$$+ \frac{\partial \underline{\mu}(\underline{\theta})^T}{\partial \theta_k} \underline{C}^{-1}(\underline{\theta})(\underline{x} - \underline{\mu}(\underline{\theta}))$$

$$- \frac{1}{2}(\underline{x} - \underline{\mu}(\underline{\theta}))^T \frac{\partial \underline{C}^{-1}(\underline{\theta})}{\partial \theta_k}(\underline{x} - \underline{\mu}(\underline{\theta})) = 0 \qquad k = 1, 2, \ldots, p$$

PRODUCES NECESSARY CONDITIONS FOR MLE

IMPORTANT EXAMPLE: GENERAL LINEAR MODEL

$$\underline{x} = \underset{N \times p}{\underline{H}} \underline{\theta} + \underset{N(\underline{0}, \underline{C})}{\underline{w}}$$

$$\sim N(\underline{H}\underline{\theta}, \underline{C})$$

$\uparrow$ DOES NOT DEPEND ON $\underline{\theta}$

$$\Rightarrow \frac{\partial \ln p}{\partial \theta_k} = \frac{\partial \underline{\mu}(\underline{\theta})^T}{\partial \theta_k} \underline{C}^{-1}(\underline{x} - \underline{\mu}(\underline{\theta}))$$

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

$$\frac{\partial \ln p}{\partial \underline{\theta}} = \underbrace{\begin{bmatrix} \frac{\partial \underline{H}\underline{\theta}}{\partial \theta_1}^T \\ \vdots \\ \frac{\partial \underline{H}\underline{\theta}}{\partial \theta_p}^T \end{bmatrix}}_{p \times N} \underbrace{\underline{C}^{-1}(\underline{x} - \underline{H}\underline{\theta})}_{N \times 1}$$

$1 \times N$

$$\frac{\partial \underline{H}\underline{\theta}}{\partial \theta_k} = \frac{\partial}{\partial \theta_k} \sum_{i=1}^{p} \underline{h}_i \theta_i \qquad \underline{H} = [\underline{h}_1\ \underline{h}_2 \cdots \underline{h}_p]$$

$$= \underline{h}_k$$

$$\frac{\partial \ln p}{\partial \underline{\theta}} = \begin{bmatrix} \underline{h}_1^T \\ \vdots \\ \underline{h}_p^T \end{bmatrix} \underline{C}^{-1}(\underline{x} - \underline{H}\underline{\theta})$$

$$= \underline{H}^T \underline{C}^{-1} (\underline{x} - \underline{H}\underline{\theta}) = \underline{0}$$

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

MLE IS OPTIMAL FOR GENERAL LINEAR MODEL.
ALSO,

$$\hat{\underline{\theta}} \sim N(\underline{\theta}, (\underline{H}^T \underline{C}^{-1} \underline{H})^{-1})$$

FROM PREVIOUS WORK AND $\underline{I}(\underline{\theta}) = \underline{H}^T \underline{C}^{-1} \underline{H}$

$\Rightarrow$ ASYMPTOTIC PDF OF MLE HOLDS <u>EXACTLY</u> FOR GENERAL LINEAR MODEL

<u>SUMMARY</u>: FOR GENERAL LINEAR MODEL MLE IS

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

WHICH IS KNOWN TO BE EFFICIENT $\Rightarrow$

MLE IS OPTIMAL. ALSO, $\hat{\underline{\theta}} \sim N(\underline{\theta}, (\underline{H}^T \underline{C}^{-1} \underline{H})^{-1})$.

INTERESTING RESULT: IF EFFICIENT ESTIMATOR EXISTS, MLE WILL BE IT. (SEE PROB. 7.12). SAYS THAT MLE CAN BE USED TO FIND MVU ESTIMATOR IF IT IS ALSO EFFICIENT — INCREASES CONFIDENCE IN APPROACH.

NUMERICAL DETERMINATION OF MLE

1) NEWTON - RAPHSON:

HESSIAN $p \times p$ GRADIENT $p \times 1$

$$\underline{\theta}_{k+1} = \underline{\theta}_k - \left[\frac{\partial^2 \ln p}{\partial \underline{\theta} \partial \underline{\theta}^T} \right]^{-1} \frac{\partial \ln p}{\partial \underline{\theta}} \bigg|_{\underline{\theta} = \underline{\theta}_k}$$

WHERE $\left[\frac{\partial^2 \ln p}{\partial \underline{\theta} \partial \underline{\theta}^T} \right]_{ij} = \frac{\partial^2 \ln p}{\partial \theta_i \partial \theta_j}$ $\quad i = 1, 2, \ldots, p$; $j = 1, 2, \ldots, p$

2) SCORING:

$$\underline{\theta}_{k+1} = \underline{\theta}_k + \underline{I}^{-1}(\underline{\theta}) \frac{\partial \ln p}{\partial \underline{\theta}} \bigg|_{\underline{\theta} = \underline{\theta}_k}$$

3) EM - EXPECTATION - MAXIMIZATION

ITERATIVE BUT GUARANTEED TO CONVERGE TO AT LEAST <u>LOCAL</u> MAXIMUM.

EXPLOITS FACT THAT SOME DATA SETS ALLOW MLE TO BE DETERMINED EASIER.

<u>EXAMPLE</u> : $x[n] = \sum_{i=1}^{p} \cos 2\pi f_i n + w[n]$

$\uparrow$ WGN

ESTIMATE FREQUENCIES.

FOR MLE MUST MINIMIZE OVER $\underline{f} = [f_1\ f_2 \ldots f_p]^T$

$$J(\underline{f}) = \sum_{n=0}^{N-1} \left(x[n] - \sum_{i=1}^{p} \cos 2\pi f_i n \right)^2$$

$\Rightarrow$ VERY NONLINEAR, CAN'T DO THIS ANALYTICALLY.

SUPPOSE WE OBSERVED THE p DATA SETS

$$y_i[n] = \cos 2\pi f_i n + w_i[n] \qquad i = 1, 2, \ldots, p \quad n = 0, 1, \ldots, N-1$$

$\uparrow$ WGN WITH VARIANCE σ_i^2

ASSUMING THE $w_i[n]$'S WERE INDEPENDENT, $w_1[n]$ INDEPENDENT OF $w_2[n]$, ETC, WE HAVE A <u>DECOUPLED</u> PROBLEM SINCE

$$p(\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_p) = p(\underline{y}_1)\, p(\underline{y}_2) \ldots p(\underline{y}_p)$$

$\uparrow$ DEPENDS ON f_1 ONLY

$\Rightarrow$ TO FIND MLE OF f_i MINIMIZE

$$J(f_i) = \sum_{n=0}^{N-1} \left(y_i[n] - \cos 2\pi f_i n \right)^2$$

FOR EACH i.

ORIGINAL p-DIMENSIONAL MINIMIZATION $J(\underline{f})$ REDUCED TO p 1-DIMENSIONAL MINIMIZATIONS $J(f_i)$ $i = 1, 2, \ldots, p$.

$\{ y_1[n], y_2[n], \ldots, y_p[n] \}$ TERMED THE <u>COMPLETE</u> DATA SINCE

$$x[n] = \sum_{i=1}^{p} y_i[n]$$

IF $\quad w[n] = \sum_{i=1}^{p} w_i[n]$

$x[n]$ TERMED <u>INCOMPLETE</u> DATA.

NOTE THAT TO INSURE $w[n] = \sum_{i=1}^{p} w_i[n]$ TO BE WGN WITH VARIANCE σ^2 WE ASSUME $w_i[n]$'s ARE INDEPENDENT AND

$$\sigma^2 = \sum_{i=1}^{p} \sigma_i^2$$

DECOMPOSITION IS NOT UNIQUE SINCE IF

$$y_1[n] = \sum_{i=1}^{p} \cos 2\pi f_i n$$

$$y_2[n] = w[n]$$

$$\Rightarrow x[n] = y_1[n] + y_2[n]$$

KEY IS TO CHOOSE A DECOMPOSITION TO SIMPLIFY MAXIMIZATION PROBLEM.

GENERAL APPROACH:

ASSUME WE CAN HYPOTHESIZE DATA $\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_M$ (PREVIOUS EXAMPLE $M=p$ AND $\underline{y}_i$ IS $N \times 1$)

WE OBSERVE

$$\underline{x} = \underline{g}(\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_M) = \underline{g}(\underline{y})$$

$\uparrow$ MANY-TO-ONE TRANSFORMATION

WISH TO MAXIMIZE $\ln p(\underline{y}; \underline{\theta})$ BUT $\underline{y}$ IS UNAVAILABLE $\Rightarrow$ USE AVERAGE VALUE

$$E_{y|x}[\ln p(\underline{y};\underline{\theta})] = \int \ln p(\underline{y};\underline{\theta})\, p(\underline{y}|\underline{x};\underline{\theta})\, d\underline{y}$$

THIS INTEGRATES OUT $\underline{y}$. STILL NEED

$p(\underline{y} \mid \underline{x}; \underline{\theta})$, BUT WE DON'T KNOW $\underline{\theta}$

$\Rightarrow$ USE CURRENT GUESS, $\underline{\theta_k}$.

EXPECTATION (E): COMPUTE AVERAGE LOG-LIKELIHOOD OF COMPLETE DATA

$$U(\underline{\theta}, \underline{\theta_k}) = \int \ln p(\underline{y}; \underline{\theta})\, p(\underline{y} \mid \underline{x}; \underline{\theta_k})\, d\underline{y}$$

MAXIMIZATION (M): MAXIMIZE AVERAGE LOG-LIKELIHOOD OF COMPLETE DATA

$$\underline{\theta}_{k+1} = \arg\max_{\underline{\theta}} U(\underline{\theta}, \underline{\theta}_k)$$

TERMED EM ALGORITHM - ITERATE UNTIL CONVERGENCE

<u>EXAMPLE</u>: PREVIOUS ONE

$$x[n] = \sum_{i=1}^{p} \cos 2\pi f_i n + \underbrace{w[n]}_{\uparrow \text{WGN}}$$

$$y_i[n] = \cos 2\pi f_i n + \underbrace{w_i[n]}_{\uparrow \text{WGN}}$$

WHERE $w_i[n]$'S ARE INDEPENDENT WITH VARIANCE σ_i^2

$$\ln p_y(\underline{y}; \underline{\theta}) = \ln \prod_{i=1}^{p} p(\underline{y}_i; \underline{\theta})$$ $\underline{y}_i$'S INDEPENDENT

$$= \sum_{i=1}^{p} \ln p(\underline{y}_i;\underline{\theta})$$

$$= \sum_{i=1}^{p} \ln \left\{ \frac{1}{(2\pi\sigma_i^2)^{N/2}} e^{-\frac{1}{2\sigma_i^2}\sum_{n=0}^{N-1}(y_i[n]-\cos 2\pi f_i n)^2} \right\}$$

$$= \underset{\uparrow}{C_1} - \sum_{i=1}^{p} \frac{1}{2\sigma_i^2} \sum_n (y_i[n] - \cos 2\pi f_i n)^2$$

DOES NOT DEPEND ON $\underline{\theta}$

$$= C_1 - \sum_i \frac{1}{2\sigma_i^2} \sum_n (y_i^2[n] - 2y_i[n]\cos 2\pi f_i n + \cos^2 2\pi f_i n)$$

$$= C_1 - \underbrace{\sum_i \frac{1}{2\sigma_i^2} \sum_n y_i^2[n]}_{\text{DOESN'T DEPEND ON } \underline{\theta}}$$

$$+ \sum_i \frac{1}{2\sigma_i^2} \sum_n (2y_i[n]\cos 2\pi f_i n - \cos^2 2\pi f_i n)$$

$$= C_2 + \sum_i \frac{1}{\sigma_i^2} \sum_n (y_i[n]\cos 2\pi f_i n - \tfrac{1}{2}\cos^2 2\pi f_i n)$$

BUT $\sum_n \cos^2 2\pi f_i n \simeq N/2$

$$= C_3 + \sum_{i=1}^{p} \frac{1}{\sigma_i^2} \sum_{n=0}^{N-1} y_i[n]\cos 2\pi f_i n$$

LET $\underline{y}_i = [\, y_i[0] \ \dots \ y_i[N-1]\,]^T$

$$\underline{c}_i = [\, 1 \ \cos 2\pi f_i \ \dots \ \cos 2\pi f_i (N-1)\,]^T$$

$$\ln p_y(\underline{y};\underline{\theta}) = C_3 + \sum_{i=1}^{p} \frac{1}{\sigma_i^2} \underline{c}_i^T \underline{y}_i$$

LET $\quad \underline{c} = \begin{bmatrix} c_1/\sigma_1^2 \\ c_2/\sigma_2^2 \\ \vdots \\ c_p/\sigma_p^2 \end{bmatrix} \qquad \underline{y} = \begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \vdots \\ \underline{y}_p \end{bmatrix}$

DROP c_3 CONSTANT SINCE IT WON'T AFFECT MAXIMIZATION

$$\ln p_Y(\underline{y};\underline{\theta}) = \underline{c}^T \underline{y}$$

TO FIND EXPECTATION (E STEP):

$$U(\underline{\theta}, \underline{\theta}_k) = E\left[\ln p_Y(\underline{y};\underline{\theta}) \mid \underline{x};\underline{\theta}_k\right]$$

$$= E\left[\underline{c}^T \underline{y} \mid \underline{x};\underline{\theta}_k\right]$$

$$= \underline{c}^T E\left[\underline{y} \mid \underline{x};\underline{\theta}_k\right]$$

BUT $\underline{x}$ AND $\underline{y}$ ARE JOINTLY GAUSSIAN SINCE

$$x[n] = \sum_{i=1}^{p} y_i[n]$$

$$\underline{x} = \begin{bmatrix} \sum_{i=1}^{p} y_i[0] \\ \vdots \\ \sum_{i=1}^{p} y_i[N-1] \end{bmatrix} = \underbrace{\begin{array}{ccc} \overbrace{\qquad}^{1} & \overbrace{\qquad}^{2} & \overbrace{\qquad}^{p} \\ \multicolumn{3}{c}{\begin{bmatrix} 10\ldots0 & 10\ldots0\ldots & 10\ldots0 \\ 01\ldots0 & 01\ldots0\ldots & 01\ldots0 \\ & & \\ 00\ldots1 & 00\ldots1\ldots & 00\ldots1 \end{bmatrix}} \end{array}}_{[\underline{I}\ \underline{I}\ \ldots\ \underline{I}]} \overbrace{\begin{bmatrix} y_1[0] \\ \vdots \\ y_1[N-1] \\ y_2[0] \\ \vdots \\ y_2[N-1] \\ \vdots \\ y_p[0] \\ \vdots \end{bmatrix}}^{\underline{y}}$$

$$\underline{x} = [\overbrace{\underline{I}\ \underline{I} \ldots \underline{I}}^{p}]\ \underline{y}$$

$\uparrow$ $N \times N$

RESULT: IF $\underline{z} = \begin{bmatrix} \underline{x} \\ \underline{y} \end{bmatrix}$ IS A GAUSSIAN RANDOM VECTOR ($\underline{x}$, $\underline{y}$ ARE JOINTLY GAUSSIAN), THEN

SEE APPENDIX 10A

$E[(\underline{y}-E(\underline{y}))(\underline{x}-E(\underline{x}))^T]$

$$E(\underline{y}/\underline{x}) = E(\underline{y}) + \underline{C}_{yx}\,\underline{C}_x^{-1}(\underline{x} - E(\underline{x})).$$

BUT $\underline{y} = \begin{bmatrix} \underline{y}_1 \\ \vdots \\ \underline{y}_p \end{bmatrix} \Rightarrow E(\underline{y}) = \begin{bmatrix} \underline{c}_1 \\ \vdots \\ \underline{c}_p \end{bmatrix}$

$$E(\underline{x}) = E\left[[\underline{I}\ \underline{I} \ldots \underline{I}] \begin{bmatrix} \underline{y}_1 \\ \vdots \\ \underline{y}_p \end{bmatrix} \right]$$

$$= \sum_{i=1}^{p} E(\underline{y}_i) = \sum_{i=1}^{p} \underline{c}_i$$

TO FIND COVARIANCES:

$$\underline{C}_x = \sigma^2 \underline{I}$$

$$\underline{C}_{yx} = E[(\underline{y} - E(\underline{y}))(\underline{x} - E(\underline{x}))^T]$$

$$= E\left[\begin{bmatrix} \underline{w}_1 \\ \vdots \\ \underline{w}_p \end{bmatrix} \underline{w}^T \right]$$

$$= E\left[\begin{bmatrix} \underline{w}_1 \\ \vdots \\ \underline{w}_p \end{bmatrix} \left[[\underline{I}\ \underline{I} \ldots \underline{I}] \begin{bmatrix} \underline{w}_1 \\ \vdots \\ \underline{w}_p \end{bmatrix} \right]^T \right]$$

$$= E\left[\begin{bmatrix}\underline{w}_1\\ \vdots\\ \underline{w}_p\end{bmatrix}\begin{bmatrix}\underline{w}_1^T \cdots \underline{w}_p^T\end{bmatrix}\begin{bmatrix}\underline{I}\\ \vdots\\ \underline{I}\end{bmatrix}\right]$$

$$= \begin{bmatrix}E(\underline{w}_1\underline{w}_1^T)\cdots E(\underline{w}_1\underline{w}_p^T)\\ \ddots \\ E(\underline{w}_p\underline{w}_p^T)\end{bmatrix}\begin{bmatrix}\underline{I}\\ \vdots\\ \underline{I}\end{bmatrix}$$

$$= \begin{bmatrix}\sigma_1^2\underline{I} & \underline{0} & \cdots & \underline{0}\\ \underline{0} & \sigma_2^2\underline{I} & \cdots & \underline{0}\\ & & \ddots & \\ \underline{0} & \underline{0} & \cdots & \sigma_p^2\underline{I}\end{bmatrix}\begin{bmatrix}\underline{I}\\ \underline{I}\\ \vdots\\ \underline{I}\end{bmatrix} = \begin{bmatrix}\sigma_1^2\underline{I}\\ \sigma_2^2\underline{I}\\ \vdots\\ \sigma_p^2\underline{I}\end{bmatrix}$$

$$E(\underline{y}\,|\,\underline{x};\underline{\theta_k}) = \begin{bmatrix}\underline{c}_1\\ \vdots\\ \underline{c}_p\end{bmatrix} + \frac{1}{\sigma^2}\begin{bmatrix}\sigma_1^2\underline{I}\\ \vdots\\ \sigma_p^2\underline{I}\end{bmatrix}\left(\underline{x} - \sum_{i=1}^{p}\underline{c}_i\right)$$

$$= \begin{bmatrix}\underline{c}_1\\ \vdots\\ \underline{c}_p\end{bmatrix} + \begin{bmatrix}\sigma_1^2/\sigma^2\,(\underline{x} - \Sigma\,\underline{c}_i)\\ \vdots\\ \sigma_p^2/\sigma^2\,(\underline{x} - \Sigma\,\underline{c}_i)\end{bmatrix}$$

BASED ON $\underline{\theta_k}$

$$\Rightarrow E(\underline{y}_i\,|\,\underline{x};\underline{\theta_k}) = \underline{c}_i + \frac{\sigma_i^2}{\sigma^2}\left(\underline{x} - \sum_{i=1}^{p}\underline{c}_i\right)$$

$i = 1, 2, \ldots, p$

THIS MAY BE VIEWED AS AN ESTIMATOR OF $\underline{y}_i$ SO THAT

$$\hat{y}_i[n] = \cos 2\pi f_{i_k} n + \underbrace{\frac{\sigma_i^2}{\sigma^2}}_{\beta_i}\left(x[n] - \sum_{i=1}^{p}\cos 2\pi f_{i_k} n\right)$$

BUT $U(\underline{\theta}, \underline{\theta}_k) = \sum_{i=1}^{p} \frac{1}{\sigma_i^2} \underline{c}_i^T E(\underline{y}_i | \underline{x}; \underline{\theta}_k)$

$\Rightarrow$ MUST MAXIMIZE OVER $\underline{\theta}$

$$U(\underline{\theta}, \underline{\theta}_k) = \sum_{i=1}^{p} \frac{1}{\sigma_i^2} \underline{c}_i^T \hat{\underline{y}}_i$$

DEPENDS ON $\underline{\theta}$ DEPENDS ON DATA, $\underline{\theta}_k$

BUT $\underline{c}_i$ DEPENDS ON f_i ONLY

$$\Rightarrow f_{i_{k+1}} = \underset{f_i}{\text{ARG MAX}}\ \underline{c}_i^T \hat{\underline{y}}_i$$

SUMMARY: FOR $i = 1, 2, \ldots, p$

E STEP: $$\hat{y}_i[n] = \cos 2\pi f_{i_k} n + \beta_i \overbrace{\left(x[n] - \sum_{i=1}^{p} \cos 2\pi f_{i_k} n\right)}^{\hat{w}[n]}$$

M STEP: $$f_{i_{k+1}} = \underset{f_i}{\text{ARG MAX}} \sum_{n=0}^{N-1} \hat{y}_i[n] \cos 2\pi f_i n$$

WHERE $\sum_{i=1}^{p} \beta_i = \sum_{i=1}^{p} \sigma_i^2/\sigma^2 = 1$

EM ALGORITHM ATTEMPTS TO SPLIT $x[n]$ DATA INTO p SETS WITH EACH ONE CONSISTING OF ONE SINUSOID AND NOISE. M STEP IS MLE FOR ONE SINUSOID AND WGN.

EM ALGORITHM INCREASES LIKELIHOOD AT EACH ITERATION $\Rightarrow$ CONVERGES TO AT LEAST LOCAL MAXIMUM.

ASYMPTOTIC MLE

CONSIDER DATA $\underline{x} \sim N(\underline{0}, \underline{C}(\underline{\theta}))$. MUST MAXIMIZE

$$p(\underline{x};\underline{\theta}) = \frac{1}{(2\pi)^{N/2} |\underline{C}(\underline{\theta})|^{\frac{1}{2}}} e^{-\frac{1}{2}\underline{x}^T \underline{C}^{-1}(\underline{\theta})\underline{x}}$$

OVER $\underline{\theta}$ TO FIND MLE (SEE (7.40) FOR DERIVATIVES). NEED TO FIND $\underline{C}^{-1}$ IN CLOSED FORM.

ALTERNATE APPROACH: IF $x[n]$ IS WSS THEN $\underline{C}(\underline{\theta})$ IS TOEPLITZ, SO THAT AS $N \to \infty$ WE CAN USE APPROXIMATE LOG-PDF

ASIDE

$$\underline{x} = [x[0]\ x[1] \ldots x[N-1]]^T$$

IF $x[n]$ IS ZERO MEAN AND WSS SO THAT

$$E[x[n]x[n+k]] = r_{xx}[k], \text{ THEN}$$

$$\underline{C} = E[\underline{x}\,\underline{x}^T]$$

$$= E\begin{bmatrix} x[0] \\ x[1] \\ \vdots \\ x[N-1] \end{bmatrix} \begin{bmatrix} x[0] & x[1] & \ldots & x[N-1] \end{bmatrix}$$

$$= \begin{bmatrix} E(x^2[0]) & E(x[0]x[1]) & \ldots & E(x[0]x[N-1]) \\ E(x[1]x[0]) & E(x^2[1]) & \ldots & E(x[1]x[N-1]) \\ & & \ddots & \\ E[x[N-1]x[0]] & E[x[N-1]x[1]] & \ldots & E(x^2[N-1]) \end{bmatrix}$$

$$= \begin{bmatrix} r_{xx}[0] & r_{xx}[1] & \ldots & r_{xx}[N-1] \\ r_{xx}[1] & r_{xx}[0] & \ldots & r_{xx}[N-2] \\ & & \ddots & \\ r_{xx}[N-1] & r_{xx}[N-2] & & r_{xx}[0] \end{bmatrix}$$

$N = 4$

$$\begin{bmatrix} r_{xx}[0] & r_{xx}[1] & r_{xx}[2] & r_{xx}[3] \\ r_{xx}[1] & r_{xx}[0] & r_{xx}[1] & r_{xx}[2] \\ r_{xx}[2] & r_{xx}[1] & r_{xx}[0] & r_{xx}[1] \\ r_{xx}[3] & r_{xx}[2] & r_{xx}[1] & r_{xx}[0] \end{bmatrix}$$

SAME ELEMENT ALONG EACH NW-SE DIAGONAL

CALLED <u>TOEPLITZ</u> MATRIX

SINCE $\underline{C}(\underline{\theta})$ IS TOEPLITZ, WE CAN WRITE AS $N \to \infty$ (SEE APPENDIX 3D)

$$\ln p(\underline{x};\underline{\theta}) = -\frac{N}{2}\ln 2\pi - \frac{N}{2}\int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\ln P_{xx}(f) + \frac{I(f)}{P_{xx}(f)} \right) df$$

$P_{xx}(f)$ = POWER SPECTRAL DENSITY (DEPENDS ON $\underline{\theta}$)

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

= PERIODOGRAM OF DATA

$$= \frac{1}{N} \left| \mathcal{F}\{x[n]\} \right|^2$$

TO FIND NECESSARY CONDITIONS FOR MLE

$$\frac{\partial \ln p}{\partial \theta_i} = -\frac{N}{2} \int \left(\frac{1}{P_{xx}(f)} - \frac{I(f)}{P_{xx}^2(f)} \right) \frac{\partial P_{xx}(f)}{\partial \theta_i} df = 0$$

$$\Rightarrow \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\frac{1}{P_{xx}(f)} - \frac{I(f)}{P_{xx}^2(f)} \right) \frac{\partial P_{xx}(f)}{\partial \theta_i} df = 0$$

CAN ALSO USE WITH NEWTON-RAPHSON TO FIND DERIVATIVES EASIER

EXAMPLE: MOVING AVERAGE PROCESS - ORDER 2

MINIMUM PHASE

$$\text{WHITE GAUSSIAN NOISE } \sigma^2 = 1 \longrightarrow \boxed{B(z)} \longrightarrow x[n]$$

$$B(z) = 1 + b[1] z^{-1} + b[2] z^{-2}$$

$$P_{xx}(f) = |B(e^{j2\pi f})|^2 \sigma^2$$

$$= |1 + b[1] e^{-j2\pi f} + b[2] e^{-j4\pi f}|^2$$

WANT MLE OF $b[1], b[2]$ (MA FILTER PARAMETERS)

CAN SHOW THAT

$$r_{xx}[k] = \begin{cases} 1 + b^2[1] + b^2[2] & k = 0 \\ b[1] + b[1]\, b[2] & k = 1 \\ b[2] & k = 2 \\ 0 & k > 2 \end{cases}$$

$$\underline{C} = \begin{bmatrix} 1+b^2[1]+b^2[2] & b[1]+b[1]b[2] & b[2] & 0 & \cdots & 0 \\ & \ddots & \ddots & \ddots & & \\ & & & & & \end{bmatrix}$$

TRY INVERTING $\underline{C}$. CAN BE DONE BUT IS MESSY. NEED SEARCH TECHNIQUE.

SINCE $B(z)$ IS MINIMUM-PHASE (ZEROS WITHIN UNIT CIRCLE OR $|z_1|<1$, $|z_2|<1$), $b[1]$, $b[2]$ ARE CONSTRAINED TO FINITE SUBSET OF PLANE. $\Rightarrow$ GRID SEARCH BEST METHOD.

$$1 + b[1]z^{-1} + b[2]z^{-2} = (1 - z_1 z^{-1})(1 - z_2 z^{-1})$$

$$\Rightarrow \quad b[1] = -(z_1 + z_2)$$
$$b[2] = z_1 z_2$$

CAN SEARCH OVER $|z_1|<1$, $|z_2|<1$ (NOTE: IF ZEROS COMPLEX, $z_1 = re^{j\theta}$, $z_2 = re^{-j\theta}$, IF REAL, $z_1 = r_1$, $z_2 = r_2$) AND USE INVARIANCE PRINCIPLE TO FIND MLE OF $b[1]$, $b[2]$.

$$\ln p(\underline{x};\underline{\theta}) = -\frac{N}{2}\ln 2\pi - \frac{N}{2}\int\left(\ln P_{xx}(f) + \frac{I(f)}{P_{xx}(f)}\right)df$$

BUT $\int \ln P_{xx}(f)\,df = 0$ (SEE PROB 7.23)

MUST MINIMIZE

$$\int \frac{I(f)}{P_{XX}(f)}\, df$$

$$= \int \frac{I(f)}{|1 + b[1] e^{-j2\pi f} + b[2] e^{-j4\pi f}|^2}\, df$$

$$= \int \frac{I(f)}{|1 - z_1 e^{-j2\pi f}|^2 \, |1 - z_2 e^{-j2\pi f}|^2}\, df$$

OVER z_1, z_2.

SIGNAL PROCESSING EXAMPLES

RANGE ESTIMATION - SEE EXAMPLE 3.13

NEED TO MAXIMIZE OVER n_0

$$p(\underline{x}; n_0) = \prod_{n=0}^{n_0-1} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2} x^2[n]}$$

$$\cdot \prod_{n=n_0}^{n_0+M-1} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x[n] - s[n-n_0])^2}$$

$$\cdot \prod_{n=n_0+M}^{N-1} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2} x^2[n]}$$

WHERE $t_0 = n\Delta$

$$p(\underline{x}; n_0) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1} x^2[n]}$$

$$\underbrace{\cdot \prod_{n=n_0}^{n_0+M-1} e^{-\frac{1}{2\sigma^2}\left(-2x[n]s[n-n_0]+s^2[n-n_0]\right)}}_{\text{MAXIMIZE OVER } n_0}$$

$$= e^{-\frac{1}{2\sigma^2}\sum_{n=n_0}^{n_0+M-1}\left(-2x[n]s[n-n_0]+s^2[n-n_0]\right)}$$

MUST MINIMIZE $\sum_{n=n_0}^{n_0+M-1}\left(-2x[n]s[n-n_0]+s^2[n-n_0]\right)$

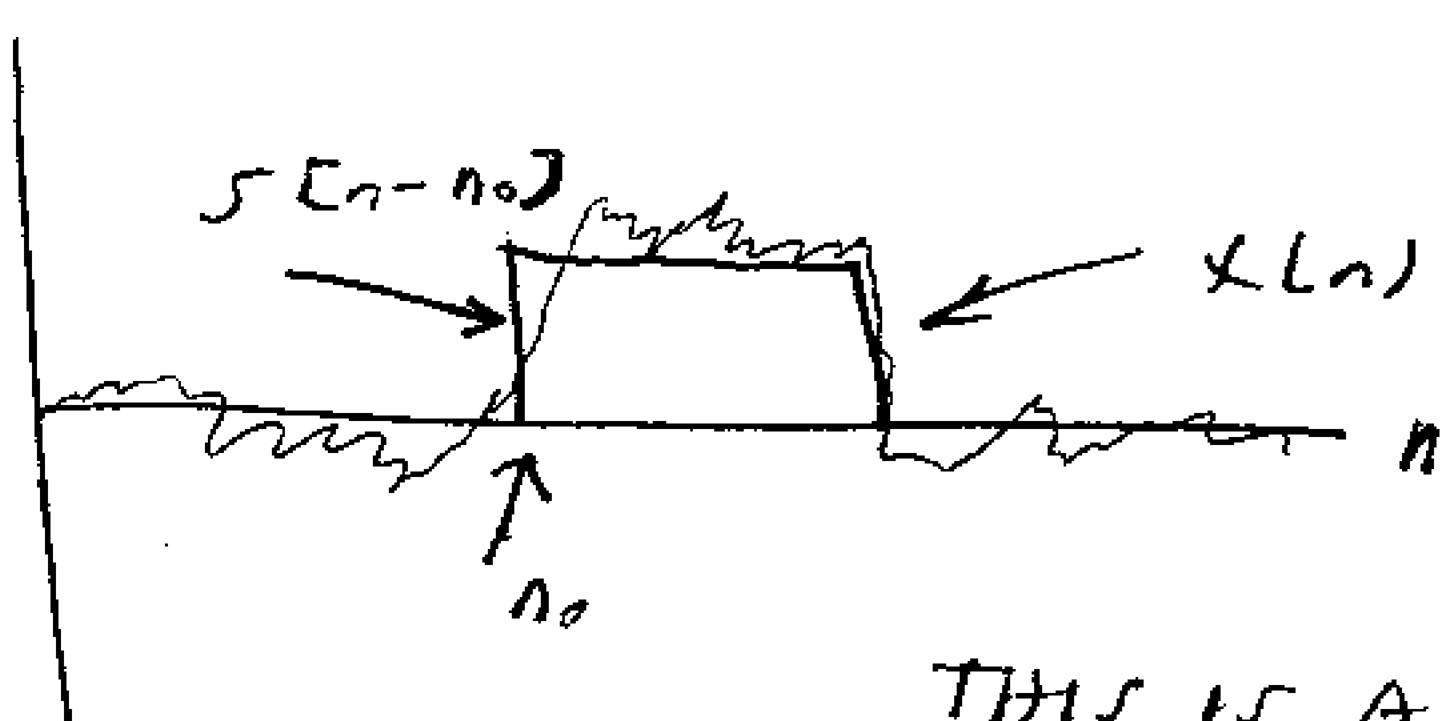
BUT $\sum_{n=n_0}^{n_0+M-1} s^2[n-n_0] = \sum_{n=0}^{M-1} s^2[n] =$ ENERGY

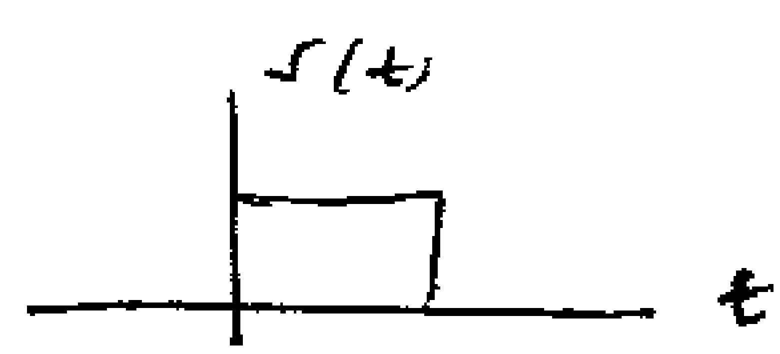
AND DOES NOT DEPEND ON n_0.

$\Rightarrow$ MINIMIZE $\sum_{n=n_0}^{n_0+M-1} -2x[n]s[n-n_0]$

OR MAXIMIZE

$$\sum_{n=n_0}^{n_0+M-1} x[n]s[n-n_0]$$





THIS IS A CORRELATOR

SINCE $R = c\,t_0/2 = c\,n_0\Delta/2$

$$\Rightarrow \hat{R} = \left(\frac{c\Delta}{2}\right)\hat{n}_0$$

EXAMPLE : SINUSOIDAL PARAMETER ESTIMATION (SEE EXAMPLE 3.14)

ESTIMATE A, f_0, ϕ FOR

$$x(n) = A\cos(2\pi f_0 n + \phi) + w(n)$$

↑ WGN

$$p(\underline{x};\underline{\theta}) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x(n) - A\cos(2\pi f_0 n + \phi))^2}$$

MLE FOUND BY MINIMIZING

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

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

$\Rightarrow$ NONQUADRATIC IN f_0, ϕ

USE TRANSFORMATION (1-1 TRANSFORMATION)

$$\alpha_1 = A\cos\phi$$
$$\alpha_2 = -A\sin\phi$$