Assignment-1

1.) MMSE estimator of θ is a function of $u = g(x)$ that minimizes the MSE $E((\theta - a)^2)$

To prove, $a = E[\theta]$.

Given,

$$E(\theta - a)^2 = \int (\theta - a)^2 p(\theta)\, d\theta$$

To minimize, Differentiating with respect to a. and equate to 0.

$$-2 \int (\theta - a)\, p(\theta)\, d\theta = 0$$

$$\int \theta p(\theta) d\theta - \int a\, p(\theta)\, d\theta = 0.$$

$$a \int p(\theta)\, d\theta = \int \theta\, p(\theta)\, d\theta.$$

Now $\int p(\theta)\, d\theta = 1$. Because θ is a Random Variable & its' pdf $= 1$

and $\int \theta\, p(\theta)\, d\theta = E[\theta] \Rightarrow$ mean

$$a = E[\theta].$$

$\therefore$ Expection of MSE mimizes to mean of distribution

2.) Given Linear estimator.

a.) $y_k = \theta_1 x_k + \theta_2 \quad k = 1, 2 \dots T$

$y_1 = \theta_1 x_1 + \theta_2$

$y_2 = \theta_1 x_2 + \theta_2$

$y_3 = \theta_1 x_3 + \theta_2$

$\vdots$

$y_T = \theta_1 x_T + \theta_2$

$$\Longrightarrow \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_T \end{bmatrix}}_{Y} = \underbrace{\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ \vdots & \\ x_T & 1 \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}}_{\theta}.$$

$\boxed{Y = X\theta}$ where

$$X = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ \vdots & \vdots \\ x_T & 1 \end{bmatrix}$$

b.) Now

Error function $= Y - X\theta$

The mean Squared Error $= \sum_{k=1}^{T} (y_k - \theta_1 x_k - \theta_2)^2.$

$= (Y - X\theta)^2.$

$= (Y - X\theta)^T (Y - X\theta)$

$= (Y^T - \theta^T X^T)(Y - X\theta)$

$= Y^T Y - Y^T X\theta - \theta^T X^T Y + \theta^T X^T X\theta.$

Now,

$$y^T x\theta = \theta^T x^T y = \text{Scalar value}$$

$$\text{mse } E(\theta) = y^T y - 2\theta^T x^T y + \theta^T x^T x\,\theta. \quad (2)$$

c.) The gradient of error is given by

$\nabla \text{MSE}(E(\theta)) \Rightarrow$ Differentiate with respect to θ in above equation

$$= \underbrace{\frac{d}{d\theta}(y^T y)}_{0} - \frac{d}{d\theta}(2\theta^T x^T y) + \frac{d}{d\theta}(\theta^T x^T x\theta)$$

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

$$= -2x^T y + 2x^T x\,\theta.$$

Now to find optimum $\underline{\underline{\theta}}$, we set gradient to zero.

$$\nabla \text{MSE}, E(\theta) = 0$$

$$-2x^T y + 2x^T x\theta = 0$$

$$x^T x\theta = x^T y.$$

$$\boxed{\theta = (x^T x)^{-1} x^T y.}$$

This is generally called Closed form solution in general Linear Estimator Journals