{-1} y - 2\theta^T X^T \Sigma^{*-1} y + \theta^T X^T \Sigma^{*-1} X \theta$$

$$= -2X^T \Sigma^{*-1} y + 2X^T \Sigma^{*-1} X\theta$$

let $-2X^T \Sigma^{*-1} y + 2X^T \Sigma^{*-1} X\theta = 0$

$$\underline{\theta_{MLE} = (X^T \Sigma^{*-1} X)^{-1} (X^T \Sigma^{*-1} y)} \ \#$$

Double A

Subject : ....

No. :
Date : ....../....../......

Problem 2.3

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

$y = X\theta^* + \varepsilon \quad , \quad \varepsilon \sim N(0, \Sigma^*)$

$y \sim N(X\theta^*, \Sigma)$

$\hat{\theta} \sim N((X^T X)^{-1} X^T X \theta^*, \ (X^T X)^{-1} X^T \Sigma^* [(X^T X)^{-1} X^T]^T)$

Problem . 2.4

By Invariance of MLE

$\hat{y} = X^T_{new} \hat{\theta}$

$= X^T_{new} (X^T X)^{-1} X^T y$

Problem 2.5.

Since $\hat{y} = X^T_{new} \hat{\theta}$

$\hat{y} \sim N(X^T_{new} \theta^*, \ X^T_{new} (X^T X)^{-1} X^T \Sigma^* [(X^T X)^{-1} X^T]^T X_{new})$

Subject :

No. :

Date :

Problem 2.6

$$\underset{\Sigma}{\text{argmax}} \frac{1}{\sqrt{(2\pi)^d |\Sigma|}} \times e^{-\frac{1}{2}(y-X\theta^*)^T \Sigma^{-1}(y-X\theta^*)}$$

$$\xrightarrow{\log} \underset{\Sigma}{\text{argmax}} -\frac{d}{2}\log 2\pi - \frac{1}{2}|\Sigma| - \frac{1}{2}(y-X\theta^*)^T \Sigma^{-1}(y-X\theta^*)$$

$$\propto -\frac{d}{2}\log 2\pi - \frac{1}{2}|\Sigma| - \frac{1}{2}\,\mathrm{tr}\left(\Sigma^{-1}(y-X\theta^*)(y-X\theta^*)^T\right)$$

$$\frac{\partial}{\partial \Sigma^{-1}} \Rightarrow \frac{1}{2}\Sigma - \frac{1}{2}(y-X\theta^*)(y-X\theta^*)^T = 0$$

$$\Sigma = (y-X\theta^*)(y-X\theta^*)^T$$

2.7.

In this question, we have to assume homoskedasticity, otherwise it's not solvable.

(a.)

$$\Sigma_{mle} = \sigma^2 I = 457.32 \times I \#$$

$$\hat{\theta} = (X^TX)^{-1}X^Ty = \begin{array}{cl} \begin{bmatrix} 0.1159 \\ 1.0605 \\ -50.4572 \end{bmatrix} & \begin{array}{l} \text{Height} \\ \text{Weight} \\ \text{Intercept} \end{array} \end{array}$$

(b)

$$\hat{y} = [175, 170, 1] \cdot \hat{\theta}$$

$$\doteqdot 150.1319$$

(c)

Variance for $\hat{y}$ $= \sigma^2 x_{new}^T (X^TX)^{-1} x_{new} = 133.54$

z value when $\alpha = 0.05 \Rightarrow 1.96$

95% Confidence interval for y $= [150.13 - \sqrt{133.54} \times 1.96,$
$150.13 + \sqrt{133.54} \times 1.96]$

$$\doteqdot [127.48, 172.78]$$

Subject : .... No. : Date : ..../......./.......

(d) Under 95% confidence level

$cov(\hat{\theta}) = \sigma^2 (X^T X)^{-1}$

$$\frac{\hat{\theta}^2_{height}}{V^2_{height}} = \frac{(0.1159)^2}{0.6725} \doteqdot 0.01999$$

Cumulated density $(0.01999) \doteqdot 11.2\% < 95\%$. Height is not significant #

(e) Under 95% confidence level

$$\frac{\hat{\theta}^2_{Height}}{V^2_{weight}} = \frac{(1.0605)^2}{0.2479} = 4.5372$$

Cumulated density of $4.5372 \doteqdot 96.6\% > 95\%$.

Weight is significant. #