PARAMETRIC MODELING (1) (time series)

BASIC IDEA: Approximate deterministic processes, or stochastic processes by a finite parametter model

- Select a model
- Estimate the parameters of the assumed model using the available data samples
- Obtain power spectral density (PSD) estimate using the estimated parameters in the assumed model

PARAMETRIC MODELING (2)

NOTES:

- Accuracy of model is of outmost importance. Inaccurate models result in errors in the estimated power spectral density

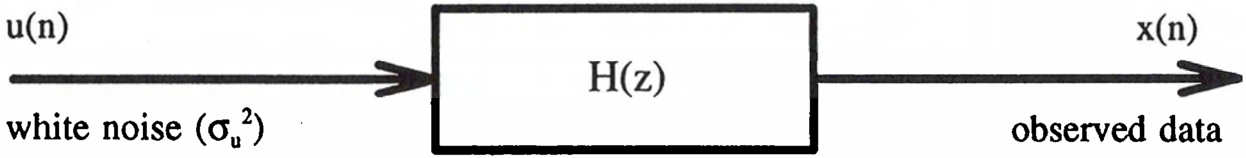
- Only the output process is available to analysis (Important difference from system identification)

- More realistic assumptions are made about data outside observed interval (i.e., data are not considered to be zero or periodic outside the observed interval). Windowing of the data is avoided

- Better performance than classical methods, especially for short data records

PARAMETRIC MODELING (3)

- **Random WSS processes**



Linear difference equation:

$$x(n) = -\sum_{k=1}^{P} a_k . a(n-k) + \sum_{k=0}^{q} b_k . u(n-k)\,, \quad a_0 = b_0 = 1$$

Transfer function:

$$H(z) = \frac{B(z)}{A(z)} = \frac{1 + \sum_{k=1}^{q} b_k . z^{-k}}{1 + \sum_{k=1}^{p} a_k . z^{-k}} = \sum_{k} h_k z^{-k}$$

$$z = e^{j2\pi f}\,, \qquad 0 \le f \le 1$$

PARAMETRIC MODELING (4)

- **POWER SPECTRUM**

$$P_x(z) = P_u(z) \,.\, H(z) \,.\, H^*\left(\frac{1}{z^*}\right) = \frac{B(z)\, B^*\left(\frac{1}{z^*}\right)}{A(z)\, A^*\left(\frac{1}{z^*}\right)} \,.\, P_u(z)$$

or, $(z = e^{j2\pi f})$

$$P_x(f) = P_u(f) \,.\, H(f) \,.\, H^*(f) = \sigma_u^2 \,.\, |\, H(f)\, |^2$$

ARMA (p,q): $H(f) = \dfrac{B(f)}{A(f)}$, pole-zero model

$$P_x(f) = \sigma_u^2 \,.\, \frac{|\, B(f)\, |^2}{|\, A(f)\, |^2} = \sigma_u^2 \, \frac{\left|\, 1 + \sum_{k=1}^{q} b_k \,.\, e^{-j2\pi fk} \,\right|^2}{\left|\, 1 + \sum_{k=1}^{p} a_k \,.\, e^{-j2\pi fk} \,\right|^2}$$

PARAMETRIC MODELING (5)

- **AR (p):** $H(f) = \dfrac{1}{A(f)}$, Autoregressive (all pole) model

$$P_x(f) = \sigma_u^2 \frac{1}{|A(f)|^2} = \sigma_u^2 \frac{1}{|1 + \sum_{k=1}^{p} a_k . e^{-2j\pi fk}|^2}$$

- **MA (q):** $H(f) = B(f)$, Moving average (all zero) model

$$P_x(f) = \sigma_u^2 . |B(f)|^2 = \sigma_u^2 . |1 + \sum_{k=1}^{q} b_k e^{-j2\pi fk}|^2$$

PARAMETRIC MODELING (6)

WORLD DECOMPOSITION THEOREM (1938)

Any stationary (WSS) ARMA or AR process of finite variance can be represented by a unique MA model of generally infinite order.

Equivalently (Kolmogorov, 1941)

Any ARMA or MA process can be represented by a unique AR model of generally infinite order.

i.e.,

$$\textbf{ARMA (p,q)} \approx \textbf{AR} (\infty)$$

$$\textbf{ARMA (p,q)} \approx \textbf{MA} (\infty)$$

Conclusion: Assuming the wrong model has been chosen, a high enough model can give reasonable approximations.

PARAMETRIC MODELING (7)

- Let, $C(z) = \sum_{k=0}^{\infty} c_k z^{-k}$, $\quad c_0 = 1$. $\quad$ Then,

$$\frac{B(z)}{A(z)} = \frac{1}{C(z)} \Rightarrow B(z)\,C(z) = A(z) \qquad (\star\star)$$

1) ARMA (p,q) $\rightarrow$ AR (∞)

$$c_n = \begin{cases} 1 & , n=0 \\ -\sum_{k=1}^{q} b_k c_{n-k} + a_n & , 1 \le n \le p \\ -\sum_{k=1}^{q} b_k c_{n-k} & , n > p \end{cases} \qquad (\star)$$

$$c_{-1} = \ldots = c_{-q} = 0$$

PARAMETRIC MODELING (8)

2) AR (∞) $\rightarrow$ ARMA (p,q)

- MA coefficients of ARMA (p,q) (repeat $\bigstar$ for $n=p+1,\ldots,p+q$)

$$\begin{bmatrix} c_p & c_{p-1} & \cdots & c_{p-q+1} \\ c_{p+1} & c_p & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ c_{p+q-1} & c_{p+q-2} & \cdots & c_p \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_q \end{bmatrix} = - \begin{bmatrix} c_{p+1} \\ c_{p+2} \\ \cdot \\ \cdot \\ c_{p+q} \end{bmatrix}$$

(Easily solved by Levinson Recursion)

- AR coefficients of ARMA (p,q) (from $\bigstar$, $\bigstar\bigstar$)

$$a_n = c_n + \sum_{k=1}^{q} b_k c_{n-k}, \qquad 1 \le n \le p$$

PARAMETRIC MODELING (9)

- Let $F(z) = \sum_{k=0}^{\infty} f_k z^{-k}$, $f_0 = 1$. Then,

$$\frac{B(z)}{A(z)} = F(z) \;\Rightarrow\; B(z) = F(z)A(z) \qquad (\star\star)$$

1) ARMA (p,q) $\rightarrow$ MA (∞)

$$f_n = \begin{cases} 1 & , n=0 \\ -\sum_{k=1}^{p} a_k f_{n-k} + b_n & , 1 \le n \le q \\ -\sum_{k=1}^{q} a_k f_{n-k} & , n > q \end{cases} \qquad (\star)$$

PARAMETRIC MODELING (10)

2) MA (∞) $\rightarrow$ ARMA (p,q)

- AR coefficients of ARMA (p,q) (repeat ★ for $n=q+1,\ldots,q+p$)

$$\begin{bmatrix} f_q & f_{q-1} & \cdot\ \cdot\ \cdot & f_{q-p+1} \\ f_{q+1} & f_q & \cdot\ \cdot\ \cdot & \cdot \\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot \\ f_{p+q-1} & f_{q+p-2} & \cdot\ \cdot\ \cdot & f_p \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ a_p \end{bmatrix} = - \begin{bmatrix} f_{q+1} \\ f_{q+2} \\ \cdot \\ \cdot \\ f_{q+p} \end{bmatrix}$$

- MA coefficients of ARMA (p,q) (from ★, ★★)

$$b_n = f_n + \sum_{k=1}^{p} a_k f_{n-k}, \qquad 1 \le n \le q$$

Relation of ARMA parameters - Autocorrelation.

Let $x(n) = -\sum_{k=1}^{p} a_k \, . \, x(n-k) + \sum_{k=0}^{q} b_k u(n-k)$

1) By multiplying both sides by $x^*(n-m)$ and then taking expectation
2) By assumming that {u(n)} is white noise (σ_u^2) and

$R_x(m) = E\{x(n)\, x^*(n-m)\}$

$$R_x(m) = \begin{cases} R_x^*(-m) & , m<0 \\ -\sum_{k=1}^{p} a_k R_x(m-k) + \sigma_u^2 \, . \sum_{k=m}^{q} b_k h_{k-m}^* & , 0 \le m < q \\ -\sum_{k=1}^{p} a_k R_x(m-k) & , m>q \end{cases}$$

- For m=q+1,...,q+p → Modified Yule-Walker equations

$$\begin{bmatrix} R_x(q) & . \; . & R_x(q-p+1) \\ . & . & . \\ R_x(q+p-1) & . \; . & R_x(q) \end{bmatrix} \begin{bmatrix} a_1 \\ . \\ a_p \end{bmatrix} = - \begin{bmatrix} R_x(q+1) \\ . \\ R_x(q+p) \end{bmatrix}$$

- Nonlinear relation for $\{b_k\}$, $k=1,\ldots,q$

PARAMETRIC MODELING (12)

- **Relation of AR(p) parameters - Autocorrelation**

Let $x(n) = -\sum_{k=1}^{p} a_k . x(n-k) + u(n)$

Following a similar procedure, obtain Yule-walker or Normal Equations

$$\begin{bmatrix} R_x(0) & R_x(-1) & \cdots & R_x(-p) \\ R_x(1) & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ R_x(p) & R_x(p-1) & \cdots & R_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \cdot \\ a_p \end{bmatrix} = \begin{bmatrix} \sigma_u^2 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

- Use Levinson Recursion
- First p autocorrelation lags are sufficient to determine parameters of AR(p) model
- $P_x^{AR}(f) = \dfrac{\sigma_u^2}{|1 + \sum_{k=1}^{p} a_k . e^{-2j\pi fk}|^2} = \underbrace{\sum_{m=-\infty}^{\infty} R_x(m) e^{-j2\pi fm}}_{}$

$P_x^{BT}(f)$ with infinite autocorrelation lags

PARAMETRIC MODELING (13)

- **Relation of MA(q) parameters - Autocorrelation**

Let $x(n) = \sum_{k=0}^{q} b_k \cdot u(n-k)$, $(b_k \equiv h_k)$

$$R_x(m) = \sigma_u^2 \cdot \sum_{k=m}^{q} b_k b_{k-m}^{*} = \sigma_u^2 \sum_{k=m}^{q} h_k h_{k-m}^{*} \quad , \quad 0 \le m \le q$$

- Nonlinear (convolution) relation for $\{b_k\}$, $k=0,\dots,q$

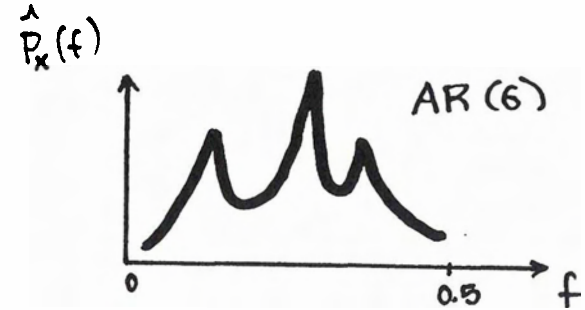
- $R_x(m)=0$, $|m|>q$

- $P_x^{MA}(f) = \sigma_u^2 \left| 1 + \sum_{k=1}^{q} b_k \cdot e^{-2j\pi fk} \right|^2 = \underbrace{\sum_{m=-q}^{q} R_x(m) e^{-j2\pi fm}}_{P_x^{BT}(f) \text{ with } (2q+1) \text{ autocorrelation lags}}$

PARAMETRIC MODELING (14)

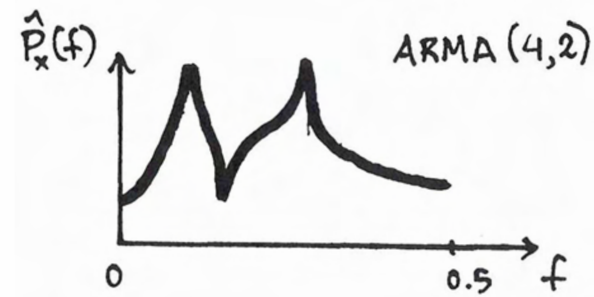
How do we choose a model ?

- **AR(p)** suitable for spectra with sharp peaks (all pole-model)



- **MA(q)**: suitable for spectra with spectral nulls (all zero-model)

- **ARMA(p,q)**: suitable for spectra with both sharp peaks and spectral nulls (zero-pole model)

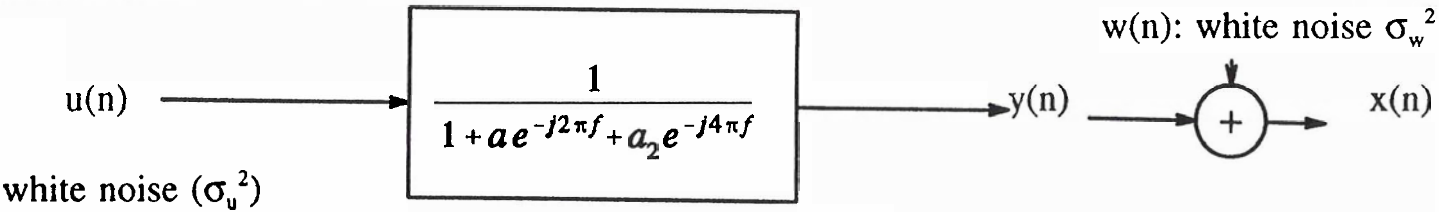


■ **p,q = ?**

Choose the model with fewest possible parameters (principle of parsimony)

PARAMETRIC MODELING (15)

- **Example**



w(n), y(n): uncorrelated

$$P_x(f) = P_y^{AR}(f) + \sigma_w^2 = \sigma_u^2 \frac{1}{|A(f)|^2} + \sigma_w^2 = \frac{\sigma_u^2 + \sigma_w^2 \cdot |1 + a_1 e^{-j2\pi f} + a_2 e^{-j4\pi f}|^2}{|1 + a_1 e^{-j2\pi f} + a_2 e^{-j4\pi f}|^2}$$

~ AR(2) if σ_w^2 low (SNR high)
~ ARMA(2,2) if σ_w^2 high (SNR low)

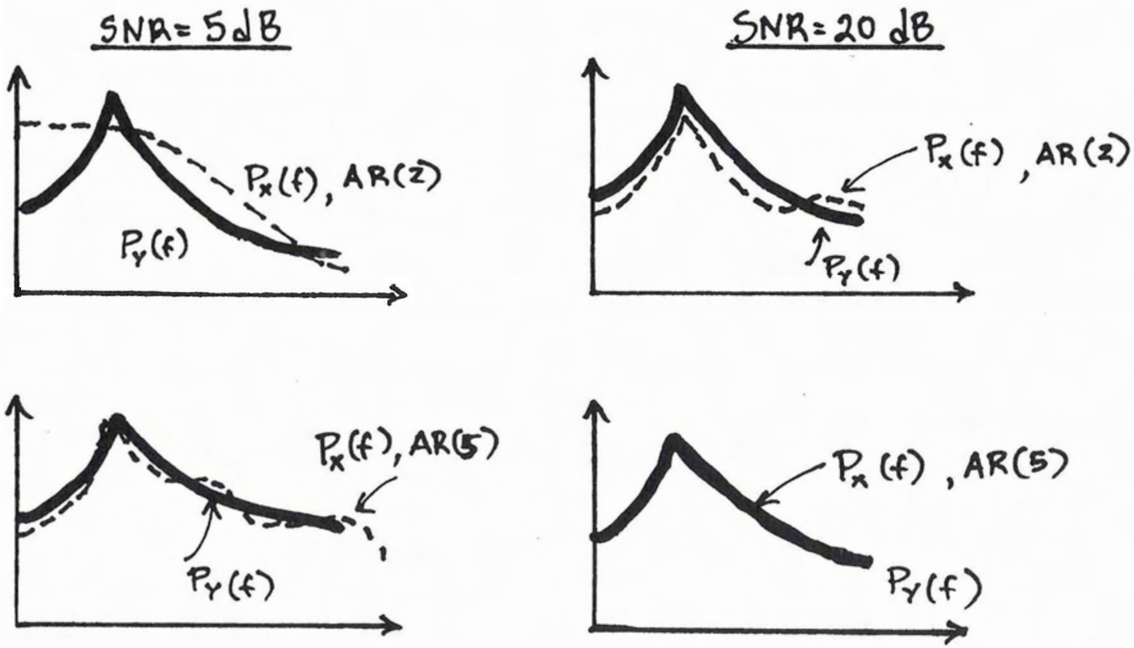
* Note that an ARMA(2,2) can be approximated by an AR(M), where M>>2. Thus due to the additive noise present in the data a higher order than the true model order is required.

Example (cont')

Let us

1) Obtain $R_x(m) = F^{-1}[P_y^{AR}(f) + \sigma_w^2] = R_y(m) + \sigma_w^2 \delta(m)$

2) Use $R_x(m)$ with Yule-Walker equations to obtain an AR(M) model



PARAMETRIC MODELING (17)

Let: $x(n) = A\cos(2\pi f_0 n + \phi)$. where $\phi \sim U(0,2\pi)$

Then: $R_x(m) = \frac{A^2}{2} \cos(2\pi f_0 m)$

$R_x(m)$ can be approximated by the autocorrelation function of an AR(2) model (process) with poles at $\underline{e^{\pm j2\pi f_0}}$

Thus: Given a harmonic process consisting of L sinusoids we must choose an AR(M) model of order $M \geq 2L$ for adequate representation

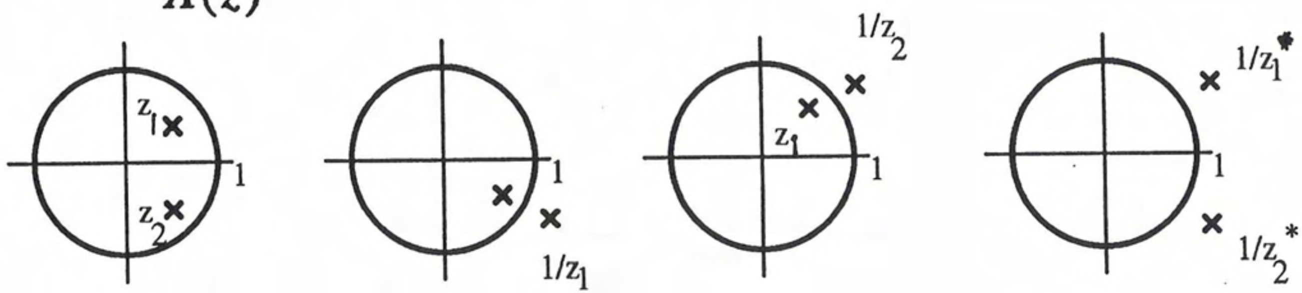
PARAMETRIC MODELING (18)

■ UNIQUENESS OF SOLUTION (1)

Let: $$P_x^{ARMA}(f) = \sigma_u^2 \frac{B(z)\,B^*(1/z^*)}{A(z)\,A^*(1/z^*)}$$

1) If z_i is a root of B(z) or A(z) then $1/z_i^*$ is a root of $B^*(1/z^*)$ or $A^*(1/z^*)$ and vice versa
2) If $|z_i|<1$, then $|1/z_i^*|>1$
3) If a model has p poles and q zeros, then, there are 2(p+q) equivalent models with the same PSD

Ex. $AR(2)$: $\dfrac{1}{A(z)}$ has two poles z_1 and z_2



PARAMETRIC MODELING (19)

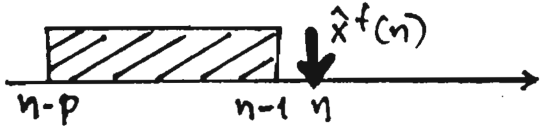
■ UNIQUENESS OF SOLUTION (2)

1) From PSD spectral estimate point of view it makes no difference which of the equivalent model is used

2) If realization of the filter is required, then the chosen model should be stable and causal (i.e., A(z) be minimum phase with zeros inside the unit circle)

3) If no constaints are given, then choose a minimum phase ARMA model (i.e., all zeros, poles inside the unit circle)

■ Relation to linear prediction (forward)



causal model

$$\hat{x}^f(n) = -\sum_{k=1}^{p} a_{p,k}^f . x(n-k)$$

Forward prediction error $e_p^f(n) = x(n) - \hat{x}^f(n) = x(n) + \sum_{k=1}^{p} a_{p,k}^f x(n-k)$

Orthogonality principle: $E\{x^*(n-k).e_p^f(n)\}=0, \quad k=0,1,\ldots,p$

(Equivalent to minimizing $E\{|e_p^f(n)|^2\}$)

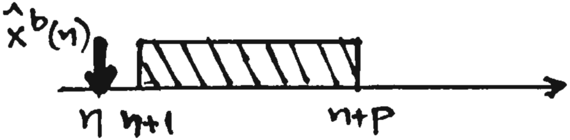
<u>Normal equations</u>

$$\begin{bmatrix} R_x(0) & R_x(-1) & \cdots & R_x(-p) \\ R_x(1) & R_x(0) & \cdots & . \\ . & . & \cdots & . \\ R_x(p) & R_x(p-1) & \cdots & R_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_{p,1}^f \\ . \\ a_{p,p}^f \end{bmatrix} = \begin{bmatrix} \sigma_{u,p}^2 \\ 0 \\ . \\ 0 \end{bmatrix}$$

$\sigma_{u,p}^2$: VAR{of driving noise) = VAR{min $e_p^f(n)$}

- **Relation to linear prediction (backward)**

anticausal model



$$\hat{x}^b(n) = -\sum_{k=1}^{p} a_{p,k}^{b} \,.\, x(n+k)$$

Backward prediction error $e_p^b(n) = x(n) - \hat{x}^b(n) = x(n) + \sum_{k=1}^{p} a_{p,k}^{b} x(n+k)$

Orthogonality principle: $E\{x^*(n+k).e_p^b(n)\}=0$, $k=0,1,\ldots p$

(Equivalent to minimizing $E\{|e_p^b(n)|^2\}$)

<u>**Normal equations**</u>

$$\begin{bmatrix} R_x(0) & R_x(-1) & \cdots & R_x(-p) \\ R_x(1) & R_x(0) & \cdots & . \\ . & . & \cdots & . \\ R_x(p) & R_x(p-1) & \cdots & R_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_{p,1}^b \\ . \\ a_{p,p}^b \end{bmatrix} = \begin{bmatrix} \sigma_{u,p}^2 \\ 0 \\ . \\ 0 \end{bmatrix}$$

Thus: $\sigma_{u,p}^2$: VAR{driving noise) = VAR{min $e_p^f(n)$} = VAR{min $e_p^b(n)$}

Also: $a_{p,k}^f = [a_{p,k}^b]^*$

■ ANOTHER INTERPRETATION OF REFLECTION COEFFICIENTS
(Makhoul, 1975)

$$K_m = a^f_{m,m} = a^{b^*}_{m,m} = \frac{-E\{e^f_{m-1}(n)\, e^{b^*}_{m-1}(n-1)\}}{\sqrt{E\{|e^f_{m-1}(n)|^2\}}.\sqrt{E\{|e^b_{m-1}(n-1)|^2\}}}\ , \quad m=1,2,\ldots,p$$

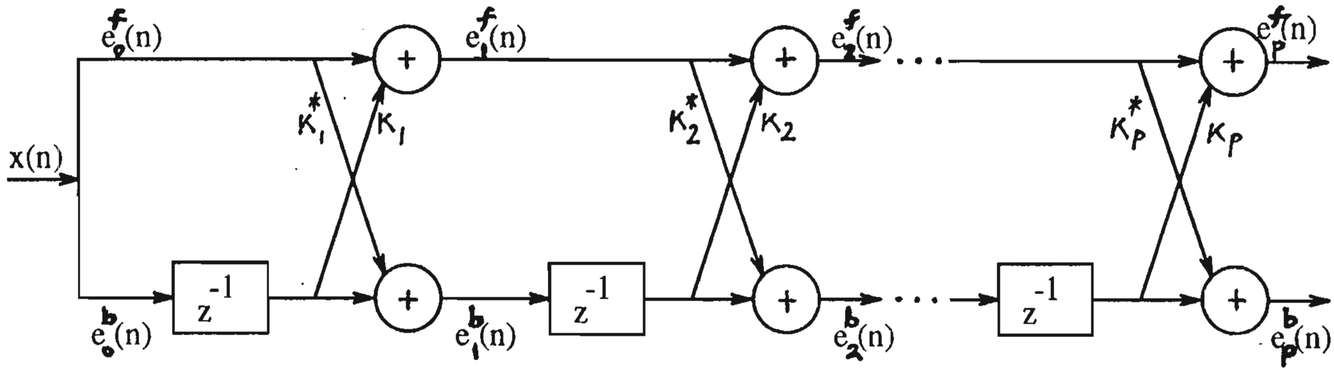
where,

$$e^f_m(n) = e^f_{m-1}(n) + K_m e^b_{m-1}(n-1)$$

$$e^b_m(n) = e^b_{m-1}(n-1) + K^*_m e^f_{m-1}(n)$$

i.c.: $x(n) = e^f_0(n) = e^b_0(n)$ (initial conditions)

■ LATICE FILTER



■ THE YULE - WALKER METHOD (1) (or, autocorrelation method)

Given $\{x(n)\}$, $n=1,2,\ldots,N$

- Estimate biased autocorrelation lags, i.e.,

$$\hat{R}_x(m) = \frac{1}{N}\sum_{n=1}^{N-|m|} x^*(n)\,x(n+m)$$

$$\hat{R}_x(-m) = \hat{R}^*(m) \qquad , \; m=0,1,2,\ldots,p$$

Form the YW equations ("normal equations")

$$\begin{bmatrix} \hat{R}_x(0) & \hat{R}_x(-1) & \ldots & \hat{R}_x(-p) \\ \hat{R}_x(1) & \hat{R}_x(0) & \ldots & . \\ . & . & \ldots & . \\ \hat{R}_x(p) & \hat{R}_x(p-1) & \ldots & \hat{R}_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ \hat{a}_{p,1} \\ . \\ \hat{a}_{p,p} \end{bmatrix} = \begin{bmatrix} \hat{\sigma}^2_{u,p} \\ 0 \\ . \\ 0 \end{bmatrix}$$

- Solve the YW equations using the Levinson - Durbin algorithm

- Obtain PSD: $$\hat{P}_x^{YW}(f) = \frac{\hat{\sigma}^2_{u,p}}{\left|\, 1 + \sum_{k=1}^{p} \hat{a}_{p,k} \cdot e^{-2j\pi fk} \,\right|^2} , \qquad 0 \le f \le 1$$

■ THE YULE - WALKER METHOD (2)

- Known as the "Autocorrelation method" (Makhoul, 1975) in least-squares linear prediction analysis

Minimize the estimate of the prediction error power

$$\frac{1}{N}\sum_{n=-\infty}^{\infty} |e_p^f(n)|^2 = \frac{1}{N}\sum_{n=1}^{N+p} |x(n) + \sum_{k=1}^{p} a_{p,k}\, x(n-k)|^2 \quad \text{w.r.t.} \quad \{a_{p,k}\},\; k=1,\dots,p$$

- Data samples involved: $\underbrace{x(1-p),\dots,x(0)}_{\text{(assumed to be 0)}},\ \underbrace{x(1),\dots x(N)}_{\text{observed}},\ \underbrace{x(N+1),\dots,x(N+p)}_{\text{(assumed to be 0)}}$

NOTES:

1) YW method exhibits poor resolution for short data records ($N \not\gg p$)
2) It is well suited for wideband processes
3) It produces a stable AR filter
4) Line splitting may be observed
   (Can be avoided by using unbiased autocorrelation estimates)

LEVINSON RECURSION [Levinson, 1949, Durbin 1960]

- Efficient way to solve linear systems with Toeplitz matrix structure, such as the Normal Equations

- Complexity ~ $O(M^2)$
  Procedure: Given $R_x(m)$, m=0,...,M one can form "normal equations" for the following AR(k) models

- k=1

$$\begin{bmatrix} R_x(0) & R_x(-1) \\ R_x(1) & R_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_{1,1} \end{bmatrix} = \begin{bmatrix} \sigma_1^2 \\ 0 \end{bmatrix}$$

- k=2

$$\begin{bmatrix} R_x(0) & R_x(-1) & R_x(-2) \\ R_x(1) & R_x(0) & R_x(-1) \\ R_x(2) & R_x(1) & R_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_{2,1} \\ a_{2,2} \end{bmatrix} = \begin{bmatrix} \sigma_2^2 \\ 0 \\ 0 \end{bmatrix}$$

- k=M

$$\begin{bmatrix} R_x(0) & \cdots & R_x(-M) \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ R_x(M) & \cdots & R_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_{M,1} \\ \cdot \\ a_{M,M} \end{bmatrix} = \begin{bmatrix} \sigma_M^2 \\ 0 \\ \cdot \\ 0 \end{bmatrix}$$

LEVINSON RECURSION (2)

- Obtain the solution of all AR models of order 1 to M recursively, i.e.,

$$
\begin{array}{llllll}
k=1 & [1, & a_{1,1}] , & \sigma_1^2 & & \\
k=2 & [1, & a_{2,1}, & a_{2,2}] , & \sigma_2^2 & \\
 & & & \vdots & & \\
k=M & \multicolumn{5}{l}{\underbrace{[1, \quad a_{M,1}, \ldots\ldots\ldots\ldots, a_{M,M}], \quad \sigma_M^2}_{\text{MESE solution}}}
\end{array}
$$

- **REFLECTION COEFFICIENTS** (or partial autocorrelations)

  $\{a_{1,1}; \quad a_{2,2}; \ldots\ldots\ldots; a_{M,M}\}$ Lattice structure

- $$
\begin{tikzcd}
 & \boxed{R_x(m), \quad m=0,\ldots,M} \arrow[dl, leftrightarrow] \arrow[dr, leftrightarrow] & \\
\boxed{\sigma_M^2, a_{M,1},\ldots,a_{M,M}} \arrow[rr, leftrightarrow] & & \boxed{R_x(0), a_{1,1},\ldots,a_{M,M}}
\end{tikzcd}
$$

LEVINSON RECURSION (3)

■ ALGORITHM

$$\sigma_0^2 = R_x(0)$$

k=1,2,...,M

$$a_{k,k} = \frac{-[R_x(k) + \sum_{j=1}^{k-1} a_{k-1,j} \cdot R_x(k-j)]}{\sigma_{k-1}^2}$$

j=1,2,...,k-1

$$a_{k,j} = a_{k-1,j} + a_{k,k} \cdot a_{k-1,k-j}^*$$

$$\sigma_k^2 = (1 - |a_{k,k}|^2)\,\sigma_{k-1}^2$$

- $0 \le \sigma_k^2 \le \sigma_{k-1}^2$
- $\sigma_M^2 = \prod_{k=1}^{M} (1 - |a_{k,k}|^2)$
- AR(M) filter stable if and only if $|a_{k,k}| < 1$, k=1,...,M
- Stop recursion at k=M, or when $a_{k,k} \ge 1$ (if a stable filter is required)

LEVINSON RECURSION (4)

EXAMPLE:

Given $R_x(0)$, $R_x(1)$, $R_x(2)$

$\rightarrow$ $\sigma_0^2 = R_x(0)$

$\rightarrow k=1$ $a_{1,1} = -\dfrac{R(1)}{R(0)}$, $\sigma_1^2 = (1 - |a_{1,1}|^2)\,\sigma_0^2$

$\rightarrow k=2$ $a_{2,2} = \dfrac{-[R_x(2) + a_{1,1} R_x(1)]}{\sigma_1^2}$

$$a_{2,1} = a_{1,1} + a_{2,2}\, a_{1,1}^*$$

$$\sigma_2^2 = (1 - |a_{2,2}|^2)\,\sigma_1^2$$

■ THE COVARIANCE METHOD (1)

Given $\{x(n)\}$, n=1,...N

- Estimate the autocorrelation lags as follows:

$$\hat{R}_x(m,l) = \frac{1}{N-p}\sum_{n=p+1}^{N} x^*(n-m)\,x(n-l) = \hat{R}_x^*(l,m)\,, \quad m,l = 0,1,\ldots p$$

- Form the linear system of equations

$$\begin{bmatrix} \hat{R}_x(0,0) & \hat{R}_x(0,1) & \cdots & \hat{R}_x(0,p) \\ \hat{R}_x(1,0) & \hat{R}_x(1,1) & \cdots & \hat{R}_x(1,p) \\ \cdot & \cdot & \cdots & \cdot \\ \hat{R}_x(p,0) & \cdot & \cdots & \hat{R}_x(p,p) \end{bmatrix} \begin{bmatrix} 1 \\ \hat{a}_{p,1} \\ \cdot \\ \hat{a}_{p,p} \end{bmatrix} = \begin{bmatrix} \hat{\sigma}^2_{u,p} \\ 0 \\ \cdot \\ 0 \end{bmatrix}$$

Hermitian but not Toeplitz

- Solve with efficient algorithm ($\sim p^2$ operations)
  [Marple recursive algorithm, 1980 (Marple, 1987, Apx 8.c)]
- Obtain PSD estimate by substituting $\{\hat{a}_{p,k}\}$, k=1,...,p into $\hat{P}_x^{AR}(f)$, $0 \le f \le 1$

■ THE COVARIANCE METHOD (2)

- "Covariance" term used for historical reasons [Makhoul, 1975]

- Equivalent to the least squares linear prediction analysis problem: minimize the estimate of the prediction error power

$$\frac{1}{N-p}\sum_{n=p+1}^{N} |e_p^f(n)|^2 = \frac{1}{N-p}\sum_{n=p+1}^{N} |x(n) + \sum_{k=1}^{p} a_{p,k}x(n-k)|^2 \quad \text{w.r.t.} \quad \{a_{p,k}\}\ ,\ k=1,\dots,p$$

- Only the observed data samples are employed (i.e., $x(1),\dots,x(n)$)
- It exhibits higher resolution than YW method
- It is well suited for harmonic processes (narrowband) (similarities with Prony method to be examined later)
- It does not guarantee a stable AR filter
- For N>>p gives similar spectral estimate with the YW method
- Nonsingular autocorrelation matrix if N≥2p (necessary condition)

MODEL ORDER SELECTION

- In practice we always have noisy observations of a time series.
  A resonable approximation of the underlaying model can be obtained by choosing "somehow" the order of the model (i.e., p $\rightarrow$ AR(p)).

- If the chosen order p is too low, we get smoothed PSD estimate.

- If p is too high, we might observe spurious peaks in estimated PSD
- The prediction error power $\sigma^2_{u,m}$ decreases monotonically as m $\rightarrow \infty$.

$$(\sigma^2_{u,m} = (1 - |K_m|^2)\,\sigma^2_{u,m-1})\ , \quad |K_m| < 1\ , \quad m = 1,2,\ldots,\infty$$

  Thus, $\sigma^2_{u,m}$ is not sufficient to determine the order of p.
- Design criteria that combine the decreasing property of $\sigma^2_{u,m}$ and the increasing property of the estimation variance.

- For a chosen p, different methods produce different estimates.

■ MODEL ORDER SELECTION CRITERIA (1)

- FINAL PREDICTION ERROR (FPE) (Akaike, 1970)

$$FPE(m) = \frac{N+m}{N-m} \cdot \sigma^2_{u,m}$$, N: # of data samples; m=1,2,3,...

Increases with m   Decreases with m

- Find the minimum value of FPE(m) vs m, i.e., $FPE_{min}(m) = FPE(p)$

- AKAIKE INFORMATION CRITERION (Akaike, 1974)

$$AIC(m) = N \ln[\sigma^2_{u,m}] + 2m$$

Increases with m   Decreases with m

- $AIC_{min}(m) = AIC(p)$

- Not consistent estimator (tends to overestimate order)

■ MODEL ORDER SELECTION CRITERIA (2)

• CRITERION AUTOREGRESSIVE TRANSFER FUNCTION (CAT)

(Parzen, 1976)

$$CAT(m) = \left[\frac{1}{N}\sum_{k=1}^{N}\frac{1}{\hat{Q}_k}\right] - \frac{1}{\hat{Q}_m}, \qquad \hat{Q}_m = \frac{N}{N-m}\cdot\sigma^2_{u,m}$$

- $CAT_{min}(m) = CAT(p)$

- Takes into account errors in the estimation of $\{a_{m,k}\}$.

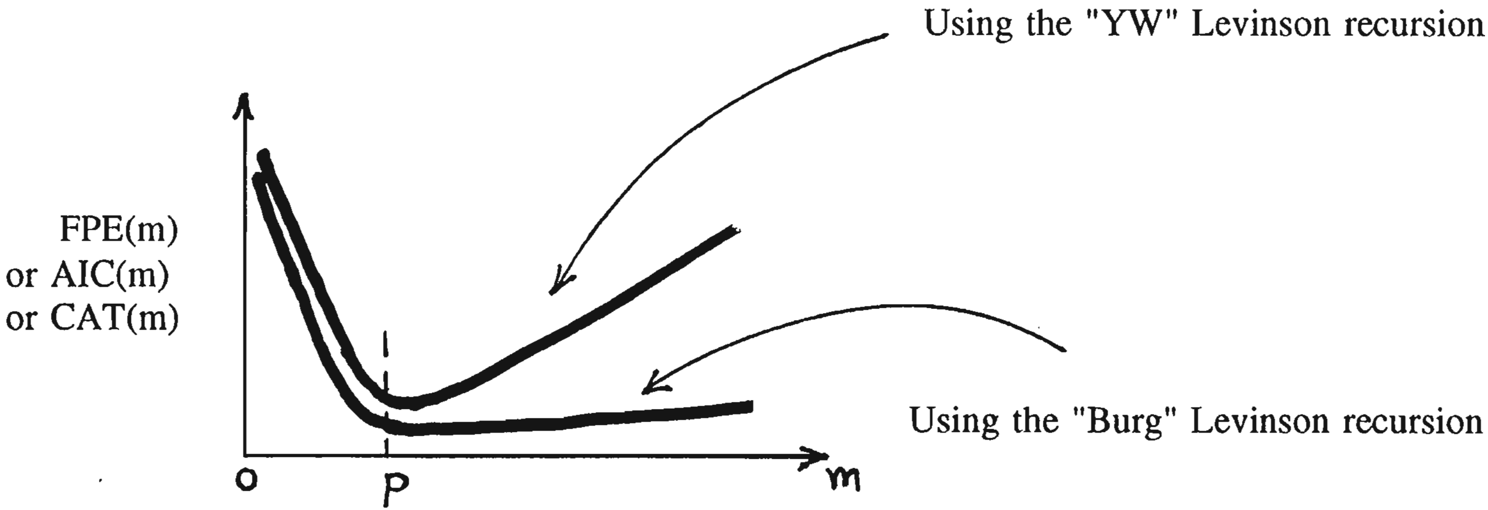
NOTES:

1) For large N, AIC(m)≅N ln FPE(m) and the two criteria become equivalent
2) For short data records AIC(m) should be preferred over FPE(m)

3) Several modifications of the above criteria exist

■ MODEL ORDER SELECTION CRITERIA (3)

Ex.



■ Rules of thumb

- Given $x(n)$, n=1,2,...,N $\quad p \sim (\frac{N}{3}\ to\ \frac{N}{2})$

- Given L sinusoids in time series, $p \geq 2L$

■ **MA MODEL** MA(q)

$$x(n) = \sum_{k=0}^{q} b_{q,k} u(n-k) \;,\; b_0 = 1$$

- $u(n)$: white noise, zero-mean, variance σ_u^2

- $E\{x(n)u(k)\}=0$ for $n<k$ or $k<n-q$

■ **MA PSD**

$$P_x^{MA}(f) = \sigma_u^2 \cdot \left| 1 + \sum_{k=1}^{q} b_{q,k} e^{-j2\pi fk} \right|^2 , \quad 0 \leq f \leq 1$$

- Appropriate for PSD that exhibits deep spectral nulls or broad peaks
- Higher resolution than classical method

$$P_x^{MA}(f) = \sigma_u^2 \cdot \left| 1 + \sum_{k=1}^{q} b_{q,k}\, e^{-j2\pi fk} \right|^2 = \sum_{m=-q}^{q} R_x(m)\, e^{-j2\pi fm}$$

BT method with (2q+1) autocorrelation lags

$$R_x(m) = \begin{cases} \sigma_u^2 \sum_{k=m}^{q} b_k b_{k-m}^*, & m=0,1,\ldots,q \\ R_x^*(-m) & \\ 0 & , \quad \text{otherwise} \end{cases}$$

NONLINEAR OPTIMIZATION APPROACH

Given $\{x(n)\}$, $n=1,\ldots,N$

1) Estimate the autocorrelation lags

$$\begin{cases} \hat{R}_x(m) = \dfrac{1}{N} \displaystyle\sum_{n=1}^{N-|m|} x^*(n)\, x(n+m) \\ \hat{R}_x(-m) = R_x^*(m) \ , \quad m = 0,\ldots,q \end{cases}$$

2) Solve the nonlinear system of equations

$$\begin{cases} R_x(m) = \sigma_u^2 \displaystyle\sum_{k=m}^{q} b_k\, b_{k-m}^* \ , \quad m=0,\pm 1,\ldots,\pm q \\ R_x(m) = 0 \ , \quad \text{otherwise} \end{cases}$$

- Difficult optimization approaches such as spectral factorization techniques [Box, Jenkins, 1970]

3) Obtain $P_x^{MA}(f)$, $0 \leq f \leq 1$

■ **DURBIN'S METHOD (1)** Given $\{x(n)\}$, $n=1,\ldots,\infty$

MA(q) model:

$$x(n) = \sum_{k=0}^{q} b_{q,k} u(n-k)$$

$$B(z) = \sum_{k=0}^{q} b_{q,k} z^{-k}, \quad b_{q,0} = 1$$

AR(∞) model:

$$x(n) = -\sum_{k=1}^{\infty} a_{\infty,k} x(n-k) + u(n)$$

$$A(z) = \sum_{k=0}^{\infty} a_{\infty,k} z^{-k}, \quad a_{\infty,0} = 1$$

Approximation:

$$B(z) \cong \frac{1}{A(z)} \Rightarrow B(z)A(z) \cong 1 \text{ or } a_{\infty,m} + \sum_{k=1}^{q} b_{q,k} a_{\infty,m-k} = \delta(m) = \begin{cases} 1, & m=0 \\ 0, & m \neq 0 \end{cases}$$

■ **DURBIN'S METHOD (2)** Given $\{x(n)\}$, $n=1,\ldots,N$

1) Choose $L \gg q$ (but $L \ll N$)
2) Estimate biased autocorrelation lags

$$\hat{R}_x(m) = \frac{1}{N}\sum_{n=1}^{N-m} x^*(n)\,x(n+m), \quad \hat{R}_x(-m) = R_x^*(m)\ , \qquad m=0,\ldots,L$$

3) Obtain coefficients of AR(L) model from the YW equations and Levinson recursion

$$\begin{bmatrix} \hat{R}_x(0) & \hat{R}_x(-1) & \cdots & \hat{R}_x(-L) \\ \hat{R}_x(1) & \hat{R}_x(0) & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \hat{R}_x(L) & \cdot & \cdots & \hat{R}_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_{L,1} \\ \cdot \\ a_{L,L} \end{bmatrix} = \begin{bmatrix} \sigma^2_{u,L} \\ 0 \\ \cdot \\ 0 \end{bmatrix}$$

(One may also use the covariance or modified covariance method)

4) Approximate AR(L) to MA(q) (Deterministic AR(q) model)

$$a_{L,m} + \sum_{k=1}^{q} b_{q,k}\, a_{L,m-k} = \delta(m)$$

■ DURBIN'S METHOD (3)

5) Estimate biased autocorrelation lags

$$\hat{R}_a(m) = \frac{1}{L+1}\sum_{n=0}^{L-m} a^*(n)\,a(n+m), \qquad \hat{R}_a(-m) = \hat{R}_a^*(m)\,, \qquad m=1,\dots,q$$

6) Obtain coefficients of MA(q) model using Levinson recursion

$$\begin{bmatrix} \hat{R}_a(0) & \hat{R}_a(-1) & \cdots & \hat{R}_a(-q+1) \\ \hat{R}_a(1) & \hat{R}_a(0) & \cdots & . \\ . & . & \cdots & . \\ \hat{R}_a(q-1) & . & \cdots & \hat{R}_a(0) \end{bmatrix} \begin{bmatrix} b_{q,1} \\ b_{q,2} \\ . \\ b_{q,q} \end{bmatrix} = - \begin{bmatrix} \hat{R}_a(1) \\ \hat{R}_a(2) \\ . \\ \hat{R}_a(q) \end{bmatrix}$$

7) Obtain $\hat{P}_x^{MA}(f) = \sigma_{u,L}^2 \cdot \left| 1 + \sum_{k=1}^{q} b_{q,k}\, e^{-j2\pi fk} \right|^2$

- Minimum phase property of MA(q) coefficients
- Model order selection criterion $AIC_{min}(m)=AIC(q)$

  AIC : $[AIC(m)=N \ln \hat{\sigma}_u(m) + 2m]$ $m=1,2,\dots$

- <u>ARMA (p,q) MODEL</u>

$$x(n) = -\sum_{k=1}^{p} a_k x(n-k) + \sum_{k=0}^{q} b_k u(n-k) \ , \quad b_0 = 1$$

$u(n)$: white noise, zero-mean, variance σ_u^2

- <u>ARMA (p,q) PSD</u>

$$P_x^{ARMA}(f) = \frac{\sigma_u^2 \cdot |1 + \sum_{k=1}^{q} b_k e^{-j2\pi fk}|^2}{|1 + \sum_{k=1}^{p} a_k e^{-j2\pi fk}|^2} \ , \quad 0 \leq f \leq 1$$

- Appropriate for PSD that exhibits both sharp spectral peaks and deep spectral nulls

■ Relation of ARMA(p,q) parameter and autocorrelation

$$R_x(m) = \begin{cases} R_x^*(-m) & , \quad m<0 \\ -\sum_{k=1}^{p} a_k R_x(m-k) + \sigma_u^2 \cdot \sum_{k=m}^{q} b_k h_{k-m}^* , & 0 \le m \le q \\ -\sum_{k=1}^{p} a_k R_x(m-k) , & m>q \end{cases}$$

where $\{h(k)\}$ is the impulse response of the filter

■ Optimum solution

- Highly nonlinear with respect to the $\{b_k\}$, $\{h_k\}$
- Requires nonlinear optimization techniques (Box, Jenkins, 1970) (Convergence to the global solution is not guaranteed)

■ **In practice suboptimum linear solutions are employed**

■ **No best ARMA method exists.**

■ THE MODIFIED YW METHOD (1)

Suboptimum solution:

- First generate the AR coefficients, then the MA coefficients

Given $\{x(n)\}$, $n=1,\dots,N$

1) Generate the autocorrelation estimates $\hat{R}_x(m)$, $m=q-p+1,\dots,q+p-1$
2) Form the modified YW equations

(from $\hat{R}_x(m) = -\sum_{k=1}^{p} a_k R_x(m-k)$, $|m| = q+1,\dots,q+p$)

i.e.,

$$\begin{bmatrix} \hat{R}_x(q) & \cdots & \hat{R}_x(q-p+1) \\ \cdot & \cdots & \cdot \\ \hat{R}_x(q+p-1) & \cdots & \hat{R}_x(q) \end{bmatrix} \cdot \begin{bmatrix} \hat{a}_1 \\ \cdot \\ \hat{a}_p \end{bmatrix} = - \begin{bmatrix} \hat{R}_x(q+1) \\ \cdot \\ \hat{R}_x(q+p) \end{bmatrix}$$

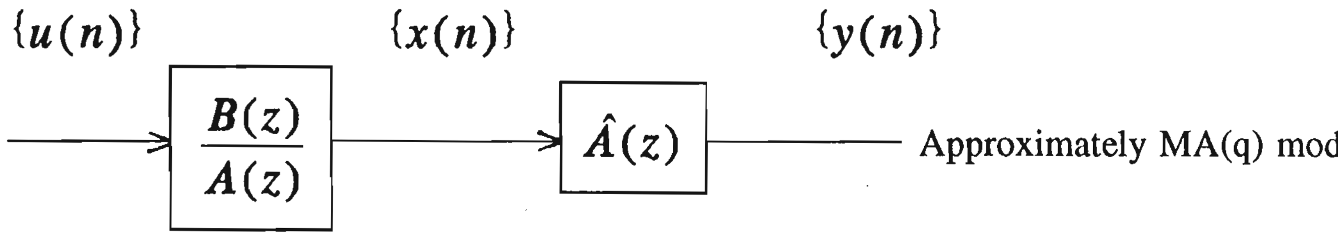
3) Obtain $\{\hat{a}_m\}$, $m=1,\dots,p$ using the modified Levinson recursion

- Stable filter is not guaranteed, nonsingular matrix is possible

THE MODIFIED YW METHOD (2)

4) Filter observed data with a filter having transfer function

$$\hat{A}(z) = 1 + \sum_{k=1}^{p} \hat{a}_k z^{-k} \quad , \quad \text{i.e.,}$$



5) Apply MA spectrum estimation methods on $\{y(n)\}$ to obtain $\{\hat{b}_k\}$, $k=1,\ldots,q$

■ THE MODIFIED YW METHOD (3)

6) $$\hat{P}_x^{ARMA}(f) = \frac{\sigma_u^2 \cdot \left|1 + \sum_{k=1}^{q} \hat{b}_k e^{-j2\pi fk}\right|^2}{\left|1 + \sum_{k=1}^{p} \hat{a}_k e^{-j2\pi fk}\right|^2}, \quad 0 \le f \le 1$$

or alternatively $$\hat{P}_x^{ARMA}(f) = \frac{\sum_{m=-q}^{q} \hat{R}_y(m)\, e^{-j2\pi fm}}{\left|1 + \sum_{k=1}^{p} \hat{a}_k e^{-j2\pi fk}\right|^2}, \quad 0 \le f \le 1$$

- Sometimes good, sometimes poor performance

- Very sensitive to AR(p) model order selection

- Positive power spectral density is not guaranteed

MODIFIED LEVINSON RECURSION ALGORITHM

(For systems with Toeplitz but non-symmetric structure)

$$a_{1,1} = -\frac{R_x(q+1)}{R_x(q)}, \qquad c_{1,1} = -\frac{R_x(q-1)}{R_x(q)}$$

$k = 2,3,\ldots,p$

$$a_{k,k} = \frac{-[R_x(q+k) + \sum_{j=1}^{k-1} a_{k-1,j}.R_x(q+k-j)]}{\sigma^2_{u,k-1}}$$

$$c_{k,k} = \frac{-[R_x(q-k) + \sum_{j=1}^{k-1} c_{k-1,j}.R_x(q-k-j)]}{\sigma^2_{u,k-1}}$$

$j = 1,2,\ldots,k-1$

$$a_{k,j} = a_{k-1,j} + a_{k,k}.c_{k-1,k-j}$$

$$c_{k,j} = b_{k-1,j} + b_{k,k}.c_{k-1,k-j}$$

$$\sigma^2_{u,k} = (1 - a_{k,k}c_{k,k})\,\sigma^2_{u,k-1}$$

By placing $c_{k,j} = a^*_{k,j} \rightarrow$ Levinson recursion