8. MMSE Estimation

Previously we have seen the Maximum Likelihood (ML) Estimation. Now, let us see another methodology for Estimation, which is Minimum Mean Square Estimation (MMSE) / MMSE philosophy.

- To understand MMSE Estimation, let us consider an observation vector $\bar{y} = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix}$ and an unknown parameter vector $\bar{h} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_M \end{bmatrix}$. However, in this, $\bar{y}, \bar{h}$ are random in nature. This estimation philosophy is known as Bayesian philosophy (unknown quantity is RANDOM).

The Fundamental difference between ML and MMSE is

ML	MMSE
⊙ $\bar{y}$ is Random	⊙ $\bar{y}$ is Random
⊙ $\bar{h}$ is Deterministic, yet unknown	⊙ $\bar{h}$ is Random

The cost-function for the MMSE estimate of $\bar{h}$ is given as

$$\underbrace{\min \underbrace{E\left\{\underbrace{\| \underbrace{\hat{h} - \bar{h}}_{\text{Error}} \|^2}_{\text{Squared Error}}\right\}}_{\text{Mean of Squared Error}}}_{\text{Minimum of the Mean of Squared Error.}}$$

This is why the name Minimum Mean Square Error (MMSE) principle.

> **Recall,** cost-function for the ML Estimate
>
> $$\min \underbrace{\| \bar{y} - X\bar{h} \|^2}_{\text{Least squares.}}$$

The General expression for the MMSE estimate for any observation vector $\bar{y}$ is given by

$$\boxed{\hat{h} = E\{\bar{h} \mid \bar{y}\}}$$

This looks simple, but challenging in Practice !! coz this is conditional expectation of $\bar{h}$ given $\bar{y}$. To find the conditional expectation, first we have to find the conditional PDF and then take the expected value of that conditional PDF. So, in general, it is fairly difficult to evaluate!

For Unknown parameter vector ($\bar{h}$), observation vector ($\bar{y}$) } Jointly Gaussian,

(i) $\bar{h}$, $\bar{y}$ both are zero-mean Gaussian Random vectors,

(ii) $E\{\bar{h}\} = 0$, $E\{\bar{y}\} = 0$,

the MMSE estimate is given by

$$\boxed{\hat{h} = E\{\bar{h} \mid \bar{y}\} = R_{hy} R_{yy}^{-1} \bar{y}}$$

This MMSE estimate of $\hat{h}$ is valid only when $\bar{h}, \bar{y}$ are jointly gaussian.

where,

- R_{yy} is the Covariance Matrix of $\bar{y}$

$$\Rightarrow R_{yy} = E\{\bar{y}\,\bar{y}^T\}$$

$$= E\left\{ \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix} \begin{bmatrix} y(1) & y(2) & \cdots & y(N) \end{bmatrix} \right\}_{N \times N}$$

- R_{hy} is the Cross Covariance Matrix of $\bar{h}, \bar{y}$

$$\Rightarrow R_{hy} = E\{\bar{h}\,\bar{y}^T\}$$

$$= E\left\{ \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_M \end{bmatrix} \begin{bmatrix} y(1) & y(2) & \cdots & y(N) \end{bmatrix} \right\}_{M \times N}$$

- $R_{(\hat{h}-\bar{h})(\hat{h}-\bar{h})}$ is the Error Covariance Matrix

$$\Rightarrow R_{(\hat{h}-\bar{h})(\hat{h}-\bar{h})} = E\left\{(\hat{h}-\bar{h})(\hat{h}-\bar{h})^T\right\}$$

$$= R_{hh} - R_{hy} R_{yy}^{-1} R_{yh}$$

-

Consider now the MISO Channel estimation model:

$$\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix}_{N\times 1} = \begin{bmatrix} \bar{x}^T(1) \\ \bar{x}^T(2) \\ \vdots \\ \bar{x}^T(N) \end{bmatrix}_{N\times M} \cdot \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_M \end{bmatrix}_{M\times 1} + \begin{bmatrix} v(1) \\ v(2) \\ \vdots \\ v(N) \end{bmatrix}_{N\times 1}$$

Observation vector

Pilot Matrix

$$\Rightarrow \bar{y} = X\bar{h} + \bar{v}$$

Assume $\bar{h}$ is gaussian, $\bar{v}$ is gaussian.

Since $X\bar{h} + \bar{v}$ is a linear combination of the random vectors $\bar{h}$, $\bar{v}$, $\bar{y}$ is also Gaussian.

Hence, this MISO channel estimation model is also known as Linear Gaussian Model / Linear Estimation model.

Now, for this linear estimation model, let us derive the MMSE Estimate ($\hat{h}$).

We have the unknown parameter vector $\bar{h} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_M \end{bmatrix}$.

Let the elements h_i be zero-mean iid gaussian of variance σ_h^2.

(i) $E\{h_i\} = 0$

$$E\{h_i^2\} = \sigma_h^2, \quad i = j$$

$$E\{h_i h_j\} = E\{h_i\} \cdot E\{h_j\} = 0, \quad i \neq j \quad (\because \text{iid})$$

The covariance Matrix of $\bar{h}$ is

$$R_{hh} = E\{\bar{h}\,\bar{h}^T\} = E\left\{ \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_M \end{bmatrix} \begin{bmatrix} h_1 & h_2 & \cdots & h_M \end{bmatrix} \right\}$$

h_i^2 - Diagonal Elements $(i=j)$

$E\{h_i^2\} = \sigma_h^2$

$$= E\left\{ \begin{bmatrix} h_1^2 & h_1 h_2 & & \\ h_2 h_1 & h_2^2 & & \\ \vdots & & \ddots & \\ & & & h_M^2 \end{bmatrix}_{M\times M} \right\}$$

$h_i h_j$ - Off Diagonal Elements

$E\{h_i h_j\} = 0$

$$= \begin{bmatrix} \sigma_h^2 & 0 & & 0 \\ 0 & \sigma_h^2 & & \\ \vdots & & \ddots & \\ 0 & & & \sigma_h^2 \end{bmatrix}_{M\times M}$$

$$= \sigma_h^2 I$$ M×M Matrix

Week 6 : Session 2

Similarly, let the elements $v(i)$ be zero-mean iid gaussian of variance σ^2.

The Noise vector $\bar{v} = \begin{bmatrix} v(1) \\ v(2) \\ \vdots \\ v(N) \end{bmatrix}$.

The covariance Matrix of $\bar{v}$ is

$$R_{vv} = E\{\bar{v}\,\bar{v}^T\} = \sigma^2 I$$ N×N Matrix

For the MISO channel estimation model $\bar{y} = X\bar{h} + \bar{v}$, the MMSE estimate $(\hat{h})$ is $\hat{h} = (R_{hy})(R_{yy}^{-1})\bar{y}$.

$M \times N$ Wide Matrix

$N \times N$ Matrix

$$\Rightarrow R_{yy} = E\{\bar{y}\bar{y}^T\}$$

$$= E\{(X\bar{h}+\bar{v})(X\bar{h}+\bar{v})^T\}$$

$$= E\{X\bar{h}\bar{h}^T X^T + \bar{v}\bar{h}^T X^T + X\bar{h}\bar{v}^T + \bar{v}\bar{v}^T\}$$

$\bar{v}$: Thermal Noise at Receiver

$\bar{h}$: Random channel because of Scattering environment

$\Rightarrow \bar{h}, \bar{v}$ are independent

$\Rightarrow E\{\bar{h}\bar{v}^T\} = E\{\bar{v}\bar{h}^T\} = 0$

$$= X E\{\bar{h}\bar{h}^T\}X^T + \cancelto{0}{E\{\bar{v}\bar{h}^T\}}X^T + X\cancelto{0}{E\{\bar{h}\bar{v}^T\}} + E\{\bar{v}\bar{v}^T\}$$

$$= X E\{\bar{h}\bar{h}^T\}X^T + E\{\bar{v}\bar{v}^T\}$$

$$= X \underline{\sigma_h^2 I} X^T + \underline{\sigma^2 I}$$

$$= \underbrace{\sigma_h^2 X X^T + \sigma^2 I}_{N \times N}$$

$$\Rightarrow R_{hy} = E\{\bar{h}\bar{y}^T\}$$

$$= E\{\bar{h}(X\bar{h}+\bar{v})^T\}$$

$$= E\{\bar{h}\bar{h}^T X^T + \bar{h}\bar{v}^T\}$$

$$= E\{\bar{h}\bar{h}^T\}X^T + \cancelto{0}{E\{\bar{h}\bar{v}^T\}}$$

$$= \sigma_h^2 X^T$$

Therefore, the MMSE Estimate is given as

$$\hat{h} = E\{\bar{h}|\bar{y}\} = R_{hy} R_{yy}^{-1}\bar{y}$$

$$= \sigma_h^2 \underline{X^T} \underbrace{(\sigma_h^2 X X^T + \sigma^2 I)}_{N \times N}^{-1} \bar{y}$$

Note:

$$\sigma_h^2 X^T X X^T + \sigma^2 X^T = \sigma_h^2 X^T X X^T + \sigma^2 X^T$$

$$\Rightarrow X^T(\sigma_h^2 X X^T + \sigma^2 I) = (\sigma_h^2 X^T X + \sigma^2 I) \cdot X^T$$

$$\Rightarrow (\sigma_h^2 X^T X + \sigma^2 I)^{-1} X^T = X^T (\sigma_h^2 X X^T + \sigma^2 I)^{-1}$$

Use the above result in MMSE estimate.

$$\Rightarrow \hat{h} = \sigma_h^2 \underbrace{(\sigma_h^2 X^T X + \sigma^2 I)^{-1}}_{M \times M} X^T \bar{y}$$

Typically, $M << N$.

So, the modified MMSE estimate is easier to evaluate than the original MMSE estimate.

On further simplification,

$$\hat{h} = \sigma_h^2 (\sigma_h^2 X^T X + \sigma^2 I)^{-1} X^T \bar{y}$$

$$= \left(X^T X + \frac{\sigma^2}{\sigma_h^2} I\right)^{-1} X^T \bar{y}.$$

$SNR = \dfrac{\sigma_h^2}{\sigma^2}$

$$\boxed{\underset{M \times 1}{\hat{h}} = \left(X^T X + \frac{1}{SNR} \underset{M \times M}{I}\right)^{-1} X^T \underset{N \times 1}{\bar{y}}}$$

This is very convenient to remember.

High SNR scenario: As $SNR \to \infty$, $\frac{1}{SNR} \to 0$

$$\Rightarrow \hat{h} = (X^T X)^{-1} X^T \bar{y}.$$

$\Rightarrow$ ML Estimate.

Therefore, MMSE estimate approaches the ML estimate at high SNR !!

Week 6 : Session 3

Consider Observation vector $(\bar{y}) = \begin{bmatrix} -2 \\ -3 \\ 1 \\ -2 \end{bmatrix}_{N\times 1}$

Pilot Matrix $(X) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & -1 \\ -1 & 1 \end{bmatrix}_{N\times M}$

What is the MMSE estimate $(\hat{h})$ when SNR = −6 dB = $\frac{1}{4}$

$$\boxed{\text{The MMSE estimate is given as} \quad \hat{h} = \left(X^TX + \frac{1}{SNR} I\right)^{-1} X^T \bar{y}.}$$

$$X^TX = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} = 4I.$$

$$X^TX + \frac{1}{SNR} I = 4I + \frac{1}{1/4} I = 4I + 4I = 8I$$

$$\left(X^TX + \frac{1}{SNR} I\right)^{-1} = \frac{1}{8} I$$

$$\left(X^TX + \frac{1}{SNR} I\right)^{-1} X^T = \frac{1}{8} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

$$= \frac{1}{8} \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

$$\left(X^TX + \frac{1}{SNR} I\right)^{-1} X^T \bar{y} = \frac{1}{8} \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ -3 \\ 1 \\ -2 \end{bmatrix}$$

$$= \frac{1}{8} \begin{bmatrix} -4 \\ -2 \end{bmatrix}$$

Therefore, MMSE estimate of $\bar{h}$

$$\boxed{\hat{h} = \begin{bmatrix} -1/2 \\ -1/4 \end{bmatrix}}$$

Week 7 : Session 1

MMSE Error Covariance

The Error covariance Matrix is given by

$$R_{(\hat{h}-\bar{h})(\hat{h}-\bar{h})} = E\left\{(\hat{h}-\bar{h})(\hat{h}-\bar{h})^T\right\}$$

$$= R_{hh} - R_{hy} R_{yy}^{-1} R_{yh}.$$

Note that, $R_{yh} = E\{\bar{y}\bar{h}^T\} = E\{(\bar{h}\bar{y}^T)^T\}$

$$= \left(E\{\bar{h}\bar{y}^T\}\right)^T = (R_{hy})^T$$

$$= (\sigma_h^2 X^T)^T = \sigma_h^2 X.$$

Therefore, one obtains the error covariance as

$$\sigma_h^2 I - \sigma_h^2 X^T(\sigma_h^2 XX^T + \sigma^2 I)^{-1}\sigma_h^2 X$$

$$\Rightarrow \sigma_h^2 I - \sigma_h^2(\sigma_h^2 X^T X + \sigma^2 I)^{-1}\sigma_h^2 X^T X$$

$N \times N$

$$X^T(\sigma_h^2 XX^T + \sigma^2 I)^{-1} = (\sigma_h^2 X^T X + \sigma^2 I)^{-1} X^T$$

$M \times M$

Add & Subtract $\sigma^2 I$

$$\Rightarrow \sigma_h^2 I - \sigma_h^2(\sigma_h^2 X^T X + \sigma^2 I)^{-1}\cdot(\underbrace{\sigma_h^2 X^T X + \sigma^2 I}_{I} - \sigma^2 I)$$

$$\Rightarrow \sigma_h^2 I - \sigma_h^2 (I - \sigma^2 I)(\sigma_h^2 X^T X + \sigma^2 I)^{-1}$$

$$\Rightarrow \cancel{\sigma_h^2 I} - \cancel{\sigma_h^2 I} + \sigma_h^2\sigma^2(\sigma_h^2 X^T X + \sigma^2 I)^{-1}$$

$$\Rightarrow \sigma_h^2\sigma^2(\sigma_h^2 X^T X + \sigma^2 I)^{-1}$$

$$\Rightarrow \sigma^2\left(X^T X + \frac{\sigma^2}{\sigma_h^2} I\right)^{-1}$$

$$\Rightarrow \sigma^2\left(X^T X + \frac{1}{SNR} I\right)^{-1}$$

$$SNR = \frac{\sigma_h^2}{\sigma^2}$$

Thus, the MMSE Error Covariance Matrix,

$$R_{(\hat{h}-h)(\hat{h}-h)} = E\left\{ (\hat{h}-h)(\hat{h}-h)^T \right\}$$

$$= R_{hh} - R_{hy} R_{yy}^{-1} R_{yh}$$

$$= \sigma^2 \left(X^T X + \frac{1}{SNR} I \right)^{-1}$$

The Mean Square Error (MSE),

$$E\left\{ \|\hat{h} - \bar{h}\|^2 \right\} = Tr\left\{ \text{Error Covariance} \right\}$$

$$= Tr\left\{ \sigma^2 \left(X^T X + \frac{1}{SNR} I \right)^{-1} \right\}$$

Example.

Consider $\bar{y} = \begin{bmatrix} -2 \\ -3 \\ 1 \\ -2 \end{bmatrix}$, $X = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & -1 \\ -1 & 1 \end{bmatrix}$.

What is Error Covariance and MSE when $SNR = -6\ dB = \frac{1}{4}$ and $\sigma^2 = 3\ dB = 2$.

MMSE Error Covariance,

$$R_{(\hat{h}-h)(\hat{h}-h)} = \sigma^2 \left(X^T X + \frac{1}{SNR} I \right)^{-1}$$

$$= 2\left(4I + \frac{1}{1/4} I \right)^{-1}$$

$$= 2\,(8I)^{-1}$$

$$= \frac{2}{8} I$$

$$= \frac{1}{4} I$$

Mean Square Error (MSE),

$$E\{\|\hat{h} - h\|^2\} = Tr\{\text{Error covariance}\}$$

$$= Tr\left\{\begin{bmatrix} 1/4 & 0 \\ 0 & 1/4 \end{bmatrix}\right\}$$

$$= \frac{1}{4} + \frac{1}{4}$$

$$= \frac{2}{4}$$

$$= \frac{1}{2}$$