$$A = \sqrt{B_1^2 + B_2^2}$$
$$\phi = \text{ARCTAN}\left(\frac{-B_2}{B_1}\right)$$

LET $\underline{c} = [1 \ \cos 2\pi f_0 \ \ldots \ \cos 2\pi f_0 (N-1)]^T$
$\underline{s} = [0 \ \sin 2\pi f_0 \ \ldots \ \sin 2\pi f_0 (N-1)]^T$

$$J'(B_1, B_2, f_0) = \sum_n \left(x(n) - B_1 \cos 2\pi f_0 n - B_2 \sin 2\pi f_0 n\right)^2$$

$$= (\underline{x} - B_1 \underline{c} - B_2 \underline{s})^T (\underline{x} - B_1 \underline{c} - B_2 \underline{s})$$

$$= (\underline{x} - \underline{H}\underline{\alpha})^T (\underline{x} - \underline{H}\underline{\alpha})$$

WHERE $\underline{H} = [\underline{c} \ \underline{s}]$, $\underline{\alpha} = [B_1 \ B_2]^T$.

RECALL LINEAR MODEL SOLUTION (SEE (4.2), (4.3))

$$\hat{\underline{\alpha}} = (\underline{H}^T \underline{H})^{-1} \underline{H}^T \underline{x}$$

$$J'(\hat{B}_1, \hat{B}_2, f_0) = (\underline{x} - \underline{H}\hat{\underline{\alpha}})^T (\underline{x} - \underline{H}\hat{\underline{\alpha}})$$
$$= (\underline{x} - \underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T\underline{x})^T (\quad)$$
$$= \underline{x}^T (\underline{I} - \underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T)(\quad)^T \underline{x}$$

BUT $\underline{I} - \underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T$ IS IDEMPOTENT
OR $\underline{A}^2 = \underline{A}$ (SHOW THIS).

$$= \underline{x}^T (\underline{I} - \underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T)\underline{x}$$

$$= \text{FUNCTION OF } f_0 \text{ ONLY } \quad (\underline{H} = [\underline{c}\ \underline{s}])$$

$\Rightarrow$ MUST MAXIMIZE $\underline{x}^T \underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T\underline{x}$ OR

$$\underline{x}^T [\underline{c}\ \underline{s}]\left(\begin{bmatrix}\underline{c}^T \\ \underline{s}^T\end{bmatrix}[\underline{c}\ \underline{s}]\right)^{-1}\begin{bmatrix}\underline{c}^T \\ \underline{s}^T\end{bmatrix}\underline{x}$$

$$= \underline{x}^T \begin{bmatrix}\underline{c}^T \\ \underline{s}^T\end{bmatrix}^T\left(\begin{bmatrix}\underline{c}^T\underline{c} & \underline{c}^T\underline{s} \\ \underline{s}^T\underline{c} & \underline{s}^T\underline{s}\end{bmatrix}\right)^{-1}\begin{bmatrix}\underline{c}^T \\ \underline{s}^T\end{bmatrix}\underline{x}$$

$$= \begin{bmatrix}\underline{c}^T\underline{x} \\ \underline{s}^T\underline{x}\end{bmatrix}^T \begin{bmatrix}\underline{c}^T\underline{c} & \underline{c}^T\underline{s} \\ \underline{s}^T\underline{c} & \underline{s}^T\underline{s}\end{bmatrix}^{-1}\begin{bmatrix}\underline{c}^T\underline{x} \\ \underline{s}^T\underline{x}\end{bmatrix}$$

FOR EXACT MLE WE MAXIMIZE THIS OVER $0 < f_0 < \frac{1}{2}$ AND THEN FIND $\hat{\beta}_1$, $\hat{\beta}_2$ FROM

$$\hat{\underline{\alpha}} = (\hat{\underline{H}}^T\hat{\underline{H}})^{-1}\hat{\underline{H}}^T\underline{x}$$

AND FINALLY
$$\hat{A} = \sqrt{\hat{\beta}_1^2 + \hat{\beta}_2^2}$$
$$\hat{\phi} = \arctan\left(\frac{-\hat{\beta}_2}{\hat{\beta}_1}\right)$$

FOR APPROXIMATE MLE ASSUME $f_0$ NOT NEAR 0 OR $1/2$, THEN

$$\frac{1}{N}\underline{c}^T\underline{s} = \frac{1}{N}\sum_{n=0}^{N-1} \cos 2\pi f_0 n \sin 2\pi f_0 n$$

$$= \frac{1}{2N}\sum_n \sin 4\pi f_0 n \simeq 0$$

AND SIMILARLY

$$\frac{\underline{c}^T\underline{c}}{N} \approx \frac{1}{2} \qquad \frac{\underline{s}^T\underline{s}}{N} \approx \frac{1}{2}$$

MAXIMIZE $\begin{bmatrix} \underline{c}^T\underline{x} \\ \underline{s}^T\underline{x} \end{bmatrix}^T \begin{bmatrix} N/2 & 0 \\ 0 & N/2 \end{bmatrix}^{-1} \begin{bmatrix} \underline{c}^T\underline{x} \\ \underline{s}^T\underline{x} \end{bmatrix}$

$$= \frac{2}{N} \left[ (\underline{c}^T\underline{x})^2 + (\underline{s}^T\underline{x})^2 \right]$$

$$= \frac{2}{N} \left[ \left( \sum_n x(n) \cos 2\pi f_0 n \right)^2 + \left( \sum_n x(n) \sin 2\pi f_0 n \right)^2 \right]$$

$$= \frac{2}{N} \left| \sum_{n=0}^{N-1} x(n) e^{-j2\pi f_0 n} \right|^2$$

OR TO FIND APPROXIMATE MLE OF $f_0$ MAXIMIZE PERIODOGRAM

$$I(f_0) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x(n) e^{-j2\pi f_0 n} \right|^2$$

ALSO,

$$\hat{\underline{\alpha}} = (\hat{\underline{H}}^T\hat{\underline{H}})^{-1}\hat{\underline{H}}^T\underline{x}$$

$$\approx \frac{2}{N} \begin{bmatrix} \hat{\underline{c}}^T\underline{x} \\ \hat{\underline{s}}^T\underline{x} \end{bmatrix} = \begin{bmatrix} 2/N \sum_n x(n) \cos 2\pi \hat{f}_0 n \\ 2/N \sum_n x(n) \sin 2\pi \hat{f}_0 n \end{bmatrix}$$

$$\hat{A} = \sqrt{\hat{B}_1^2 + \hat{B}_2^2} = \frac{2}{N}\left| \sum_{n=0}^{N-1} x(n)\, e^{-j2\pi \hat{f}_0 n} \right|$$

$$\hat{\phi} = \arctan \frac{-\hat{B}_2}{\hat{B}_1} = \arctan \frac{-\sum_n x(n) \sin 2\pi \hat{f}_0 n}{\sum_n x(n) \cos 2\pi \hat{f}_0 n}.$$

## LEAST SQUARES

NO PROBABILISTIC ASSUMPTIONS ABOUT DATA - ONLY SIGNAL MODEL ASSUMED. HENCE, NO OPTIMALITY PROPERTIES.

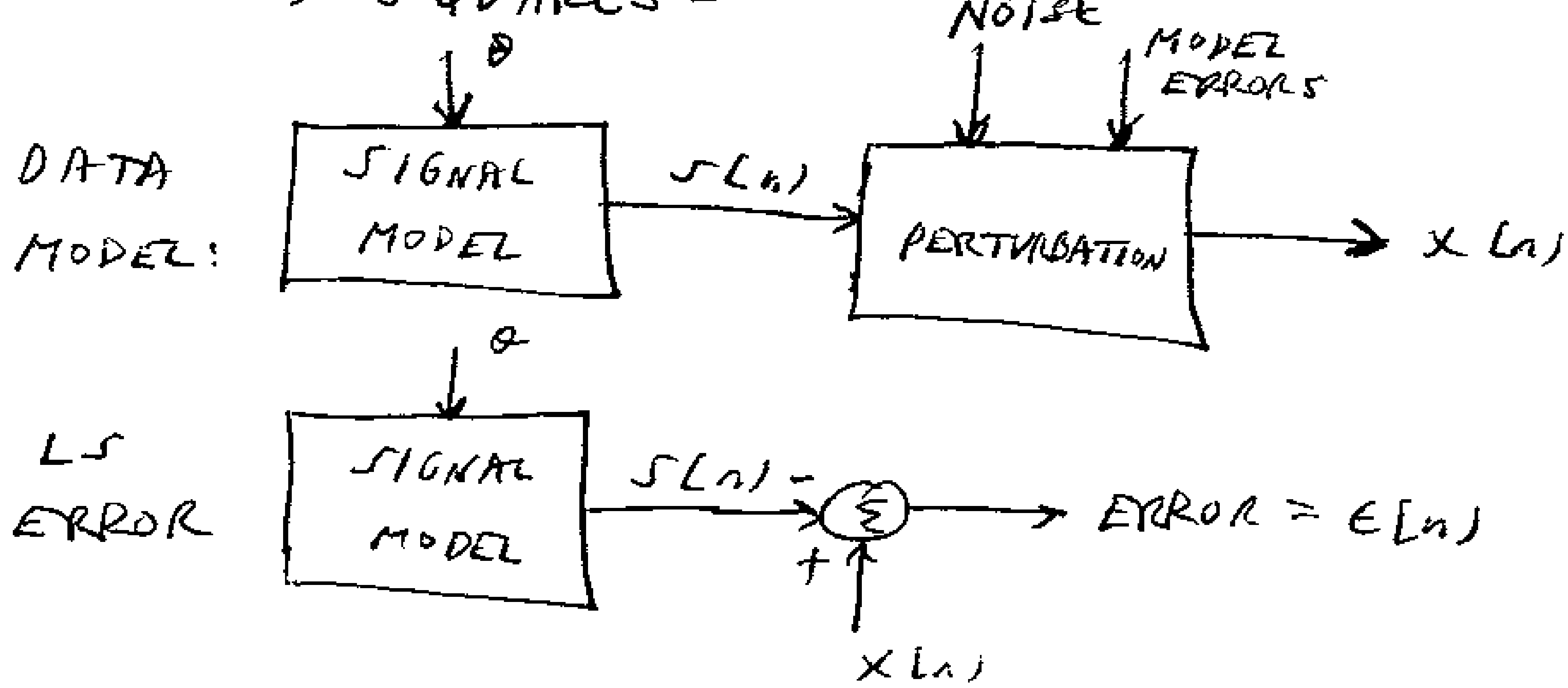
MVU CRITERION - MINIMIZE

$$E((\hat{\theta} - \theta)^2)$$

$$\theta = E(\hat{\theta})$$

NEED PDF OF $\underline{x}$ TO EVALUATE

LEAST SQUARES -



WE CHOOSE OUR ESTIMATE OF $\theta$ TO MAKE $s[n]$ CLOSE TO $x[n]$ – ANY SUBSEQUENT ERROR IS DUE TO NOISE OR MODEL ERRORS.

LS ERROR CRITERION

$$J(\theta) = \sum_{n=0}^{N-1} (x[n] - s[n])^2$$

↑DEPENDS ON $\theta$

VALUE OF $\theta$ THAT MINIMIZES $J$ IS LEAST SQUARES ESTIMATOR (LSE).

CAN FIND $\hat{\theta}$ WITH ONLY KNOWLEDGE OF $x[n]$ AND $s[n]$ – DON'T NEED PDF.

$\Rightarrow$ EASILY APPLIED IN PRACTICE TO "SIGNAL" IN NOISE PROBLEMS

EXAMPLE: DC LEVEL SIGNAL IN ??

SIGNAL MODEL: $s[n] = A$

$$J(A) = \sum_{n=0}^{N-1} (x[n] - A)^2$$

MINIMIZING $J$ OVER $A \Rightarrow \hat{A} = \bar{x}$

NOT OPTIMAL (ONLY IF $X[n] = A + \underset{\uparrow \text{WGN}}{w[n]}$)

WHAT HAPPENS IF $x[n] = A + w[n]$ AND $E(w[n]) \neq 0$?

$$w[n] = E(w[n]) + \underset{\uparrow \text{ZERO MEAN}}{w'[n]}$$

$$x[n] = \underbrace{A + E(w[n])}_{\bar{x} \text{ ESTIMATES THIS}} + w'[n]$$

UNDERLYING ASSUMPTION IS THAT ERRORS $\epsilon[n] = x[n] - s[n]$ TEND TO BE ZERO ON THE AVERAGE

<u>EXAMPLE</u> : SINUSOIDAL FREQ. ESTIMATION

$$s[n] = \cos 2\pi f_0 n$$

$$J(f_0) = \sum_{n=0}^{N-1} (x[n] - \cos 2\pi f_0 n)^2$$

HIGHLY NONLINEAR IN $f_0$ SINCE SIGNAL MODEL IS NOT <u>LINEAR</u> IN $\theta$.

PREVIOUS EXAMPLE $s(n) = \theta$

THIS " $s(n) = \cos 2\pi\theta n$

IF SIGNAL IS LINEAR IN $\theta \Rightarrow J$ IS QUADRATIC $\Rightarrow$ EASY MINIMIZATION

TERMED A LINEAR LS PROBLEM

OTHERWISE, NONLINEAR LS PROBLEM

EXAMPLE : SINUSOIDAL AMPLITUDE ESTIMATION

$$s(n) = A \cos 2\pi f_0 n$$

$$J(A) = \sum_{n=0}^{N-1} \left( x(n) - A \cos 2\pi f_0 n \right)^2$$

TO ESTIMATE $A$ FOR $f_0$ KNOWN $\Rightarrow$ LINEAR LS

" " $f_0$ " $A$ " $\Rightarrow$ NONLINEAR LS

TO ESTIMATE $A$ AND $f_0 \Rightarrow$ SEPARABLE LS

CAN MINIMIZE OVER $A$ AND THEN SUBSTITUTE (RECALL MLE PROBLEM) TO OBTAIN

$$J(\hat{A}, f_0) = \sum_{n=0}^{N-1} \left( x(n) - \underset{\substack{\uparrow \\ \text{FUNCTION OF } f_0}}{\hat{A}} \cos 2\pi f_0 n \right)^2$$

NOW HAVE

## LINEAR LS

MUST ASSUME $s[n] = \theta \underbrace{h[n]}_{\uparrow \text{KNOWN}}$

(SEE ALSO BLUE DISCUSSION)..

$$J(\theta) = \sum_{n=0}^{N-1} (x[n] - \theta h[n])^2$$

MINIMIZING THIS

$$\hat{\theta} = \frac{\sum_n x[n] h[n]}{\sum_n h^2[n]}$$

ALSO

$$J_{MIN} = J(\hat{\theta}) = \sum_n (x[n] - \hat{\theta} h[n])(x[n] - \hat{\theta} h[n])$$

$$= \sum_n x[n] (x[n] - \hat{\theta} h[n]) - \hat{\theta} \underbrace{\sum_n h[n] (x[n] - \hat{\theta} h[n])}_{S = 0}$$

$$= \sum_n x^2[n] - \hat{\theta} \sum_n x[n] h[n]$$

OR

$$J_{MIN} = \underbrace{\sum_n x^2[n]}_{\text{ORIGINAL ENERGY}} - \underbrace{\frac{\left(\sum_n x[n] h[n]\right)^2}{\sum_n h^2[n]}}_{\text{REDUCTION DUE TO}}$$

EXAMPLE : $s[n] = A \qquad h[n] = 1$

$$\hat{A} = \bar{x}$$

$$J_{MIN} = \sum_n x^2[n] - \hat{\theta} \sum_n x[n]\, h[n]$$

$$= \sum_n x^2[n] - \bar{x} \sum_n x[n]$$

$$= \sum_n x^2[n] - N\bar{x}^2$$

FOR NO NOISE $x[n] = A \Rightarrow \bar{x} = A$

$$J_{MIN} = NA^2 - NA^2 = 0$$

FOR A LOT OF NOISE (ZERO MEAN) OR

$$E(w^2[n]) \gg A^2$$

$$\frac{1}{N} \sum_n x^2[n] \gg \bar{x}^2$$

AND $J_{MIN} \approx \sum_{n=0}^{N-1} x^2[n]$

IN GENERAL

$$0 \le J_{MIN} \le \sum_n x^2[n]$$

PERFECT FIT

POOR FIT

IN GENERAL WE REQUIRE FOR A LINEAR LS PROBLEM

$$\underline{s} = \underset{N\times p}{\underline{H}}\ \underset{p\times 1}{\underline{\theta}}$$

LINEAR MODEL WITHOUT NOISE PDF ASSUMPTION.

TO FIND LSE MINIMIZE:

$$J(\underline{\theta}) = \sum_n (x[n] - s[n])^2$$

$$= (\underline{x} - \underline{H}\underline{\theta})^T(\underline{x} - \underline{H}\underline{\theta})$$

$$= \underline{x}^T\underline{x} - \underline{x}^T\underline{H}\underline{\theta} - \underline{\theta}^T\underline{H}^T\underline{x} + \underline{\theta}^T\underline{H}^T\underline{H}\underline{\theta}$$

$$= \underline{x}^T\underline{x} - 2\underline{\theta}^T\underline{H}^T\underline{x} + \underline{\theta}^T\underline{H}^T\underline{H}\underline{\theta}$$

$$\frac{\partial J}{\partial \underline{\theta}} = 0 - 2\underline{H}^T\underline{x} + 2\underline{H}^T\underline{H}\underline{\theta} = \underline{0}$$

$$\hat{\underline{\theta}} = (\underline{H}^T\underline{H})^{-1}\underline{H}^T\underline{x}. \qquad (\underline{H} \text{ IS FULL RANK})$$

SAME FUNCTIONAL FORM AS BLUE

IF $E(\underline{x}) = \underline{H}\underline{\theta}$, $\underline{C}_x = \sigma^2\underline{I} \Rightarrow$ BLUE

SAME FUNCTIONAL FORM AS EFFICIENT (MVU) ESTIMATOR FOR LINEAR MODEL

IF $E(\underline{x}) = \underline{H}\underline{\theta}$, $\underline{C}_x = \sigma^2\underline{I}$, $\underline{x}$ IS GAUSSIAN

$\Rightarrow$ MVU

TO FIND $J_{MIN}$:

$$\begin{aligned}
J_{MIN} &= J(\hat{\underline{\theta}}) \\
&= (\underline{x} - \underline{H}\hat{\underline{\theta}})^T(\underline{x} - \underline{H}\hat{\underline{\theta}}) \\
&= (\underline{x} - \underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T\underline{x})^T(\underline{x} - \underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T\underline{x}) \\
&= \underline{x}^T\underbrace{(\underline{I} - \underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T)(\underline{I} - \underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T)}_{\underline{I} - \underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T}\underline{x}
\end{aligned}$$

MATRIX IS IDEMPOTENT ($\underline{A}^2 = \underline{A}$)

$$J_{MIN} = \underline{x}^T(\underline{I} - \underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T)\underline{x}$$

OR $$J_{MIN} = \underline{x}^T(\underline{x} - \underline{H}\hat{\underline{\theta}})$$

## WEIGHTED LS

CONSIDER $J(\underline{\theta}) = (\underline{x} - \underline{H}\underline{\theta})^T\underline{W}(\underline{x} - \underline{H}\underline{\theta})$

↑ $N \times N$ POSITIVE DEFINITE WEIGHTING MATRIX

EXAMPLE: $\underline{W} = \begin{bmatrix} w_0 & & & \underline{0} \\ & w_1 & & \\ & & \ddots & \\ \underline{0} & & & w_{N-1} \end{bmatrix}$ $\quad w_n > 0$

$$J(A) = \sum_{n=0}^{N-1} w_n (x[n] - A)^2$$

IF $\underline{W}$ NOT POSITIVE DEFINITE, WE COULD

HAVE $W_0 \leq 0 \Rightarrow (x[n]-A)^2$ ERROR TERM WOULD BE ZERO OR NEGATIVE

RATIONALE FOR $\underline{W}$ IS TO WEIGHT LS ERROR TERMS. IF $x[n] = A + w[n]$, AND $w[n]$ SAMPLES ARE UNCORRELATED WITH VARIANCE $\sigma_n^2$, THEN REASONABLE TO CHOOSE

$$\underline{W} = \begin{bmatrix} 1/\sigma_0^2 & & & \underline{0} \\ & 1/\sigma_1^2 & & \\ & & \ddots & \\ \underline{0} & & & 1/\sigma_n^2 \end{bmatrix}$$

SO THAT $J(A) = \sum_{n=0}^{N-1} \frac{(x[n]-A)^2}{\sigma_n^2}$

$$\Rightarrow \hat{A} = \frac{\sum_n \frac{x[n]}{\sigma_n^2}}{\sum_n \frac{1}{\sigma_n^2}} \qquad \text{(FAMILIAR?)}$$

RESULTS :

$$\underline{\hat{\theta}} = (\underline{H}^T \underline{W} \underline{H})^{-1} \underline{H}^T \underline{W} \underline{x}$$

$$J_{MIN} = \underline{x}^T (\underline{W} - \underline{W}\,\underline{H}(\underline{H}^T \underline{W} \underline{H})^{-1} \underline{H}^T \underline{W}) \underline{x}$$

# GEOMETRICAL INTERPRETATION

RECALL $\underline{s} = \underline{H}\,\underline{\theta}$

LET $\underset{N \times p}{\underline{H}} = [\underline{h}_1\ \underline{h}_2 \ldots \underline{h}_p]$ (COLUMNS) $N \times 1$

$$\underline{s} = [\underline{h}_1\ \underline{h}_2 \ldots \underline{h}_p] \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_p \end{bmatrix}$$

$$= \sum_{i=1}^{p} \theta_i \underline{h}_i$$ = LINEAR COMBINATION OF "SIGNALS"

EXAMPLE : $s[n] = a \cos 2\pi f_0 n + b \sin 2\pi f_0 n$

$$\underline{\theta} = [a\ b]^T$$

$$\begin{bmatrix} s[0] \\ s[1] \\ \vdots \\ s[N-1] \end{bmatrix} = \begin{bmatrix} \underbrace{\begin{matrix} 1 \\ \cos 2\pi f_0 \\ \vdots \\ \cos 2\pi f_0(N-1) \end{matrix}}_{\underline{h}_1} & \underbrace{\begin{matrix} 0 \\ \sin 2\pi f_0 \\ \vdots \\ \sin 2\pi f_0(N-1) \end{matrix}}_{\underline{h}_2} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

$$\underline{s} = a\underline{h}_1 + b\underline{h}_2$$

RECALL LS ERROR

$$J(\underline{\theta}) = (\underline{x} - \underline{H}\,\underline{\theta})^T(\underline{x} - \underline{H}\,\underline{\theta})$$

BUT EUCLIDEAN LENGTH OF $N \times 1$ VECTOR $\underline{\xi} = [\xi_1\ \xi_2\ \ldots\ \xi_N]^T$ IS

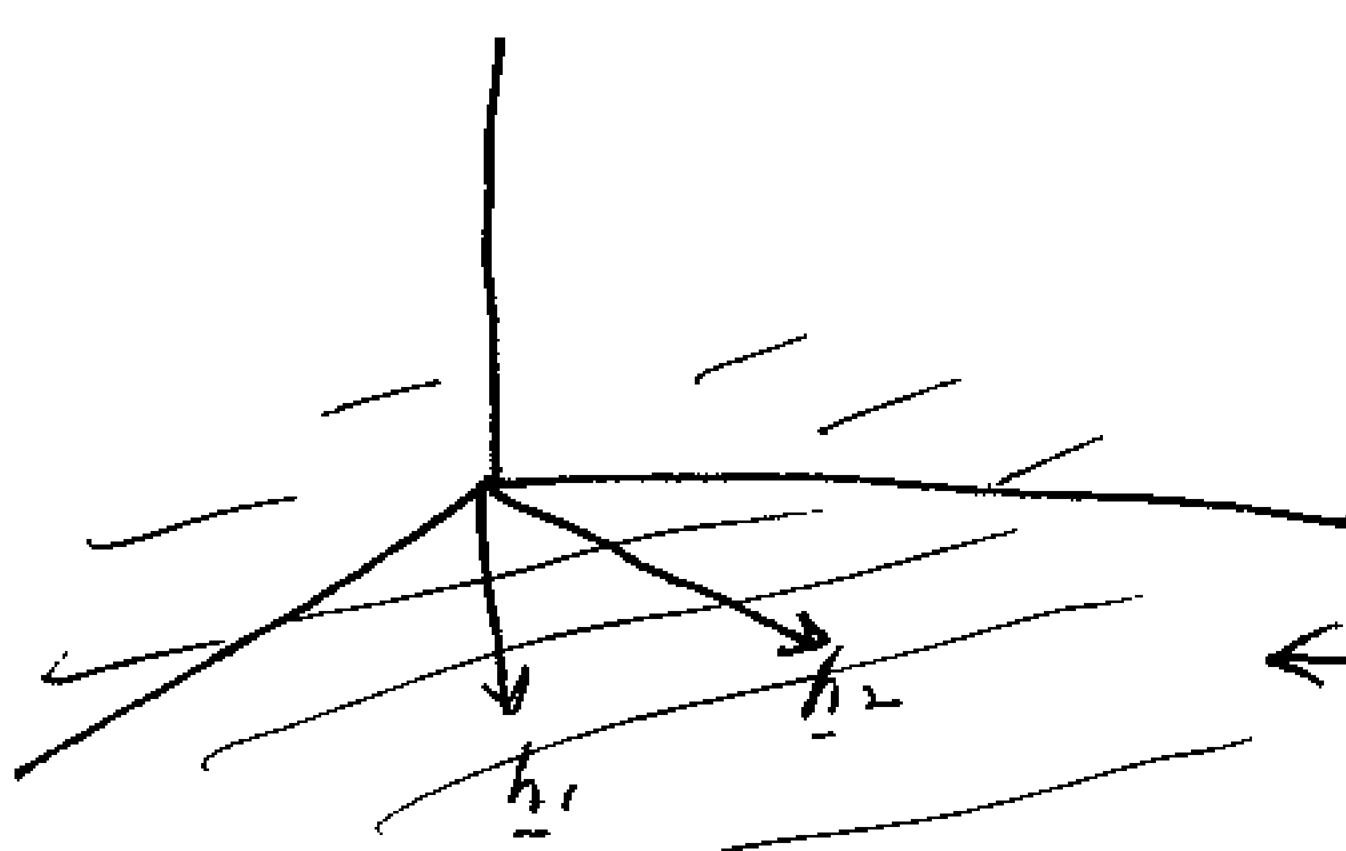
$$\|\underline{\xi}\| = \sqrt{\sum_{i=1}^{N} \xi_i^2} = \sqrt{\underline{\xi}^T\underline{\xi}}$$

$$\Rightarrow \quad J(\underline{\theta}) = \|\underline{x} - \underline{H}\,\underline{\theta}\|^2$$

$$= \|\underline{x} - \sum_{i=1}^{p} \theta_i \underline{h}_i\|^2$$

LS ATTEMPTS TO MINIMIZE THE SQUARE OF THE DISTANCE FROM $\underline{x}$ TO A SIGNAL VECTOR $\underline{s} = \sum_i \theta_i \underline{h}_i$ BY CHOOSING $\theta_i$'s.

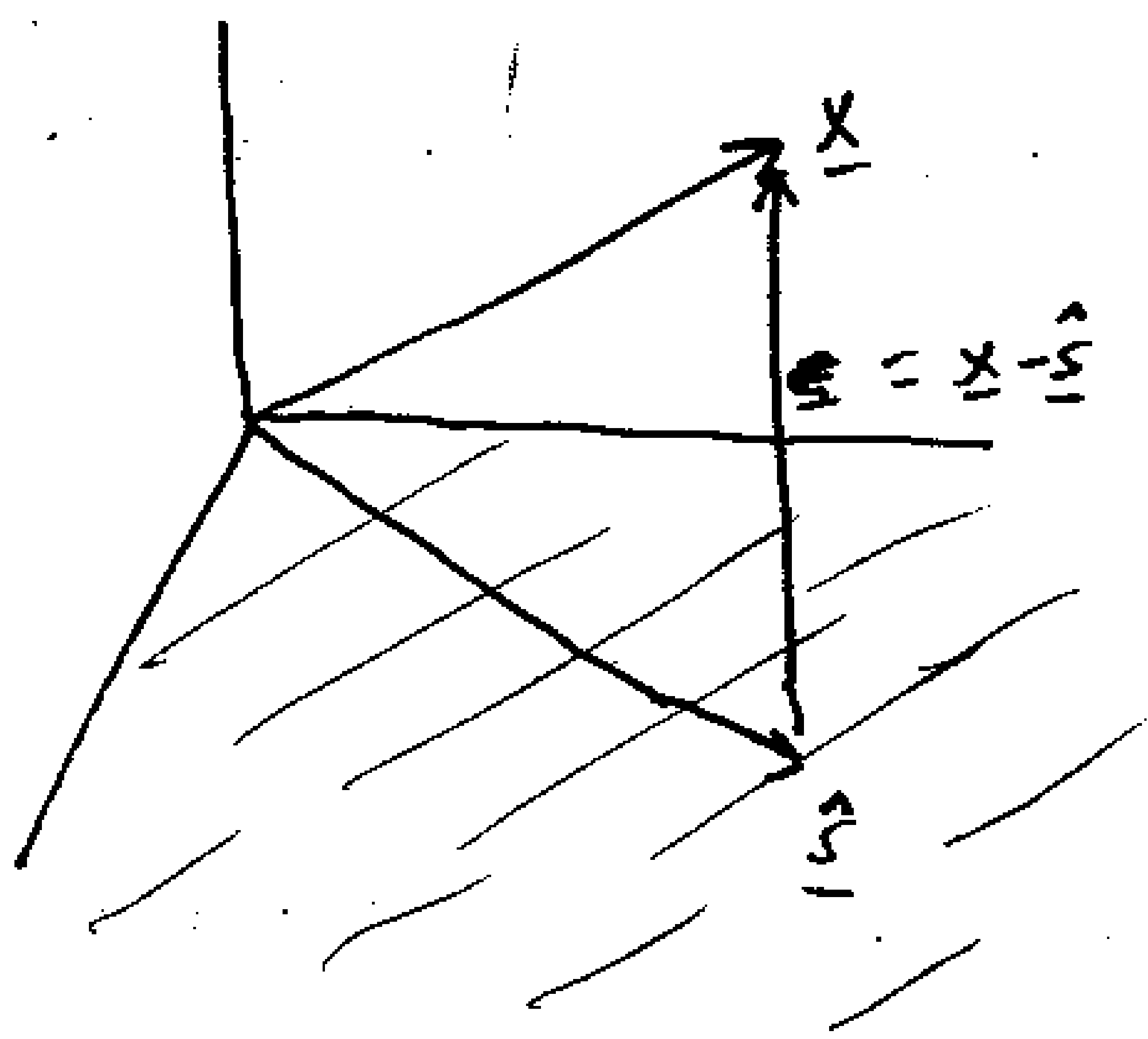
IN GENERAL, $\underline{x}$ LIES IN $R^N$ (N-DIMENSIONAL EUCLIDEAN SPACE) WHILE $\underline{s}$ MUST LIE IN $p$-DIMENSIONAL SUBSPACE, $S^p$.




$N = 3$
$p = 2$

$\leftarrow$ $S^2$ SPANNED BY $\underline{h}_1, \underline{h}_2$

TO FIND LS ESTIMATOR CHOOSE $\theta_1, \theta_2$ SO THAT $\|\underline{x} - \underline{s}\|^2$ IS MINIMUM.




$\underline{\epsilon} \perp S^2$

$\hat{\underline{s}}$ IS ORTHOGONAL PROJECTION OF $\underline{x}$ ONTO $S^2$.

ORTHOGONALITY OF $\underline{x}$ & $\underline{y}$ $\Rightarrow$ $\underline{x}^T \underline{y} = 0$

SINCE $\quad \underline{x} - \hat{\underline{s}} \perp S^2$

$$\Rightarrow \quad \begin{aligned} \underline{x} - \hat{\underline{s}} &\perp \underline{h}_1 \\ \underline{x} - \hat{\underline{s}} &\perp \underline{h}_2 \end{aligned}$$

OR
$$\begin{aligned} (\underline{x} - \hat{\underline{s}})^T \underline{h}_1 &= 0 \\ (\underline{x} - \hat{\underline{s}})^T \underline{h}_2 &= 0 \end{aligned}$$

BUT $\hat{\underline{s}} = \theta_1 \underline{h}_1 + \theta_2 \underline{h}_2$

$$(\mathbf{x} - \theta_1 \mathbf{h}_1 - \theta_2 \mathbf{h}_2)^T \mathbf{h}_1 = 0$$
$$(\mathbf{x} - \theta_1 \mathbf{h}_1 - \theta_2 \mathbf{h}_2)^T \mathbf{h}_2 = 0$$

$$(\mathbf{x} - \mathbf{H}\boldsymbol{\theta})^T \mathbf{h}_1 = 0$$
$$(\mathbf{x} - \mathbf{H}\boldsymbol{\theta})^T \mathbf{h}_2 = 0$$

$$\rightarrow \quad (\mathbf{x} - \mathbf{H}\boldsymbol{\theta})^T \underbrace{[\mathbf{h}_1 \ \mathbf{h}_2]}_{\mathbf{H}} = \mathbf{0}^T$$

$$\Rightarrow \quad \hat{\boldsymbol{\theta}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{x}$$

NOTE THAT ORTHOGONALITY CONDITION IS

$$\boldsymbol{\epsilon}^T \mathbf{H} = \mathbf{0}^T \qquad \boldsymbol{\epsilon} = \mathbf{x} - \underbrace{\mathbf{H}\boldsymbol{\theta}}_{\mathbf{s}}$$

$\Rightarrow$ ERROR IS ORTHOGONAL TO COLUMNS OF $\mathbf{H}$.

TO FIND LS ERROR

$$J_{MIN} = (\mathbf{x} - \mathbf{H}\hat{\boldsymbol{\theta}})^T(\mathbf{x} - \mathbf{H}\hat{\boldsymbol{\theta}})$$

$$= \mathbf{x}^T(\mathbf{x} - \mathbf{H}\hat{\boldsymbol{\theta}}) - \underbrace{\hat{\boldsymbol{\theta}}^T \mathbf{H}^T \boldsymbol{\epsilon}}_{= 0}$$

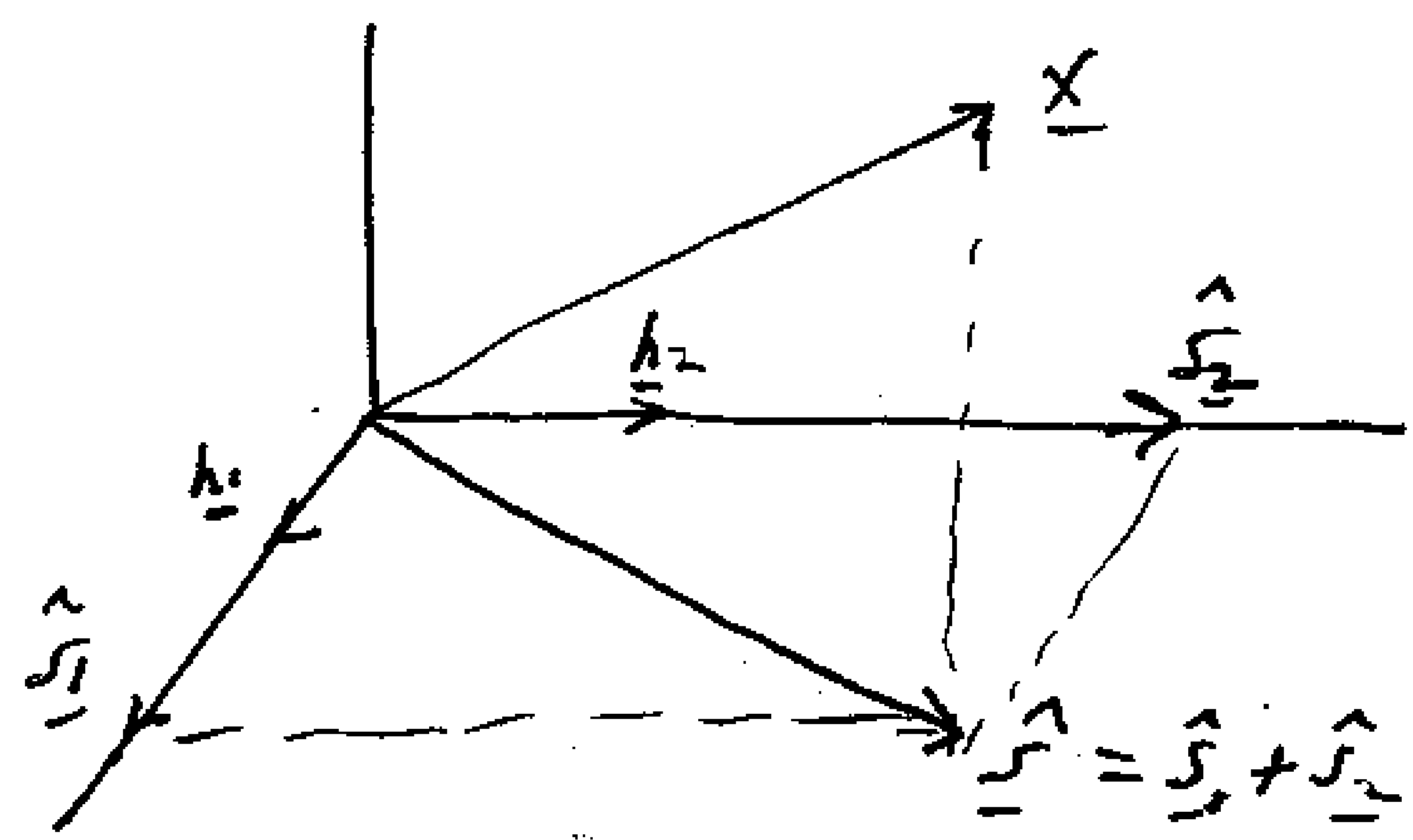
$$= \mathbf{x}^T\mathbf{x} - \mathbf{x}^T\mathbf{H}\hat{\boldsymbol{\theta}}$$

$$= \underline{x}^T\underline{x} - \underline{x}^T\underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T\underline{x}$$

$$= \underline{x}^T(\underline{I} - \underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T)\underline{x}$$

THIS IS JUST $\|\underline{x}-\hat{\underline{s}}\|^2$ OR LENGTH$^2$ OF ERROR VECTOR.

WHAT IF $\underline{h}_1, \underline{h}_2$ ARE $\perp$?




$\|\underline{h}_1\| = \|\underline{h}_2\| = 1$

$$\Rightarrow \quad \hat{\underline{s}} = \hat{\underline{s}}_1 + \hat{\underline{s}}_2 = \underbrace{(\underline{h}_1^T\underline{x})}_{} \underline{h}_1 + (\underline{h}_2^T\underline{x})\underline{h}_2$$

LENGTH OF $\underline{x}$ ALONG $\underline{h}_1$

$$= [\underline{h}_1 \; \underline{h}_2]\begin{bmatrix}\underline{h}_1^T\underline{x} \\ \underline{h}_2^T\underline{x}\end{bmatrix} = [\underline{h}_1 \; \underline{h}_2]\begin{bmatrix}\underline{h}_1^T \\ \underline{h}_2^T\end{bmatrix}\underline{x}$$

$$= \underline{H}\underbrace{\underline{H}^T\underline{x}}_{\hat{\underline{\theta}}}$$

SINCE COLUMNS OF $\underline{H}$ ARE $\perp \Rightarrow (\underline{H}^T\underline{H})^{-1}$

$$= (\underline{I})^{-1} = \underline{I}$$

OR $\hat{\underline{\theta}} = (\underline{H}^T\underline{H})^{-1}\underline{H}^T\underline{x} = \underline{H}^T\underline{x}$

SEE EXAMPLE 8.5

FOR NONORTHOGONAL COLUMNS

$$\hat{\underline{s}} = \underline{H}\hat{\underline{\theta}} = \underbrace{\underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T}_{\underline{P}\ N \times N}\underline{x}$$

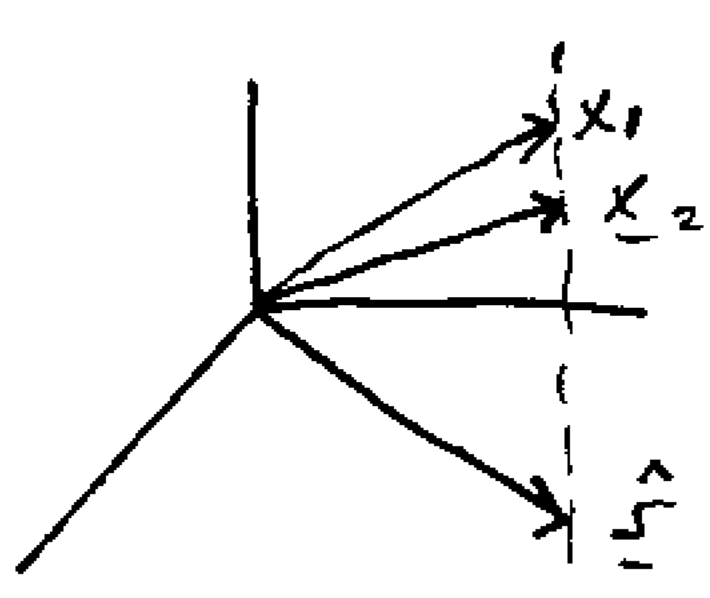
$\underline{P}$ IS CALLED THE PROJECTION MATRIX - PROJECTS $\underline{x}$ ONTO $p$- DIMENSIONAL SUBSPACE $S^p$.

PROPERTIES:

1) $\underline{P}^T = \underline{P}$ (SYMMETRIC - SEE PROB 8.11)
2) $\underline{P}^2 = \underline{P}$ (IDEMPOTENT)

$$\underline{P}^2\underline{x} = \underline{P}(\underline{P}\underline{x}) = \underline{P}\underbrace{\hat{\underline{s}}}_{\text{ALREADY IN } S^P} = \hat{\underline{s}} \Rightarrow \underline{P}^2 = \underline{P}$$

ALSO, $\underline{P}$ IS SINGULAR (RANK = $p$) WHY?



NOTE $\underline{\epsilon} = \underline{x} - \underline{\hat{s}} = \underline{x} - \underline{P}\underline{x} = \underbrace{(\underline{I} - \underline{P})}_{\underline{P}^{\perp}}\underline{x}$

$\underline{P}^{\perp}$ ALSO PROJECTION MATRIX – PROJECTS $\underline{x}$ ONTO SUBSPACE $\perp$ TO $S^P$.

$$
\begin{aligned}
J_{MIN} &= \underline{x}^T(\underline{I} - \underline{H}(\underline{H}^T\underline{H})^{-1}\underline{H}^T)\underline{x} \\
&= \underline{x}^T \underline{P}^{\perp}\underline{x} = \underline{x}^T \underline{P}^{\perp T}\underline{P}^{\perp}\underline{x} \\
&= \|\underline{P}^{\perp}\underline{x}\|^2.
\end{aligned}
$$

## ORDER-RECURSIVE LS

NOW SHOW HOW TO FIT SEVERAL MODELS TO DATA.


$s_1[n] = A$

$s_2[n] = A + Bn$

FOR MODEL 1 : $\underline{s} = \underline{H}_1 A \qquad \underline{H}_1 = [1\ 1\ \ldots\ 1]^T$

$\Rightarrow$ LSE $= \hat{A}_1 = \bar{x}$

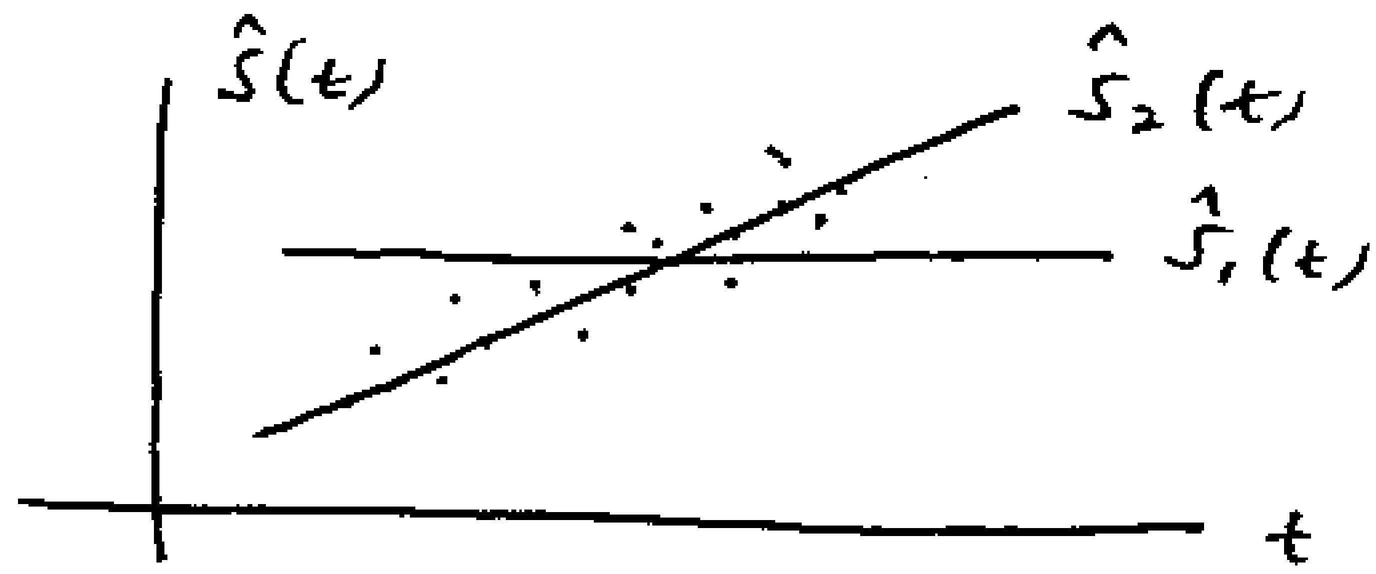
FOR MODEL 2 : $\underline{s} = \underline{H}_2 \begin{bmatrix} A \\ B \end{bmatrix}$

$$\underline{H}_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & N-1 \end{bmatrix}$$

LSE GIVEN BY (8.24). ASSUMING DATA WERE OBTAINED BY SAMPLING WE SUPPOSE

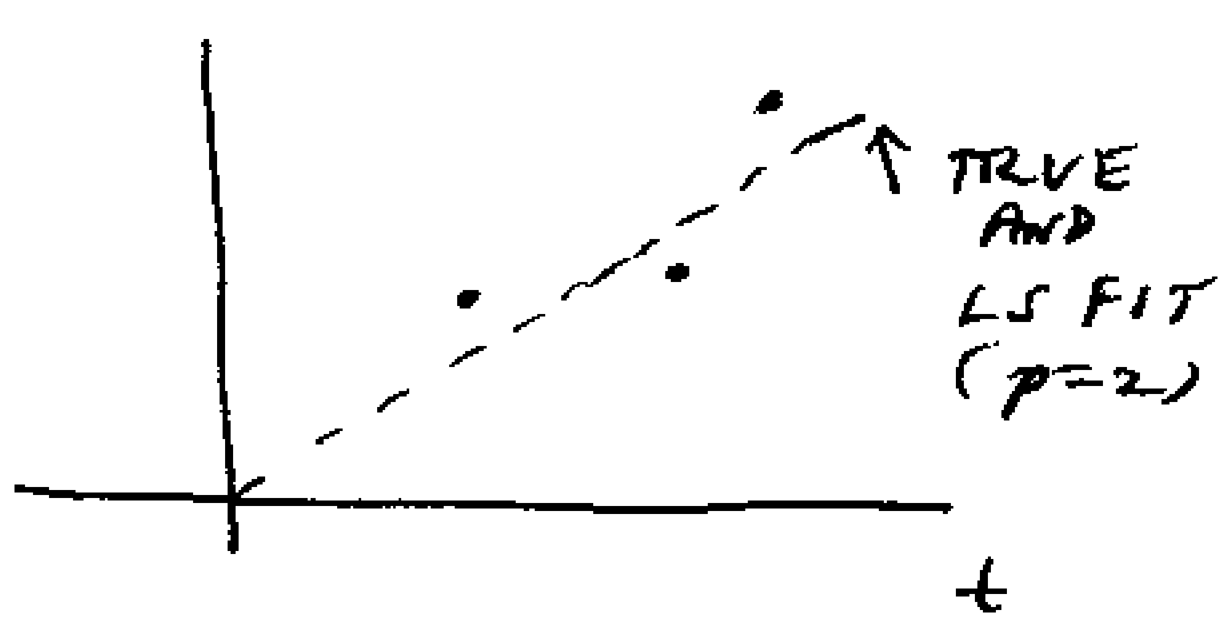
$$S_1(t) = A$$
$$S_2(t) = A + Bt$$

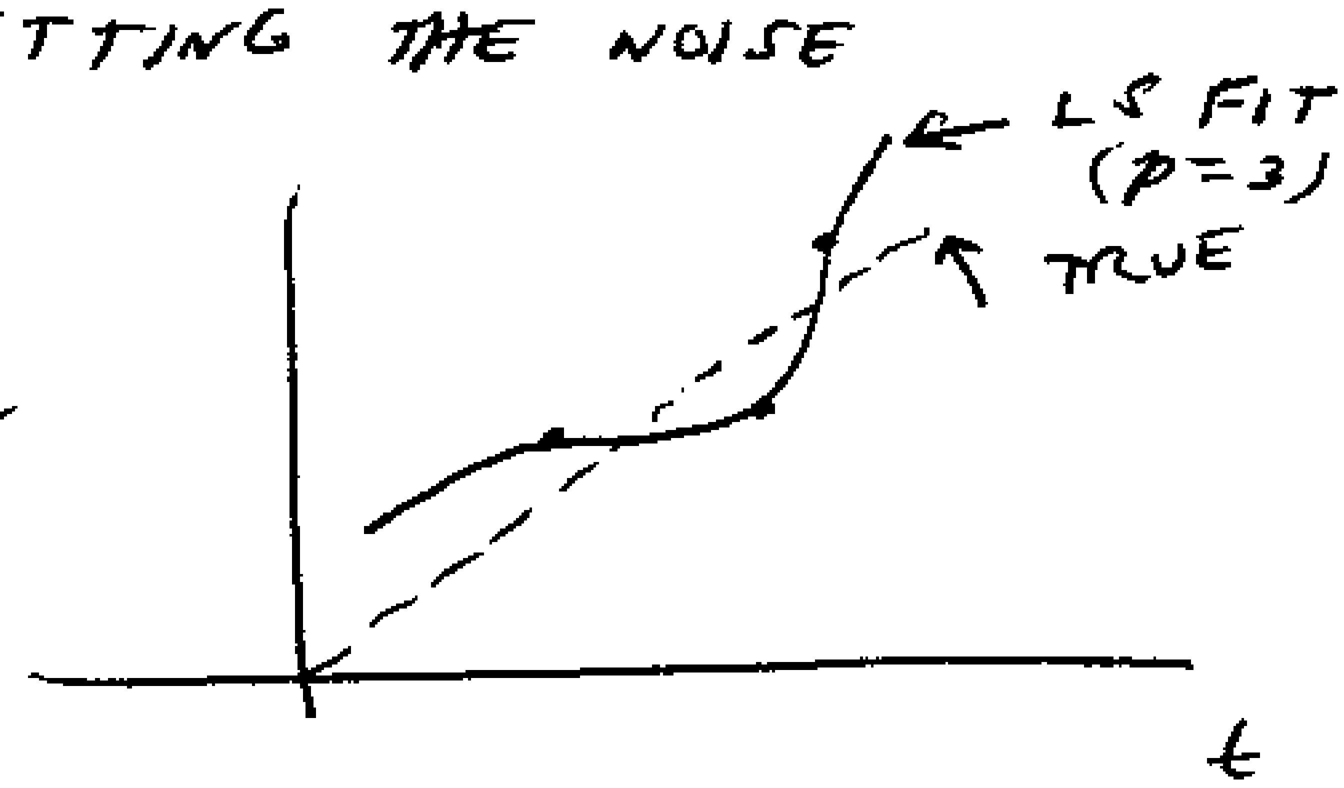
$$\Rightarrow \quad \hat{S}_1(t) = \hat{A}_1 \qquad \hat{S}_2(t) = \hat{A}_2 + \hat{B}_2 t$$




FIT WITH TWO PARAMETERS IS BETTER.
WILL SHOW THAT MINIMUM LS ERROR DECREASES AS WE ADD PARAMETERS.

NOTE: AS WE ADD PARAMETERS, HOWEVER, COULD BE FITTING THE NOISE




TRUE: $s(t) = A + Bt$