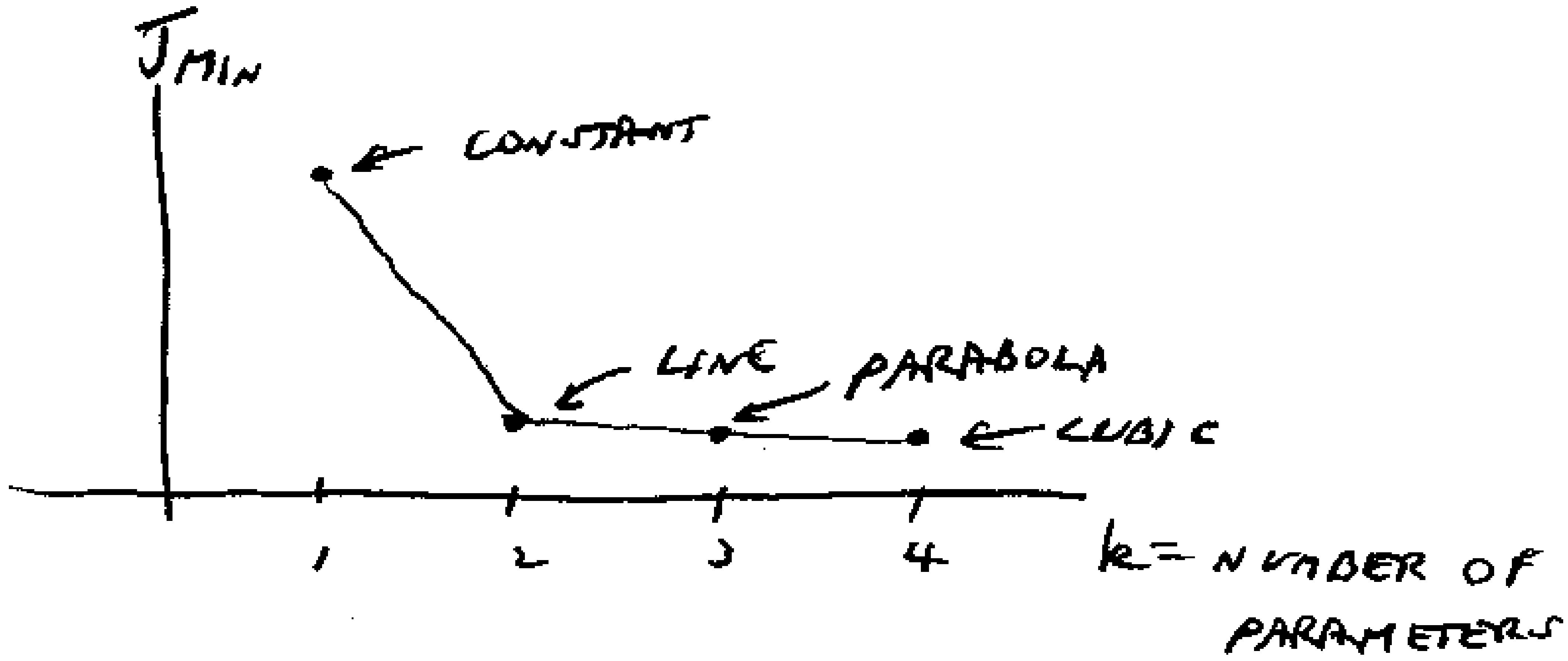
ONE SCHEME FOR DECIDING ON NUMBER OF PARAMETERS EXAMINES MINIMUM LS ERROR

FOR DATA OF FIG 8.5



WHEN J_{MIN} "BOTTOMS OUT", CHOOSE k AS CORRECT ORDER.

$J_{MIN} \simeq N\sigma^2$ AT OR ABOVE CORRECT ORDER SINCE

$$J_{MIN} = \sum_n (x[n] - \hat{s}[n])^2$$

$$= \sum_n (s[n] + w[n] - \hat{s}[n])^2$$

$$\simeq \sum_n w^2[n] \approx N\sigma^2$$

OTHERWISE, $J_{MIN} = \sum_n \left[(s[n] - \hat{s}[n]) + w[n]\right]^2$

$$\simeq \underbrace{\sum_n (s[n] - \hat{s}[n])^2}_{\text{MODELING ERROR}} + N\sigma^2$$

INSTEAD OF COMPUTING SOLUTION TO

$$\underline{H}^T\underline{H}\,\underline{\theta} = \underline{H}^T\underline{x}$$

FOR $\underline{\theta}$ 1×1, 2×1, 3×1, ETC. WE CAN USE AN ORDER UPDATE SOLUTION. (ALSO ALLOWS US TO SOLVE EQUATIONS WITHOUT INVERSION OF MATRIX)

LET $\underline{H}_k$ BE $N\times k$, $\hat{\underline{\theta}}_k$ BE LSE BASED ON $\underline{H}_k$. ASSUME THAT WE HAVE DETERMINED $\hat{\underline{\theta}}_k$ AND WE WISH TO FIND $\hat{\underline{\theta}}_{k+1}$.

IF $\underline{H}_{k+1} = (\underset{N\times k}{\underset{\uparrow}{\underline{H}_k}} \quad \underset{N\times 1\ (\text{NEW COLUMN})}{\underset{\uparrow}{\underline{h}_{k+1}}})$

$$\hat{\underline{\theta}}_{k+1} = \begin{bmatrix} \hat{\underline{\theta}}_k - \dfrac{(\underline{H}_k^T\underline{H}_k)^{-1}\underline{H}_k^T\underline{h}_{k+1}\underline{h}_{k+1}^T\underline{P}_k^{\perp}\underline{x}}{\underline{h}_{k+1}^T\underline{P}_k^{\perp}\underline{h}_{k+1}} \\ \\ \dfrac{\underline{h}_{k+1}^T\underline{P}_k^{\perp}\underline{x}}{\underline{h}_{k+1}^T\underline{P}_k^{\perp}\underline{h}_{k+1}} \end{bmatrix}$$

$$= \begin{pmatrix} k\times1 \\ 1\times1 \end{pmatrix}$$

WHERE $\underline{P}_k^{\perp} = \underline{I} - \underline{H}_k(\underline{H}_k^T\underline{H}_k)^{-1}\underline{H}_k^T$

= PROJECTION MATRIX ONTO SUBSPACE ORTHOGONAL TO SPAN $\{\underline{h}_1, \underline{h}_2, \ldots, \underline{h}_k\}$

TO AVOID INVERTING $\underline{H}_k^T \underline{H}_k$ (FOR $\underline{P}_k^{\perp}$)

LET $\underline{D}_k = (\underline{H}_k^T \underline{H}_k)^{-1}$

$$\underline{D}_{k+1} = \begin{bmatrix} \underline{D}_k + \dfrac{\underline{D}_k \underline{H}_k^T \underline{h}_{k+1} \underline{h}_{k+1}^T \underline{H}_k \underline{D}_k}{\alpha} & -\dfrac{\underline{D}_k \underline{H}_k^T \underline{h}_{k+1}}{\alpha} \\ -\dfrac{\underline{h}_{k+1}^T \underline{H}_k \underline{D}_k}{\alpha} & \dfrac{1}{\alpha} \end{bmatrix}$$

WHERE $\alpha = \underline{h}_{k+1}^T \underline{P}_k^{\perp} \underline{h}_{k+1}$

$$= \begin{bmatrix} k \times k & k \times 1 \\ 1 \times k & 1 \times 1 \end{bmatrix}$$

ALSO, $$J_{MIN_{k+1}} = J_{MIN_k} - \frac{(\underline{h}_{k+1}^T \underline{P}_k^{\perp} \underline{x})^2}{\underline{h}_{k+1}^T \underline{P}_k^{\perp} \underline{h}_{k+1}}$$

SINCE $\underline{h}_{k+1}^T \underline{P}_k^{\perp} \underline{h}_{k+1} = \| \underline{P}_k^{\perp} \underline{h}_{k+1} \|^2$

J_{MIN} DECREASES (OR STAYS SAME) WITH INCREASING ORDER.

EXAMPLE : LINE FITTING

$s_1[n] = A$

$s_2[n] = A + Bn$

$$\underline{H}_1 = [1\ 1\ \ldots\ 1]^T = \underline{1}$$

$$\underline{H}_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & N-1 \end{bmatrix} = [\ \underline{H}_1\ \underline{h}_2\ ]$$

TO START USE USUAL FORMULAS

$$\hat{A}_1 = \hat{\theta}_1 = (\underline{H}_1^T \underline{H}_1)^{-1} \underline{H}_1^T \underline{x}$$
$$= \bar{x}$$

$$J_{MIN} = (\underline{x} - \underline{H}_1 \hat{\theta}_1)^T (\underline{x} - \underline{H}_1 \hat{\theta}_1)$$
$$= \sum_n (x[n] - \bar{x})^2$$

LET $\hat{\underline{\theta}}_2 = (\hat{A}_2\ \hat{B}_2)^T$

$$\hat{\underline{\theta}}_2 = \begin{bmatrix} \hat{\theta}_1 - \dfrac{(\underline{H}_1^T \underline{H}_1)^{-1} \underline{H}_1^T \underline{h}_2 \underline{h}_2^T \underline{P}_1^{\perp} \underline{x}}{\underline{h}_2^T \underline{P}_1^{\perp} \underline{h}_2} \\ \dfrac{\underline{h}_2^T \underline{P}_1^{\perp} \underline{x}}{\underline{h}_2^T \underline{P}_1^{\perp} \underline{h}_2} \end{bmatrix}$$

$$(\underline{H}_1^T \underline{H}_1)^{-1} = 1/N$$
$$\underline{P}_1^{\perp} = \underline{I} - \underline{H}_1 (\underline{H}_1^T \underline{H}_1)^{-1} \underline{H}_1^T = \underline{I} - \tfrac{1}{N} \underline{1}\,\underline{1}^T$$

$$\underline{P}_1^{\perp} \underline{x} = \underline{x} - \bar{x}\,\underline{1}$$
$$\underline{h}_2^T \underline{P}_1^{\perp} \underline{x} = \underline{h}_2^T \underline{x} - \underline{h}_2^T \underline{H}_1 \underbrace{(\underline{H}_1^T \underline{H}_1)^{-1} \underline{H}_1^T \underline{x}}_{\hat{\theta}_1}$$
$$= \underline{h}_2^T \underline{x} - \underline{h}_2^T \underline{1}\,\bar{x}$$
$$= \sum n x[n] - \bar{x} \sum n$$

$$\underline{h}_2^T \underline{P}_1^{\perp} \underline{h}_2 = \underline{h}_2^T \underline{h}_2 - \tfrac{1}{N} (\underline{h}_2^T \underline{1})^2$$

$$= \sum n^2 - \frac{1}{N}\left(\sum n\right)^2$$

$$\hat{\underline{\theta}}_2 = \begin{bmatrix} \bar{x} - \dfrac{\frac{1}{N}\sum_n n\left(\sum_n n x[n] - \bar{x}\sum n\right)}{\sum n^2 - \frac{1}{N}\left(\sum n\right)^2} \\ \dfrac{\sum n x[n] - \bar{x}\sum n}{\sum n^2 - \frac{1}{N}\left(\sum n\right)^2} \end{bmatrix}$$

CAN BE SIMPLIFIED TO YIELD (F. 24).

NOTE: THIS PROCEDURE SOLVES ALL LS FIT PROBLEMS OF LOWER ORDER MODELS AS WELL.

PROCEDURE IS GRAM-SCHMIDT ORTHOGONALIZATION – SEE TEXT.

SEQUENTIAL LS

$$\hat{\underline{\theta}} = (\underline{H}^T\underline{H})^{-1}\underline{H}^T \begin{bmatrix} x[0] \\ x[1] \\ \vdots \\ x[N-1] \end{bmatrix}$$

IF WE NOW OBSERVE $x[N]$, CAN WE UPDATE $\hat{\underline{\theta}}$ WITHOUT RESOLVING EQUATIONS.

EXAMPLE: DC LEVEL

RECALL $\hat{A} = \frac{1}{N}\sum_{n=0}^{N-1} x[n]$

LET $\hat{A}(N-1)$ = LSE BASED ON $\{x[0], x[1], \ldots, x[N-1]\}$

WHAT IS $\hat{A}[N]$?

$$\hat{A}[N] = \frac{1}{N+1} \sum_{n=0}^{N} x[n]$$

NO NEED TO RECOMPUTE SUM SINCE

$$\hat{A}[N] = \frac{1}{N+1} \left(\sum_{n=0}^{N-1} x[n] + x[N] \right)$$

$$= \frac{N}{N+1} \hat{A}[N-1] + \frac{1}{N+1} x[N]$$

NEW ESTIMATE

OLD ESTIMATE

$$\hat{A}[N] = \hat{A}[N-1] + \underbrace{\frac{1}{N+1} \left(x[N] - \hat{A}[N-1] \right)}_{\text{CORRECTION OF } x[N]}$$

OLD ESTIMATE

WE WEIGHT "ERROR" $x[N] - \hat{A}[N-1]$ BY $\frac{1}{N+1}$

$\Rightarrow$ AS N INCREASES, LESS WEIGHT WHY?

ERROR = $x[N] - \hat{A}[N-1]$

PREDICTION OF $x[N]$
(NO NOISE $\ni$ $x[N] = A$)

NOW CONSIDER A WEIGHTED LS PROBLEM.

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

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

$$\hat{A}[N-1] = \frac{\sum_{n=0}^{N-1} x[n]/\sigma_n^2}{\sum_{n=0}^{N-1} 1/\sigma_n^2}$$

TO UPDATE IN TIME

$$\hat{A}[N] = \frac{\sum_{n=0}^{N} \frac{x[n]}{\sigma_n^2}}{\sum_{n=0}^{N} \frac{1}{\sigma_n^2}}$$

$$= \frac{\sum_{n=0}^{N-1} \frac{x[n]}{\sigma_n^2} + \frac{x[N]}{\sigma_N^2}}{\sum_{n=0}^{N} 1/\sigma_n^2}$$

$$= \frac{\sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \hat{A}[N-1]}{\sum_{n=0}^{N} 1/\sigma_n^2} + \frac{x[N]/\sigma_N^2}{\sum_{n=0}^{N} 1/\sigma_n^2}$$

$$= \hat{A}[N-1] - \frac{1/\sigma_N^2 \, \hat{A}[N-1]}{\sum_{n=0}^{N} 1/\sigma_n^2} + \frac{x[N]/\sigma_N^2}{\sum_{n=0}^{N} 1/\sigma_n^2}$$

$$\Rightarrow \hat{A}[N] = \hat{A}[N-1] + \frac{1/\sigma_N^2}{\sum_{n=0}^{N} 1/\sigma_n^2} \left(x[N] - \hat{A}[N-1] \right)$$

IF $\sigma_n^2 = \sigma^2 \Rightarrow$ PREVIOUS RESULT, OTHERWISE

IF $\sigma_N^2 \to \infty$, $\dfrac{1/\sigma_N^2}{\sum_{n=0}^{N} 1/\sigma_n^2} \to 0 \Rightarrow \hat{A}[N] = \hat{A}[N-1]$

IF $\sigma_N^2 \to 0$, $\dfrac{1/\sigma_N^2}{\sum_{n=0}^{N} 1/\sigma_n^2} \to 1 \Rightarrow \hat{A}[N] = x[N]$

GAIN FACTOR $K[N] = \dfrac{1/\sigma_N^2}{\sum_{n=0}^{N} 1/\sigma_n^2}$ DETERMINES CONFIDENCE IN NEW DATA SAMPLE.

ALSO, IF $x[n] = A + w[n]$, WHERE $w[n]$ IS ZERO MEAN UNCORRELATED NOISE WITH $\mathrm{VAR}(w[n]) = \sigma_n^2$

$\Rightarrow$ $\hat{A}[N-1]$ IS BLUE AND

$$\mathrm{VAR}(\hat{A}[N-1]) = \frac{1}{\sum_{n=0}^{N-1} 1/\sigma_n^2}$$

$$\Rightarrow \quad K[N] = \frac{1/\sigma_N^2}{\sum_{n=0}^{N} 1/\sigma_n^2} = \frac{1/\sigma_N^2}{\frac{1}{VAR(\hat{A}[N-1])} + \frac{1}{\sigma_N^2}}$$

$$= \frac{VAR(\hat{A}[N-1])}{VAR(\hat{A}[N-1]) + \sigma_N^2}$$

$$0 \le K[N] \le 1$$

THE CORRECTION TO $\hat{A}[N-1]$ DEPENDS ON GAIN, WHICH DEPENDS ON OUR CONFIDENCE IN PREVIOUS ESTIMATE (VIA $VAR(\hat{A}[N-1])$) AND NEW DATA SAMPLE (VIA σ_N^2).

WE WILL CONTINUE TO THINK OF $\frac{1}{\sum_{n=0}^{N-1} 1/\sigma_n^2}$ AS THE VARIANCE OF THE BLUE.

TO COMPUTE $VAR(\hat{A}[N])$ BASED ON $VAR(\hat{A}[N-1])$

$$VAR(\hat{A}[N]) = \frac{1}{\sum_{n=0}^{N} 1/\sigma_n^2} = \frac{1}{\sum_{n=0}^{N-1} 1/\sigma_n^2 + 1/\sigma_N^2}$$

$$= \frac{1}{\frac{1}{VAR(\hat{A}[N-1])} + 1/\sigma_N^2}$$

$$= \frac{VAR(\hat{A}[N-1])\,\sigma_N^2}{VAR(\hat{A}[N-1]) + \sigma_N^2}$$

$$= \left(1 - \frac{VAR(\hat{A}[N-1])}{VAR(\hat{A}[N-1]) + \sigma_N^2} \right) VAR(\hat{A}[N-1])$$

$$\therefore \quad VAR(\hat{A}[N]) = (1 - K[N]) \, VAR(\hat{A}[N-1])$$

SUMMARY OF SEQUENTIAL LSE

ESTIMATOR UPDATE:

$$\hat{A}[N] = \hat{A}[N-1] + K[N] \left(x[N] - \hat{A}[N-1] \right)$$

$$\text{WHERE} \quad K[N] = \frac{VAR(\hat{A}[N-1])}{VAR(\hat{A}[N-1]) + \sigma_N^2}$$

VARIANCE UPDATE:

$$VAR(\hat{A}[N]) = (1 - K[N]) \, VAR(\hat{A}[N-1])$$

INITIALIZATION: $\hat{A}[0] = x[0]$

$VAR(\hat{A}[0]) = \sigma_0^2$.

PROCEDURE:

1) INITIALIZE
2) COMPUTE $K[1]$ USING $VAR(\hat{A}[0])$ AND σ_1^2
3) COMPUTE $\hat{A}[1]$ USING $\hat{A}[0]$, $K[1]$, $x[1]$
4) COMPUTE $VAR(\hat{A}[1])$ USING $K[1]$, $VAR(\hat{A}[0])$.
5) COMPUTE $K[2]$

ETC

EXAMPLE: $A = 10$, $\sigma_n^2 = 1$ MONTE CARLO COMPUTER SIM.

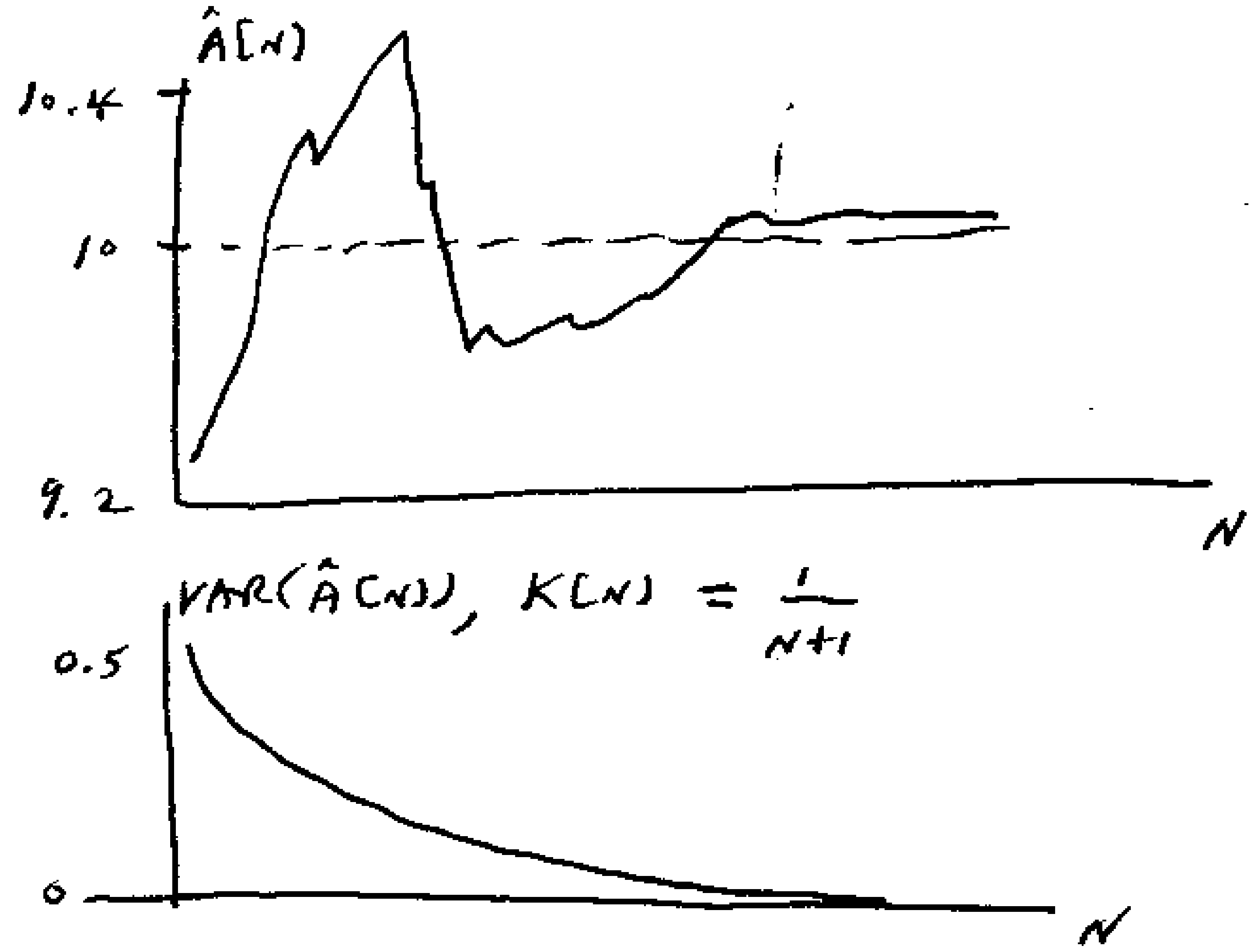


FIG. 8.9c

APPROACH IS TERMED SEQUENTIAL LS. USUAL APPROACH WHERE WE WAIT FOR ALL DATA AND PROCESS AT ONE TIME IS BATCH APPROACH.

GENERAL SEQUENTIAL LS

CONSIDER SEQUENTIAL WEIGHTED LSE OR IF $\underline{W} = \underline{C}^{-1}$ (ASSUME BLUE SETUP)

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

$$\underline{C}_{\hat{\theta}} = (\underline{H}^T \underline{C}^{-1} \underline{H})^{-1}$$

TO COMPUTE $\hat{\theta}$ SEQUENTIALLY MUST ASSUME THAT $\underline{C}$ IS DIAGONAL OR

UNCORRELATED NOISE.

LET $\underline{C}(n) = \begin{bmatrix} \sigma_0^2 & & \underline{0} \\ & \sigma_1^2 & \\ & & \ddots \\ \underline{0} & & & \sigma_n^2 \end{bmatrix}$ $(n+1)\times(n+1)$

$$\underline{H}(n) = \begin{bmatrix} \underline{H}[n-1] \\ \underline{h}^T[n] \end{bmatrix} = \begin{bmatrix} n\times p \\ 1\times p \end{bmatrix}$$

$$\underline{x}[n] = (x[0]\ x[1] \ldots x[n])^T$$

$\hat{\underline{\theta}}(n) =$ LSE BASED ON $\underline{x}[n]$ OR $n+1$ DATA SAMPLES.

BATCH ESTIMATOR:

$$\hat{\underline{\theta}}[n] = \left(\underline{H}^T(n)\, \underline{C}^{-1}[n]\, \underline{H}(n)\right)^{-1} \underline{H}^T(n)\, \underline{C}^{-1}[n]\, \underline{x}(n)$$

$$\underline{C}_{\hat{\theta}} = \underline{\Sigma}[n] = \left(\underline{H}^T(n)\, \underline{C}^{-1}[n]\, \underline{H}(n)\right)^{-1}$$

COVARIANCE OF LSE

COVARIANCE OF NOISE

ESTIMATOR UPDATE:

$$\hat{\underline{\theta}}[n] = \hat{\underline{\theta}}[n-1] + \underline{K}[n]\left(x[n] - \underline{h}^T[n]\,\hat{\underline{\theta}}[n-1]\right)$$

WHERE
$$\underset{\substack{\uparrow \\ p\times 1}}{\underline{K}[n]} = \frac{\underline{\Sigma}[n-1]\,\underline{h}[n]}{\sigma_n^2 + \underline{h}^T[n]\,\underline{\Sigma}[n-1]\,\underline{h}[n]}$$

COVARIANCE UPDATE:

$$\underset{\substack{\uparrow \\ p\times p}}{\underline{\Sigma}[n]} = \left(\underline{I} - \underline{K}[n]\,\underline{h}^T[n]\right)\underline{\Sigma}[n-1]$$

NO MATRIX INVERSIONS REQUIRED.

STEPS:
1) INITIALIZE $\Rightarrow \hat{\underline{\theta}}[0],\ \underline{\Sigma}[0]$
2) COMPUTE $\underline{K}[1]$ USING $\underline{\Sigma}[0],\ \underline{h}[1],\ \sigma_1^2$
3) COMPUTE $\hat{\underline{\theta}}[1]$ USING $\hat{\underline{\theta}}[0],\ \underline{K}[1],\ x[1]$, AND $\underline{h}[1]$
4) COMPUTE $\underline{\Sigma}[1]$ USING $\underline{K}[1],\ \underline{h}[1],\ \underline{\Sigma}[0]$
5) COMPUTE $\underline{K}[2]$

ETC.

WE HAVE SLIGHT PROBLEM IN STEP 1. TO FIND $\hat{\underline{\theta}}[0]$ ($p\times 1$) WE NEED AT LEAST $\{x[0], x[1], \ldots, x[p-1]\}$ SO THAT

$$\hat{\underline{\theta}} = \underbrace{(\underline{H}^T \underline{C}^{-1} \underline{H})}_{\substack{\text{MUST BE} \\ \text{INVERTIBLE}}}{}^{-1} \underline{H}^T \underline{C}^{-1} \underline{x}$$

USE BATCH APPROACH

HENCE, WE COULD COMPUTE $\hat{\underline{\theta}}[p-1]$, $\underline{\Sigma}[p-1]$ AND START RECURSION AT $n = p$ OR

ASSUME $\hat{\underline{\theta}}[-1]$, $\underline{\Sigma}[-1]$ AND START AT $n = 0$. THIS WILL BIAS LSE TOWARDS $\hat{\underline{\theta}}[-1]$ UNLESS WE LET $\underline{\Sigma}[-1] = \alpha \underline{I}$ WHERE $\alpha \to \infty$.

IN PRACTICE LET $\hat{\underline{\theta}}[-1] = \underline{0}$, $\underline{\Sigma}[-1] = 10^{10} \underline{I}$. FOR $n \geq p$ BOTH INITIALIZATIONS WILL PRODUCE SAME LSE.

<u>EXAMPLE</u> : FOURIER ANALYSIS SEE EX 8.4

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

f_0 KNOWN, SEQUENTIALLY ESTIMATE $\underline{\theta} = \begin{bmatrix} a \\ b \end{bmatrix}$

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

$$\begin{bmatrix} s[0] \\ s[1] \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \cos 2\pi f_0 & \sin 2\pi f_0 \\ \vdots & \vdots \end{bmatrix} \underline{\theta}$$

ASSUME $x(n) = s(n) + w(n)$

↑ DIAGONAL COVARIANCE MATRIX ⇒ UNCORRELATED

TO INITIALIZE WAIT FOR $p = 2$ DATA SAMPLES $x(0), x(1)$.

$$\hat{\underline{\theta}}(1) = \left(\underline{H}^T(1)\,\underline{C}^{-1}(1)\,\underline{H}(1)\right)^{-1}\underline{H}^T(1)\,\underline{C}^{-1}(1)\,\underline{x}(1)$$

$$\underline{x}(1) = \underset{(n+1)\times 1}{\begin{bmatrix} x(0) \\ x(1) \end{bmatrix}} \qquad \underline{H}(1) = \underset{(n+1)\times 2}{\begin{bmatrix} 1 & 0 \\ \cos 2\pi f_0 & \sin 2\pi f_0 \end{bmatrix}}$$

$$\underline{C}(1) = \begin{bmatrix} \sigma_0^2 & 0 \\ 0 & \sigma_1^2 \end{bmatrix}$$

$$\Rightarrow \qquad \underline{\Sigma}(1) = \left(\underline{H}^T(1)\,\underline{C}^{-1}(1)\,\underline{H}(1)\right)^{-1}$$

NOW LET $n = 2$

$$\underline{K}(2) = \frac{\underline{\Sigma}(1)\,\underline{h}(2)}{\sigma_2^2 + \underline{h}^T(2)\,\underline{\Sigma}(1)\,\underline{h}(2)} \qquad 2\times 1$$

WHERE $\underline{h}^T(2)$ = NEW ROW OF $\underline{H}(2)$ OR

$$\underline{h}^T(2) = \left[\cos 4\pi f_0 \quad \sin 4\pi f_0\right]$$

$$\hat{\underline{\theta}}(2) = \hat{\underline{\theta}}(1) + \underline{K}(2)\left(x(2) - \underline{h}^T(2)\,\hat{\underline{\theta}}(1)\right)$$

$$\underline{\Sigma}(2) = (\underline{I} - \underbrace{\underline{K}(2)}_{2\times 1}\underbrace{\underline{h}^T(2)}_{1\times 2})\ \underline{\Sigma}(1)$$

NEXT FIND $\underline{K}[3]$, ETC. NEED A COMPUTER TO EVALUATE.

NONLINEAR LS

GENERAL PROBLEM : $J = (\underline{x} - \underline{s}(\underline{\theta}))^T(\underline{x} - \underline{s}(\underline{\theta}))$

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

$s[n;\underline{\theta}]$ IS NONLINEAR IN $\underline{\theta}$.

PREVIOUSLY $\underline{s}(\underline{\theta}) = \underline{H}\,\underline{\theta} \Rightarrow$ LINEAR LS

HOW DO WE MINIMIZE LS CRITERION WHEN NONLINEAR?

FIRST CONSIDER METHODS TO CONVERT NONLINEAR PROBLEM TO LINEAR ONE

1) TRANSFORMATION OF PARAMETERS
2) SEPARABILITY OF PARAMETERS

1) TRANSFORM $\underline{\theta}$ TO PRODUCE $\underline{\alpha}$, WHERE SIGNAL IS LINEAR IN $\underline{\alpha}$

$$\underline{\alpha} = \underline{g}(\underline{\theta}) \qquad \underline{g} \text{ IS INVERTIBLE } p\text{-DIMENSIONAL FUNCTION}$$

FIND $\underline{g}$ SO THAT $\underline{s}(\underline{\theta}(\underline{\alpha})) = \underline{s}(\underline{g}^{-1}(\underline{\alpha}))$
$= \underline{H}\,\underline{\alpha}$

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

$$\Rightarrow \quad \hat{\underline{\theta}} = \underline{g}^{-1}(\hat{\underline{\alpha}})$$

MINIMIZATION CAN BE CARRIED OUT IN NEW SPACE AND MINIMIZING VALUE THEN CONVERTED BACK.

EXAMPLE: SINUSOIDAL AMPLITUDE, PHASE ESTIMATION

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

$A > 0$, ESTIMATE A, ϕ

$$J = \sum_n \left(x(n) - A\cos(2\pi f_0 n + \phi)\right)^2$$

$s(n)$ NONLINEAR IN ϕ

BUT $A\cos(2\pi f_0 n + \phi) = A\cos\phi \cos 2\pi f_0 n$
$- A\sin\phi \sin 2\pi f_0 n$

LET $\alpha_1 = A\cos\phi$ OR $\underline{\alpha} = \underline{g}(\underline{\theta})$
$\alpha_2 = -A\sin\phi$

$\Rightarrow s(n) = \alpha_1 \cos 2\pi f_0 n + \alpha_2 \sin 2\pi f_0 n$
LINEAR IN α_1, α_2.

OR $\underline{s} = \underline{H}\underline{\alpha}$

$$= \begin{bmatrix} 1 & 0 \\ \cos 2\pi f_0 & \sin 2\pi f_0 \\ \vdots & \\ \cos 2\pi f_0(N-1) & \sin 2\pi f_0(N-1) \end{bmatrix} \underline{\alpha}$$

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

TO FIND $\hat{\underline{\theta}} = \underline{g}^{-1}(\hat{\underline{\alpha}})$ NEED INVERSE TRANSFORMATION

$$A = \sqrt{\alpha_1^2 + \alpha_2^2}$$
$$\phi = \arctan\left(\frac{-\alpha_2}{\alpha_1}\right)$$

$$\therefore \hat{\underline{\theta}} = \begin{bmatrix} \hat{A} \\ \hat{\phi} \end{bmatrix} = \begin{bmatrix} \sqrt{\hat{\alpha}_1^2 + \hat{\alpha}_2^2} \\ \arctan(-\hat{\alpha}_2/\hat{\alpha}_1) \end{bmatrix}$$

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

2) SEPARABILITY APPROACH - SIGNAL MAY BE LINEAR IN SOME PARAMETERS ⇒ CAN MINIMIZE J OVER THESE

ASSUME $\underline{s} = \underline{H}(\underline{\alpha})\underline{\beta}$

WISH TO ESTIMATE $\underline{\theta} = \begin{bmatrix} \underline{\alpha} \\ \underline{\beta} \end{bmatrix} = \begin{bmatrix} (p-q) \times 1 \\ q \times 1 \end{bmatrix}$

$\underline{H}(\underline{\alpha})$ IS $N \times q$ AND DEPENDS ON $\underline{\alpha}$ ONLY

TO FIND LSE

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

FOR A <u>GIVEN</u> $\underline{\alpha}$ J IS MINIMIZED BY

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

$$\Rightarrow J(\underline{\alpha}, \hat{\underline{\beta}}) = \underline{x}^T \left[\underline{I} - \underline{H}(\underline{\alpha})(\underline{H}^T(\underline{\alpha})\underline{H}(\underline{\alpha}))^{-1}\underline{H}^T(\underline{\alpha}) \right] \underline{x}$$

MUST MAXIMIZE $\underline{x}^T \underline{H}(\underline{\alpha}) (\underline{H}^T(\underline{\alpha}) \underline{H}(\underline{\alpha}))^{-1} \underline{H}^T(\underline{\alpha})\underline{x}$ OVER $\underline{\alpha}$.

<u>EXAMPLE</u> : $s(n) = A_1 r^n + A_2 r^{2n} + A_3 r^{3n}$

WHERE $0 < r < 1$ AND A_1, A_2, A_3, r ARE TO BE ESTIMATED

SIGNAL IS LINEAR IN $\underline{\beta} = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix}$ AND NONLINEAR IN $\underline{\alpha} = (r)$

MUST MAXIMIZE $\underline{x}^T \underline{H}(r) (\underline{H}^T(r) \underline{H}(r))^{-1} \underline{H}^T(r) \underline{x}$ OVER $0 \le r \le 1$

WHERE

$$\underline{H}(r) = \begin{bmatrix} 1 & 1 & 1 \\ r & r^2 & r^3 \\ \vdots & & \\ r^{N-1} & r^{2(N-1)} & r^{3(N-1)} \end{bmatrix}$$

CAN USE A GRID SEARCH $\Rightarrow \hat{r}$

THEN $\hat{\underline{\beta}} = (\underline{H}^T(\hat{r}) \underline{H}(\hat{r}))^{-1} \underline{H}^T(\hat{r}) \underline{x}$

GENERAL NONLINEAR LS

MINIMIZE $J = (\underline{x} - \underline{s}(\underline{\theta}))^T (\underline{x} - \underline{s}(\underline{\theta}))$

TO FIND NECESSARY CONDITIONS

$$J = \sum_n (x[n] - s[n])^2$$

$$\frac{\partial J}{\partial \theta_j} = -2 \sum (x[n] - s[n]) \frac{\partial s[n]}{\partial \theta_j} = 0$$

OR LETTING $\left(\frac{\partial \underline{s}(\underline{\theta})}{\partial \underline{\theta}} \right)_{ij} = \frac{\partial s[i]}{\partial \theta_j}$ $\quad i = 0, 1, \ldots, N-1$ $\quad j = 1, 2, \ldots, p$