Άσκηση 1

(α)
$$\sum_{n=1}^{N}\left(y_n-\theta^T x_n\right)^2+2\lambda\theta=0$$

1ος όρος

$$\frac{\partial\sum\left(y_n-\theta^T x_n\right)^2}{\partial\theta}=$$

$$=\frac{\partial\sum\left(y_n^2-2y_n\theta^T x_n+(\theta^T x_n)^2\right)}{\partial\theta}$$

$$=\frac{\partial\sum\cancel{(y_n^2)}}{\partial\theta}\xrightarrow{}0-\frac{\partial\,2\sum y_n\theta^T x_n}{\partial\theta}+\frac{\partial\sum(\theta^T\cdot x_n)^2}{\partial\theta}$$

$$=-2\sum y_n x_n\cdot\frac{\partial\theta^T}{\partial\theta}+\frac{\partial\sum\theta^T\cdot\overbrace{\theta^T\cdot x_n\cdot x_n}^{x_n\cdot x_n^T\theta}}{\partial\theta}$$

$$=-2\sum y_n x_n+\frac{\partial\,\theta^T\cdot\theta\sum x_n\cdot x_n^T}{\partial\theta}$$

$$=-2\sum_{n=1}^{N}y_n\cdot x_n+\overbrace{\frac{\partial\theta^T\cdot\theta}{\partial\theta}}^{2\theta}\circ\sum_{i=1}^{N}x_n\cdot x_n^T$$

$$=-2\sum_{n=1}^{N}y_n\cdot x_n+2\theta\sum_{i=1}^{N}x_n\cdot x_n^T$$

Άρα

$$-\cancel{2}\sum_{n=1}^{N} y_n \cdot x_n + \cancel{2}\theta \sum_{n=1}^{N} x_n \cdot x_n^T + \cancel{2}\lambda\theta = 0 \iff$$

$$\theta \sum_{n=1}^{N} x_n \cdot x_n^T + \lambda\theta = \sum_{n=1}^{N} y_n \cdot x_n \iff$$

$$\underbrace{\sum_{n=1}^{N} x_n \cdot x_n^T}_{A} \cdot \underbrace{\theta}_{z} + \lambda \underbrace{\theta}_{z} = \sum_{n=1}^{N} y_n \cdot x_n \iff$$

$$\left(\sum_{n=1}^{N} x_n \cdot x^T + \lambda I \cdot\right)\theta = \sum_{n=1}^{N} y_n \cdot x_n \iff$$

$$\hat{\theta} = \left(\sum_{n=1}^{N} x_n \cdot x_n^T + \lambda I\right)^{-1} \cdot \sum_{n=1}^{N} y_n \cdot x_n \quad (1)$$

(b) $$\sum_{n=1}^{N} x_n \cdot x_n^T = [x_1, x_2, \ldots, x_n] \times \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = X^T \cdot X$$

Homework 1 Exercise 3

$$\sum_{n=1}^{N} y_n \cdot x_n = [x_1, x_2, \ldots, x_n] \times \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = X^T \cdot y$$

Συνεπώς $$\theta = (X^T \cdot X + \lambda I)^{-1} \cdot X^T \cdot y$$

Αντικατάσταση στο (1) από ερώτημα (a)

Άσκηση 2

(a) $MSE(\hat{\theta}) = \underbrace{E\left[(\hat{\theta} - E(\hat{\theta}))^2\right]}_{\text{Variance}} + \underbrace{(E(\hat{\theta}) - \theta_0)^2}_{\text{bias}^2}$

Επειδή είναι unbiased τότε $(E(\hat{\theta}) - \theta_0)^2 = 0$

Άρα $MSE(\hat{\theta}) = E\left[(\hat{\theta} - E[\hat{\theta}])^2\right]$

(b)

$\hat{\theta}_b = (1+a)\hat{\theta}_{MVU} \overset{E[\,]}{\Leftrightarrow}$

$E[\hat{\theta}_b] = E[(1+a)\hat{\theta}_{MVU}] \Leftrightarrow$

$E[\hat{\theta}_b] = (1+a) \cdot E[\overset{\to \theta_0}{\cancel{\hat{\theta}_{MVU}}}] \Leftrightarrow$

$E[\hat{\theta}_b] = (1+a) \cdot \theta_0$ Για $a \neq 0$ $E[\hat{\theta}_b] \neq \theta_0$
άρα $\hat{\theta}_b$ biased estimator.

(c)

$MSE(\hat{\theta}) = E\left[(\hat{\theta} - E(\hat{\theta}))^2\right] + (E(\hat{\theta}) - \theta_0)^2$

Άρα $MSE(\hat{\theta}_{MVU}) = E\left[(\hat{\theta}_{MVU} - \overset{\to \theta_0}{E[\hat{\theta}_{MVU}]})^2\right] + (\overset{\to \theta_0}{E[\hat{\theta}_{MVU}]} - \theta_0)^2 \Leftrightarrow$

$MSE(\hat{\theta}_{MVU}) = \underbrace{E\left[(\hat{\theta}_{MVU} - \theta_0)^2\right]}_{\text{Variance}} + \cancel{(\theta_0 - \theta_0)^2}$

$= \underbrace{Var(\hat{\theta}_{MVU} - \theta_0)} + E^2\left[(\theta_{MVU} - \theta_0)\right]$

$\to$ για συγκεκριμένο N θα είναι διάφορο του 0.

(d) $MSE(\hat{\theta}_b) = E\left[\left(\hat{\theta}_b - E(\hat{\theta}_b)\right)^2\right] + \left(E(\hat{\theta}_b) - \theta_0\right)^2$

$$\xlongequal{\hat{\theta}_b = (1+a)\hat{\theta}_{MVU}} E\left[\left((1+a)\hat{\theta}_{MVU} - E\left[(1+a)\hat{\theta}_{MVU}\right]\right)^2\right] + \left(E\left[(1+a)\cdot\hat{\theta}_{MVU}\right] - \theta_0\right)^2$$

$$= E\left[\left((1+a)\cdot\hat{\theta}_{MVU} - (1+a)\theta_0\right)^2\right] + \left((1+a)\theta_0 - \theta_0\right)^2$$

$$= (1+a)^2 \cdot E\left[\left(\hat{\theta}_{MVU} - \theta_0\right)^2\right] + \left(\theta_0 \cdot (1+a-1)\right)^2$$

$$= (1+a)^2 \cdot \underbrace{E\left[\left(\hat{\theta}_{MVU} - \theta_0\right)^2\right]}_{MSE(\hat{\theta}) = E[(\hat{\theta}-\theta_0)^2]} + (\theta_0 \cdot a)^2$$

$$= (1+a)^2 \cdot MSE(\hat{\theta}_{MVU}) + \theta_0^2 \cdot a^2$$

(e) $MSE_{\hat{\theta}_\delta} < MSE_{\hat{\theta}_{MVU}} \overset{(d)}{\Leftrightarrow}$

$$(1+a)^2 \cdot MSE(\hat{\theta}_{MVU}) + \theta_0^2 \cdot a^2 < MSE(\hat{\theta}_{MVU}) \Leftrightarrow$$

$$(1+a)^2 \cdot MSE(\hat{\theta}_{MVU}) - MSE(\hat{\theta}_{MVU}) + \theta_0^2 \cdot a^2 < 0 \Leftrightarrow$$

$$MSE(\hat{\theta}_{MVU}) \cdot \left((1+a)^2 - 1\right) + \theta_0^2 \cdot a^2 < 0 \Leftrightarrow$$

$$MSE(\hat{\theta}_{MVU}) \cdot \left(\cancel{1} + 2a + a^2 - \cancel{1}\right) + \theta_0^2 \cdot a^2 < 0 \Leftrightarrow$$

$$a^2 \cdot MSE(\hat{\theta}_{MVU}) + 2a\,MSE(\hat{\theta}_{MVU}) + \theta_0^2 \cdot a^2 < 0 \Leftrightarrow$$

$$a^2\left(MSE(\hat{\theta}_{MVU}) + \theta_0^2\right) + a \cdot 2MSE(\hat{\theta}_{MVU}) < 0$$

$$\Delta = \left(2MSE(\hat{\theta}_{MVU})\right)^2 = 4\left(MSE(\hat{\theta}_{MVU})\right)^2$$

$$p_{1,2} = \frac{-2MSE(\hat{\theta}_{MVU}) \pm \sqrt{4MSE^2(\hat{\theta}_{\theta MV})}}{2\left(MSE(\hat{\theta}_{MVU}) + \theta_0^2\right)} = \begin{cases} 0 \\ \dfrac{-4MSE(\hat{\theta}_{MVU})}{2(MSE(\hat{\theta}_{MVU} + \theta_0^2)} \end{cases}$$

Συνεπώς $\quad -\dfrac{2MSE(\hat{\theta}_{MVU})}{MSE(\hat{\theta}_{MVU}) + \theta_0^2} < a < 0$

(f)

$$|1+a|<1 \overset{\times|\hat{\theta}_{MVU}|}{\Longleftrightarrow}$$

$$|(1+a)\cdot\hat{\theta}_{MVU}|<|\hat{\theta}_{MVU}| \overset{(1)}{\Longleftrightarrow}$$

$$|\hat{\theta}_b|<|\hat{\theta}_{MVU}|$$

(g) $MSE(\hat{\theta}_b)=(1+a)^2 MSE(\hat{\theta}_{MVU})+a^2\theta_0^2$

$$\frac{\partial MSE(\hat{\theta}_b)}{\partial a}=2(1+a)MSE(\hat{\theta}_{MVU})+2a\theta_0^2$$

$$=2MSE(\hat{\theta}_{MVU})+2aMSE(\hat{\theta}_{MVU})+2a\theta_0^2$$

$$=2a\left(MSE(\hat{\theta}_{MVU})+\theta_0^2\right)+2MSE(\hat{\theta}_{MVU})$$

$$\frac{\partial MSE(\hat{\theta}_b)}{\partial a}=0 \Longleftrightarrow$$

$$2a\left(MSE(\hat{\theta}_{MVU})+\theta_0^2\right)+2MSE(\hat{\theta}_{MVU})=0 \Longleftrightarrow$$

$$a^*=\frac{-\cancel{2}MSE(\hat{\theta}_{MVU})}{\cancel{2}(MSE(\hat{\theta}_{MVU})+\theta_0^2)}=\frac{-MSE(\hat{\theta}_{MVU})}{MSE(\hat{\theta}_{MVU})+\theta_0^2}$$

(h) —

Exercise 3

(a) $\left(\sum_{n=1}^{N} x_n \cdot x_n^T\right) \cdot \theta_0 = \sum_{n=1}^{N} y_n \cdot x_n \overset{\gamma\iota\alpha\ x=[1]}{\Longleftrightarrow}$

$\sum_{n=1}^{N} [1][1] \cdot \theta_0 = \sum_{n=1}^{N} y_n \Longleftrightarrow$

$N \cdot \theta_0 = \sum_{n=1}^{N} y_n \Longleftrightarrow \theta_0 = \frac{\sum_{n=1}^{N} y_n}{N} = \bar{y}$

(b) $E[y_n] \overset{(3)}{=} E[\theta_0 + \eta_n] = E[\theta_0] + \cancel{E[\eta_n]}^{0} \quad \eta \sim N(\mu=0)$ $=$

$E[\theta_0] = \theta_0$

(c)

$E[\bar{y}] = E\left[\frac{1}{N} \cdot \sum_{i=1}^{N} y_n\right] = \frac{1}{N} \cdot \sum_{i=1}^{N} \cdot E[y_n]$

$= \frac{1}{\cancel{N}} \cdot \cancel{N} \cdot E[y_n] = E[y_n] \overset{(b)}{=} \theta_0$

(e)

$\hat{\theta} = \left(\sum_{i=1}^{N} x_n \cdot x_n^T + \lambda I\right)^{-1} \cdot \sum_{i=1}^{N} y_n \cdot x_n$

$= \left(\sum_{i=1}^{N} [1] \cdot [1] + \lambda \cdot [1]\right)^{-1} \cdot \sum_{i=1}^{N} y_n \cdot [1] = (N+\lambda)^{-1} \cdot \sum_{i=1}^{N} y_n$

$= \frac{\sum_{i=1}^{N} y_n}{(N+\lambda)}$

(f)

$E[\hat{\theta}] = E\left[\frac{\sum_{i=1}^{N} y_n}{(N+\lambda)}\right] = \frac{1}{N+\lambda} \cdot \sum_{i=1}^{N} E[y_n] = \frac{N}{N+\lambda} \cdot \theta_0 \neq \theta_0$

Άρα $\hat{\theta}$ είναι biased, Συνεπώς $\hat{\theta} = (1+\alpha) \cdot \hat{\theta}_{MVU}$

(g) $E[\hat{b}] = E\left[\frac{\sum_{i=1}^{N} y_n}{N+2}\right] = \frac{1}{N+2} \cdot \sum_{i=1}^{N} E[y_n] = \frac{N}{N+2} \cdot E[y_n] \overset{(b)}{=}$

$= \frac{N \cdot \theta_0}{N+2} \neq 0$ biased

(h) Απο (f) $\hat{\theta} = (1+a)\hat{\theta}_{MVU} \Leftrightarrow$

$|\hat{\theta}| = |(1+a)\cdot\hat{\theta}_{MVU}| \Leftrightarrow |\hat{\theta}| = |1+a|\cdot|\hat{\theta}_{MVU}|$ (1)

Από άσκηση 2.(f) $|1+a| < 1$ (2)

Συνεπώς απο (1),(2) φένεται $|\hat{\theta}| < |\hat{\theta}_{MVU}|$