eMasters in Communication Systems

Prof. Aditya Jagannatham

Elective Module:

Estimation for Wireless Communication

Chapter 4

Vector Parameter Estimation

Vector Parameter

$h \leftarrow$ Scalar Parameter

- Consider now a vector parameter

$$\bar{h} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_M \end{bmatrix}$$

M components

$M \times 1$

Vector parameter

Vector Parameter

- Consider now a vector parameter

$$\bar{\mathbf{h}} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_M \end{bmatrix}$$

$M \times 1$

Vector Parameter

$h_1, h_2, \ldots, h_M$

- The parameter has M components

$$\bar{\mathbf{h}} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_M \end{bmatrix}$$

Vector Parameter

- The parameter has M components

$$\bar{\mathbf{h}} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_M \end{bmatrix}$$

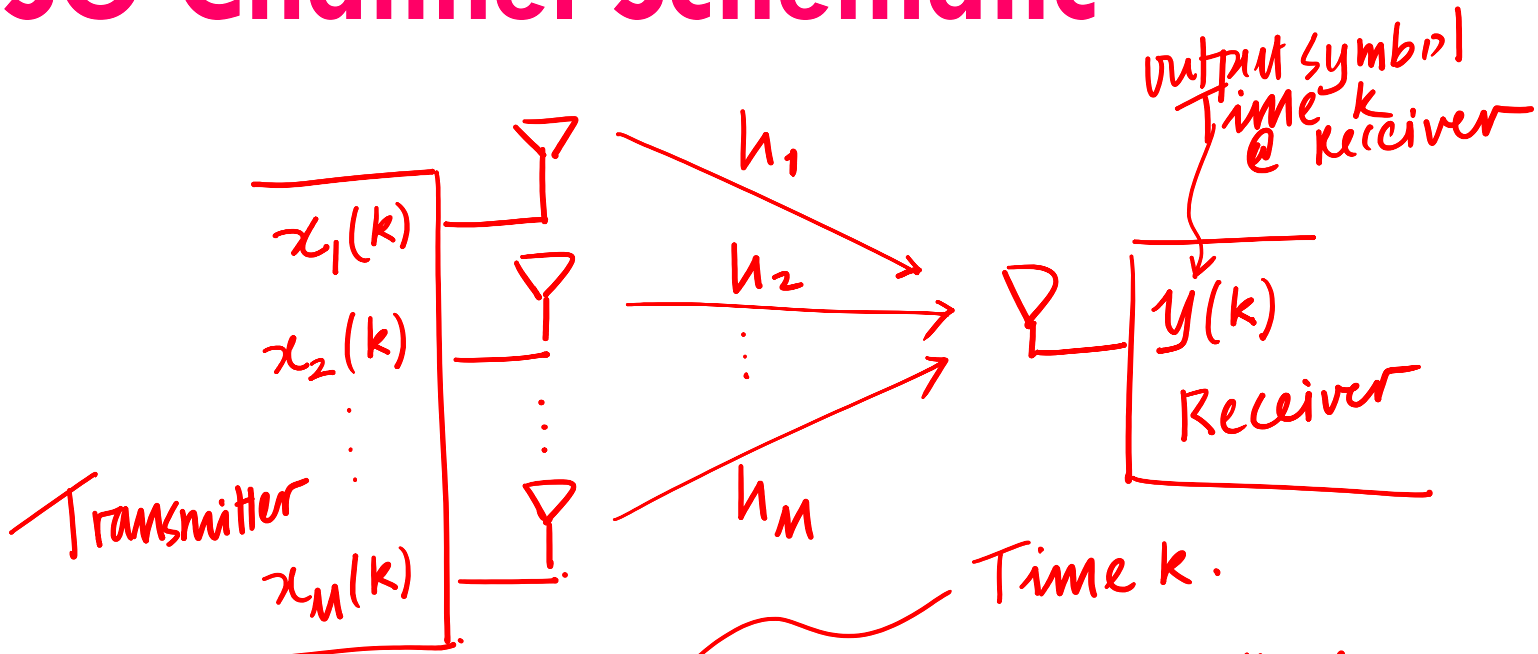
MISO Channel Estimation

MISO: Multiple input single output.

- Consider a system with M transmit antennas and single receive antenna
- Channel coefficients are

$$h_1, h_2, \ldots, h_M$$

MISO Channel Schematic



$x_i(k)$ = Symbol Transmitted From Transmit antenna i

MISO Channel Model

- The MISO system is given as

$$y(k) = x_1(k)h_1 + x_2(k)h_2 + \cdots + x_M(k)h_M + v(k)$$

$$= \underbrace{\begin{bmatrix} x_1(k) & x_2(k) & \cdots & x_M(k) \end{bmatrix}}_{\bar{x}^T(k)} \overbrace{\begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_M \end{bmatrix}}^{\bar{h}} + v(k)$$

MISO Channel Model

MISO
Multiple input
single output.

$$y(k) = \bar{x}^T(k)\,\bar{h} + v(k)$$

Pilot vector
@ Time $\underline{k}$.

$$\bar{x}(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_M(k) \end{bmatrix}.$$

MISO Channel Model

- The MISO system is given as

$$y(k) = x_1(k)h_1 + x_2(k)h_2 + \cdots + x_M(k)h_M + v(k)$$

$$= \overbrace{[x_1(k) \quad x_2(k) \quad \ldots \quad x_M(k)]}^{\bar{x}^T(k)} \overbrace{\begin{bmatrix} h_1 \\ h_1 \\ \vdots \\ h_M \end{bmatrix}}^{\bar{h}} + v(k)$$

MISO Channel Model

- The MISO system is given as

$$y(k) = [x_1(k) \quad x_2(k) \quad \ldots \quad x_M(k)] \begin{bmatrix} h_1 \\ h_1 \\ \vdots \\ h_M \end{bmatrix} + +v(k)$$

$$y(k) = \bar{\mathbf{x}}^T(k)\bar{\mathbf{h}} + v(k)$$

MISO Channel Model

- Consider now the transmission of N pilot vectors

$$y(1) = \bar{\mathbf{x}}^T(1)\bar{\mathbf{h}} + v(1)$$

$$y(2) = \bar{x}^T(2)\,\bar{h} + v(2)$$

$$\vdots$$

$$y(N) = \bar{x}^T(N)\,\bar{h} + v(N).$$

MISO Channel Model

- Consider now the transmission of N pilot vectors

N Time instants.

$$\underbrace{\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix}}_{\bar{y}} = \underbrace{\begin{bmatrix} \bar{\mathbf{x}}^T(1) \\ \bar{\mathbf{x}}^T(2) \\ \vdots \\ \bar{\mathbf{x}}^T(N) \end{bmatrix}}_{X} \bar{\mathbf{h}} + \underbrace{\begin{bmatrix} v(1) \\ v(2) \\ \vdots \\ v(N) \end{bmatrix}}_{\bar{v}}$$

MISO Channel Model

- This can be written in the matrix form vectors

$$\underbrace{\begin{bmatrix} y(1)\\ y(2)\\ \vdots \\ y(N)\end{bmatrix}}_{\bar{y}} = \underbrace{\overbrace{\begin{bmatrix} \bar{x}^T(1)\\ \bar{x}^T(2)\\ \vdots\\ \bar{x}^T(N)\end{bmatrix}}^{M\times N \text{ matrix}}}_{X} \underbrace{\overbrace{\begin{bmatrix} h_1\\ h_2\\ \vdots\\ h_M\end{bmatrix}}^{\text{parameter } M\times 1 \text{ vector}}}_{\bar{h}} + \underbrace{\overbrace{\begin{bmatrix} v(1)\\ v(2)\\ \vdots\\ v(N)\end{bmatrix}}^{N\times 1 \text{ Noise vector}}}_{\bar{v}}$$

MISO Channel Model

$\bar{y}$ = Output Vector $N \times 1$

- This can be written in the matrix form vectors

Parameter vector $M \times 1$

$$\underbrace{\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix}}_{\bar{\mathbf{y}}} = \underbrace{\begin{bmatrix} \bar{\mathbf{x}}^T(1) \\ \bar{\mathbf{x}}^T(2) \\ \vdots \\ \bar{\mathbf{x}}^T(N) \end{bmatrix}}_{\mathbf{X}} \bar{\mathbf{h}} + \underbrace{\begin{bmatrix} v(1) \\ v(2) \\ \vdots \\ v(N) \end{bmatrix}}_{\bar{\mathbf{v}}}$$

Pilot matrix $N \times M$

Noise vector $N \times 1$

MISO Channel Model

- The sizes of the various quantities are

$$\underbrace{\overset{N\times 1}{\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix}}}_{\bar{\mathbf{y}}} = \underbrace{\overset{N\times M}{\begin{bmatrix} \bar{\mathbf{x}}^T(1) \\ \bar{\mathbf{x}}^T(2) \\ \vdots \\ \bar{\mathbf{x}}^T(N) \end{bmatrix}}}_{\mathbf{X}} \overset{M\times 1}{\bar{\mathbf{h}}} + \underbrace{\overset{N\times 1}{\begin{bmatrix} v(1) \\ v(2) \\ \vdots \\ v(N) \end{bmatrix}}}_{\bar{\mathbf{v}}}$$

MISO Channel Model

- Note that X is of size $N \times M$, $N \geq M$
- This is known as a **Tall matrix**.

$$\begin{bmatrix} \bar{\mathbf{x}}^T(1) \\ \bar{\mathbf{x}}^T(2) \\ \vdots \\ \bar{\mathbf{x}}^T(N) \end{bmatrix}$$

$N \geq M$

$N \times M$

$\Rightarrow$ #rows $\geq$ #columns.

$\Longrightarrow$ "Tall matrix"

MISO Channel Model

- Therefore, the compact model is

Tall matrix

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

Likelihood

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

Gaussian
mean = 0
var = σ^2

- The likelihood for the estimation of $\bar{h}$ can be obtained as follows
- Note that PDF of $v(k)$ is

$\mathcal{N}(0, \sigma^2)$

$$\frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{1}{2\sigma^2} v^2(k)} = f_{v(k)}(v(k)).$$

Likelihood

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

- The likelihood for the estimation of $\bar{h}$ can be obtained as follows
- Note that PDF of $v(k)$ is

$$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{v^2(k)}{2\sigma^2}}$$

Likelihood

- The joint PDF of $\underbrace{v(1), v(2), \dots, v(N)}_{\text{independent identically distributed iid.}}$ is

$$\frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{v^2(1)}{2\sigma^2}} \times \cdots \times \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{v^2(N)}{2\sigma^2}}$$

$$= \underbrace{\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} \cdot e^{-\frac{1}{2\sigma^2}\sum_{k=1}^{N} v^2(k)}}_{\text{joint PDF}} = \left(\frac{1}{2\pi\sigma^2}\right)^{N/2} e^{-\frac{\|\bar{v}\|^2}{2\sigma^2}}$$

$$\sum_{k=1}^{N} v^2(k) = \|\bar{v}\|^2$$

Likelihood

- The joint PDF of $v(1), v(2), \ldots, v(N)$ is

$$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{v^2(1)}{2\sigma^2}} \times \cdots \times \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{v^2(N)}{2\sigma^2}}$$

$$= \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} e^{-\frac{1}{2\sigma^2}\sum_{k=1}^{N} v^2(k)} = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} e^{-\frac{1}{2\sigma^2}\|\bar{\mathbf{v}}\|^2}$$

Likelihood

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

Likelihood wrto $\bar{h}$

- The PDF of $\bar{\mathbf{y}}$ is

$\bar{y}$ = Gaussian
mean $= X\bar{h}$

$\bar{v} = \bar{y} - X\bar{h}$

$$\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} \cdot e^{-\frac{1}{2\sigma^2}\|\bar{y} - X\bar{h}\|^2} = p(\bar{y};\bar{h})$$

Likelihood

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

- The PDF of $\bar{\mathbf{y}}$ is

$$\overbrace{\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} e^{-\frac{1}{2\sigma^2}\|\bar{\mathbf{y}}-\mathbf{X}\bar{\mathbf{h}}\|^2}}^{p(\bar{y};\,\bar{h}).}$$

Likelihood

- Therefore, the likelihood of $\bar{\mathbf{h}}$ is

$$p(\bar{\mathbf{y}}; \bar{\mathbf{h}}) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} \cdot e^{-\frac{1}{2\sigma^2}\|\bar{y} - X\bar{h}\|^2}$$

$\bar{y} \sim N \times 1$ vector

$X \sim N \times M$ matrix

$\bar{h} \sim M \times 1$ vector

Likelihood

- Therefore, the likelihood of $\bar{\mathbf{h}}$ is

$$p(\bar{\mathbf{y}};\bar{\mathbf{h}}) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} e^{-\frac{1}{2\sigma^2}\left\|\bar{\mathbf{y}}-\mathbf{X}\bar{\mathbf{h}}\right\|^2}$$

Likelihood

- Once again, maximize the likeliho0d to estimate $\bar{\mathbf{h}}$

constant

$$\max p\left(\bar{\mathbf{y}};\bar{\mathbf{h}}\right)$$

maximize

constant

$$= \max\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} e^{-\frac{1}{2\sigma^2}\left\|\bar{\mathbf{y}}-\mathbf{X}\bar{\mathbf{h}}\right\|^2}$$

minimize

-ve sign

Least Squares

- Likelihood maximization reduces to

$$\min \underbrace{\|\bar{y} - Xh\|^2}_{\text{minimize this!!}}$$

- This is termed the **Least Squares (LS)** problem

minimize square of norm of error

Least Squares

- Likelihood maximization reduces to

$$\min \left\| \bar{\mathbf{y}} - \mathbf{X}\bar{\mathbf{h}} \right\|^2$$

- This is termed the **Least Squares (LS)** problem

Least Squares

- The solution to this problem is given as

$$\underbrace{\hat{h} = (X^T X)^{-1} X^T \bar{y}}_{\text{LS } \underline{\text{Solution}} \,..}$$

Least Squares

- The solution to this problem is given as

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

Least Squares

- The quantity $(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ is termed the pseudo-inverse of $\mathbf{X}$

X is Tall matrix ⇒ X is NOT invertible

$$(X^T X)^{-1}(X^T \cdot X) = I$$

acts as an inverse of X
⇒ Pseudoinverse.

Least Squares

- The quantity $(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ is termed the pseudo-inverse of $\mathbf{X}$

acts as a left inverse of X

$$(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T \times \mathbf{X} = \mathbf{I}$$

Properties of the LS Estimate

- The LS estimate is

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

Properties of the LS Estimate

- This can be simplified as

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

$$= \left(X^T X\right)^{-1} X^T \left(X \bar{h} + \bar{v}\right)$$

$$\hat{h} = \bar{h} + \left(X^T X\right)^{-1} X^T \bar{v}$$

Properties of the LS Estimate

- This can be simplified as

$$\hat{h} - \bar{h} = (X^T X)^{-1} X^T \bar{v}$$

$$\begin{aligned} \hat{\mathbf{h}} &= (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\bar{\mathbf{y}} \\ &= (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T(\mathbf{X}\bar{\mathbf{h}} + \bar{\mathbf{v}}) \\ &= \bar{\mathbf{h}} + (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\bar{\mathbf{v}} \end{aligned}$$

Properties of the LS Estimate

- Its mean is

$$E\{\hat{\mathbf{h}}\} = E\{\bar{h} + (X^T X)^{-1} X^T \bar{v}\}$$

$\bar{v}$ is zero mean noise

$$= \bar{h} + (X^T X)^{-1} X^T \cancelto{0}{E\{\bar{v}\}}$$

$$E\{\hat{h}\} = \bar{h} \Rightarrow \text{UNBIASED ESTIMATOR}$$

Properties of the LS Estimate

- Its mean is

$$E\{\hat{\mathbf{h}}\} = E\left\{\bar{\mathbf{h}} + \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\overline{\mathbf{v}}\right\}$$
$$= \bar{\mathbf{h}}$$

Properties of the LS Estimate

- Therefore, estimate is **unbiased**.

$$E\{\hat{\mathbf{h}}\} = E\left\{\bar{\mathbf{h}} + \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\bar{\mathbf{v}}\right\}$$
$$= \bar{\mathbf{h}} = \text{True parameter}$$

Properties of the LS Estimate

- The **covariance** of the estimate is given as

$$E\{(\hat{h}-\bar{h})(\hat{h}-\bar{h})^T\}$$

$$=E\{(X^TX)^{-1}X^T\bar{v}\bar{v}^TX(X^TX)^{-1}\}$$

$$=(X^TX)^{-1}X^TE\{\bar{v}\bar{v}^T\}X(X^TX)^{-1}$$

Properties of the LS Estimate

$$E\left\{ \bar{v}\,\bar{v}^T \right\} = E\left\{ \begin{bmatrix} v(1) \\ v(2) \\ \vdots \\ v(N) \end{bmatrix} \begin{bmatrix} v(1) & v(2) & \cdots & v(N) \end{bmatrix} \right\}$$

$$= E\left\{ \begin{bmatrix} v^2(1) & v(1)v(2) & \cdots \\ v(2)v(1) & v^2(2) & \\ \vdots & & \ddots \end{bmatrix} \right\} = \begin{bmatrix} \sigma^2 & & & \\ & \sigma^2 & & \\ & & \ddots & \\ & & & \sigma^2 \end{bmatrix}$$

$$= \sigma^2 I$$

Properties of the LS Estimate

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

$$= (X^TX)^{-1} X^T \sigma^2 I \, X (X^TX)^{-1}$$

$$= \sigma^2 \cancel{(X^TX)^{-1}} \, \cancel{X^T X} \cdot (X^TX)^{-1}$$

$$= \sigma^2 (X^TX)^{-1} = \underline{\text{Error covariance}}$$

Properties of the LS Estimate

- The **covariance** of the estimate is given as

$$
\begin{aligned}
&E\left\{\left(\hat{\mathbf{h}}-\bar{\mathbf{h}}\right)\left(\hat{\mathbf{h}}-\bar{\mathbf{h}}\right)^T\right\} \\
&= E\left\{\left(\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\bar{\mathbf{v}}\right)\left(\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\bar{\mathbf{v}}\right)^T\right\} \\
&= \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T E\{\bar{\mathbf{v}}\bar{\mathbf{v}}^T\}\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}
\end{aligned}
$$

Properties of the LS Estimate

- This can be simplified as

$$(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T E\{\bar{\mathbf{v}}\bar{\mathbf{v}}^T\}\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}$$
$$= (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\sigma^2\mathbf{I}\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}$$
$$= \sigma^2(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}$$
$$= \sigma^2(\mathbf{X}^T\mathbf{X})^{-1}$$

$\sigma^2(X^TX)^{-1}$

Error Covariance

MSE of Estimate

Mean Square Error

- MSE of the estimate is given as

Trace = Sum of Diagonal Elements of square matrix

$$\mathrm{Tr}\{\sigma^2(\mathbf{X}^T\mathbf{X})^{-1}\} = \sigma^2\mathrm{Tr}\{(\mathbf{X}^T\mathbf{X})^{-1}\}$$

MSE = Trace of Error covariance matrix

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

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

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

$$= E\{\|\bar{h}-\hat{h}\|^2\}$$

$$= \sum_{i=1}^{M} E\{(\hat{h}_i - h_i)^2\}$$

Sum of MSEs of individual Parameters,

ML Example

- Consider the pilot matrix

Pilot matrix

$$\mathbf{X} = \begin{bmatrix} 1 & 4 \\ 1 & 2 \\ 1 & 1 \\ 1 & 3 \end{bmatrix}$$

$N \times M$

$N = 4$ Pilot vectors

$M = 2$

$\Rightarrow$ # antennas

u_1, u_2

channel coefficients.

ML Example

- The output vector is

4x1.

$$\bar{\mathbf{y}} = \begin{bmatrix} -1 \\ 3 \\ -2 \\ -1 \end{bmatrix}$$

ML Example

- Start as follows

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

$$\mathbf{X}^T\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 4 & 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 1 & 2 \\ 1 & 1 \\ 1 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 10 \\ 10 & 30 \end{bmatrix} = X^T X$$

ML Example

- Start as follows

$$\mathbf{X}^T\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 4 & 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 1 & 2 \\ 1 & 1 \\ 1 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 10 \\ 10 & 30 \end{bmatrix}$$

MxM
2x2

ML Example

$$(\mathbf{X}^T\mathbf{X})^{-1} = \frac{1}{20}\begin{bmatrix} 30 & -10 \\ -10 & 4 \end{bmatrix}$$

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

ML Example

$$(\mathbf{X}^T\mathbf{X})^{-1} = \frac{1}{20}\begin{bmatrix} 30 & -10 \\ -10 & 4 \end{bmatrix}$$

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

ML Example

$$\hat{\mathbf{h}} = \overbrace{\frac{1}{20}\begin{bmatrix} 30 & -10 \\ -10 & 4 \end{bmatrix}}^{(X^TX)^{-1}} \overbrace{\begin{bmatrix} 1 & 1 & 1 & 1 \\ 4 & 2 & 1 & 3 \end{bmatrix}}^{X^T} \overbrace{\begin{bmatrix} -1 \\ 3 \\ -2 \\ -1 \end{bmatrix}}^{y}$$

$$= \frac{1}{20}\begin{bmatrix} -10 & 10 & 20 & 0 \\ 6 & -2 & -6 & 2 \end{bmatrix}\begin{bmatrix} -1 \\ 3 \\ -2 \\ -1 \end{bmatrix}$$

$$= \frac{1}{20}\begin{bmatrix} 0 \\ -2 \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 0 \\ -1 \end{bmatrix} = \hat{h}$$

$$= \begin{bmatrix} 0 \\ -\frac{1}{10} \end{bmatrix}$$

ML Example

$$\hat{\mathbf{h}} = \frac{1}{20}\begin{bmatrix} 30 & -10 \\ -10 & 4 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 4 & 2 & 1 & 3 \end{bmatrix}\begin{bmatrix} -1 \\ 3 \\ -2 \\ -1 \end{bmatrix}$$

$$= \frac{1}{20}\begin{bmatrix} -10 & 10 & 20 & 0 \\ 6 & -2 & -6 & 2 \end{bmatrix}\begin{bmatrix} -1 \\ 3 \\ -2 \\ -1 \end{bmatrix} = \frac{1}{20}\begin{bmatrix} 0 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{1}{10} \end{bmatrix}$$

ML Example

- Given $\sigma^2 = \frac{1}{2}$, what is the error covariance matrix and MSE?
- Error covariance is $= \sigma^2 (X^T X)^{-1}$

$$= \frac{1}{2} \times \frac{1}{20} \begin{bmatrix} 30 & -10 \\ -10 & 4 \end{bmatrix}$$

$$= \frac{1}{40} \begin{bmatrix} 30 & -10 \\ -10 & 4 \end{bmatrix}$$

ML Example

- Given $\sigma^2 = \frac{1}{2}$, what is the error covariance matrix and MSE?
- Error covariance is

$$\sigma^2(\mathbf{X}^T\mathbf{X})^{-1} = \frac{1}{2} \times \frac{1}{20}\begin{bmatrix} 30 & -10 \\ -10 & 4 \end{bmatrix}$$

$$= \underbrace{\frac{1}{40}\begin{bmatrix} 30 & -10 \\ -10 & 4 \end{bmatrix}}_{\text{Error Covariance}}$$

ML Example

- MSE is $E\{\|\hat{u}-\bar{u}\|^2\}$.

mean square error

$$= Tr(EC)$$

$$= Tr\left(\frac{1}{40}\begin{bmatrix}30 & -10\\ -10 & 4\end{bmatrix}\right)$$

$$= \frac{1}{40}\times 34 = \frac{17}{20}.$$

ML Example

- MSE is

$$Tr(\sigma^2(\mathbf{X}^T\mathbf{X})^{-1}) = Tr\left(\frac{1}{40}\begin{bmatrix} 30 & -10 \\ -10 & 4 \end{bmatrix}\right)$$

$$= \frac{34}{40} = \frac{17}{20}$$

Instructors may use this white area (14.5 cm / 25.4 cm) for the text.
Three options provided below for the font size.

Font: Avenir (Book), Size: 32, Colour: Dark Grey

Font: Avenir (Book), Size: 28, Colour: Dark Grey

Font: Avenir (Book), Size: 24, Colour: Dark Grey

Do not use the space below.