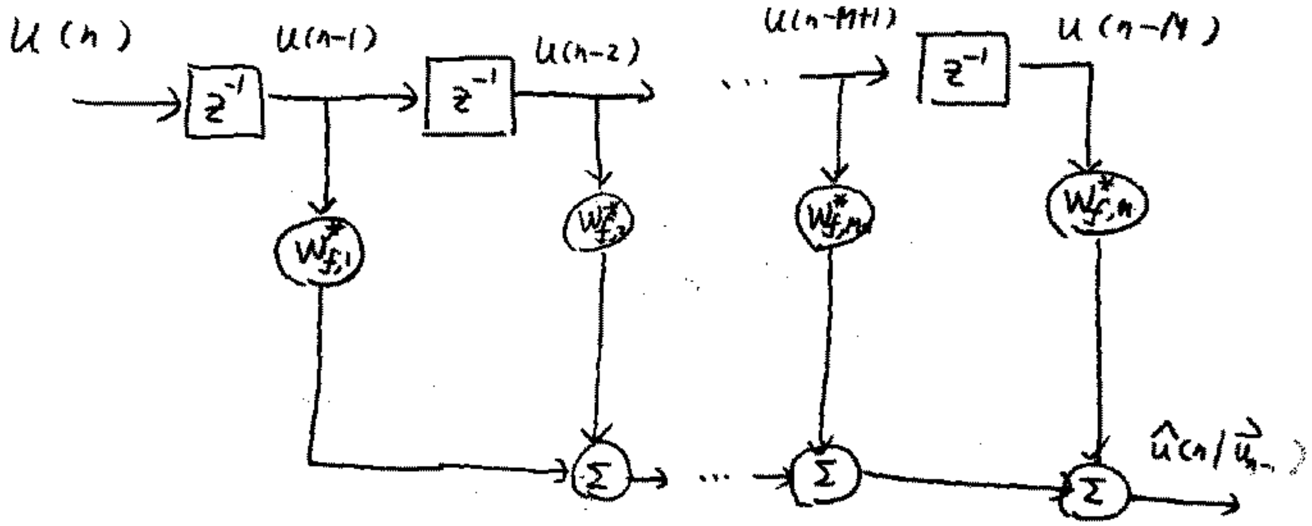
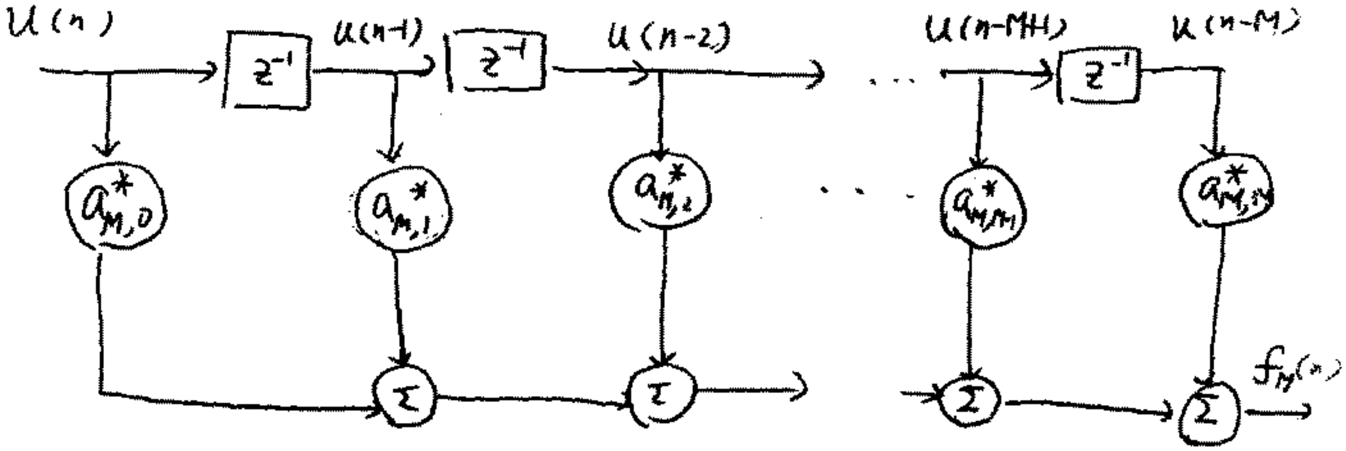
Chapter 3 Linear Prediction



a linear predictor (one step)



a prediction-error filter

$(a_{M,0} = 1)$

The predicted value in a linear predictor can be written as

$$\hat{u}(n|\vec{U}_{n-1}) = \sum_{k=1}^{M} w_{f,k}^{*} u(n-k)$$

The desired response is $d(n) = u(n)$.

Accordingly, the forward prediction error is $f_M(n) = u(n) - \hat{u}(n|\vec{U}_{n-1})$ and its mean-square value is

$$P_M = E\left[|f_M(n)|^2\right].$$

Based on this linear predictor system, the Wiener-Hopf Equation can be formulated.

Let's denote

$$\vec{w}_f = [w_{f,1}, w_{f,2}, \cdots, w_{f,M}]^T$$

$$\vec{u}(n-1) = [u(n-1), u(n-2), \cdots, u(n-M)]^T$$

$$\tilde{R} = E[\vec{u}(n-1)\vec{u}^H(n-1)]$$

$$= \begin{bmatrix} r(0) & r(1) & \cdots & r(M-1) \\ r^*(1) & r(0) & \cdots & r(M-2) \\ \vdots & & & \vdots \\ r^*(M-1) & r^*(M-2) & \cdots & r(0) \end{bmatrix}$$

The cross-correlation vector

$$\vec{p} = \vec{r} = E[\vec{u}(n-1)\, \underbrace{u^*(n)}_{d^*(n)}]$$

$$= \begin{bmatrix} r^*(1) \\ r^*(2) \\ \vdots \\ r^*(M) \end{bmatrix} = \begin{bmatrix} r(-1) \\ r(-2) \\ \vdots \\ r(-M) \end{bmatrix}$$

The Wiener-Hopf Equations based on this linear predictor is

$$\tilde{R}\, \vec{w}_f = \vec{r}$$

and the minimum mean square prediction-error is

$$P_M = r(0) - \vec{r}^H \vec{w}_f$$

※ Forward Prediction-Error Filter

If we just add one more connection to the linear predictor, namely, the desired signal $u(n)$, then we can construct a prediction-error filter. The output of the prediction-error filter is just the prediction error sequence $f_M(n)$, such that

$$f_M(n) = u(n) - \sum_{k=1}^{M} w_{f,k}^{*} u(n-k)$$

$$\Rightarrow f_M(n) = \sum_{k=0}^{M} a_{M,k}^{*} u(n-k),$$

where $a_{M,k} = \begin{cases} 1, & k=0 \\ -w_{f,k}, & k=1,2,\dots,M \end{cases}$

※ Augmented Wiener-Hopf Equations for Forward Prediction:

We can combine Eq. (3.9) and (3.10) to construct an augmented Wiener-Hopf Equation such that

$$\begin{bmatrix} r(0) & \vec{r}^{H} \\ \vec{r} & \tilde{R} \end{bmatrix} \begin{bmatrix} 1 \\ -\vec{w}_f \end{bmatrix} = \begin{bmatrix} P_M \\ \vec{0} \end{bmatrix}$$

or $\displaystyle\sum_{l=0}^{M} a_{M,l}\, r(l-i) = \begin{cases} P_M, & i=0 \\ 0, & i=1,2,\dots,M \end{cases}$

where $\vec{a}_M = \begin{bmatrix} 1 \\ -\vec{w}_f \end{bmatrix}$

Example: Compute the $\vec{a}_M$ for a linear predictor of order $M = 1, 2$

Equation (3.14) yields

$M=1$,
$$\begin{bmatrix} r(0) & r(1) \\ r^*(1) & r(0) \end{bmatrix} \begin{bmatrix} a_{1,0} \\ a_{1,1} \end{bmatrix} = \begin{bmatrix} P_1 \\ 0 \end{bmatrix}$$

$$a_{10} = \frac{P_1}{\Delta r} r(0), \quad a_{1,1} = -\frac{P_1}{\Delta r} r^*(1),$$

where $\Delta r = \det\left(\begin{bmatrix} r(0) & r(1) \\ r^*(1) & r(0) \end{bmatrix} \right)$

$$= r^2(0) - |r(1)|^2$$

$$P_1 = \frac{\Delta r}{r(0)} \quad \text{since} \quad a_{1,0} = 1$$

$$\Rightarrow \quad a_{1,1} = -\frac{r^*(1)}{r(0)}$$

$M=2$,
$$\begin{bmatrix} r(0) & r(1) & r(2) \\ r^*(1) & r(0) & r(1) \\ r^*(2) & r^*(1) & r(0) \end{bmatrix} \begin{bmatrix} a_{2,0} \\ a_{2,1} \\ a_{2,2} \end{bmatrix} = \begin{bmatrix} P_2 \\ 0 \\ 0 \end{bmatrix}$$

$$a_{2,0} = \frac{P_2}{\Delta r} [r^2(0) - |r(1)|^2],$$

$$a_{2,1} = -\frac{P_2}{\Delta r} [r^*(1) r(0) - r(1) r^*(2)],$$

$$a_{2,2} = \frac{P_2}{\Delta r} [(r^*(1))^2 - r(0) r^*(2)]$$

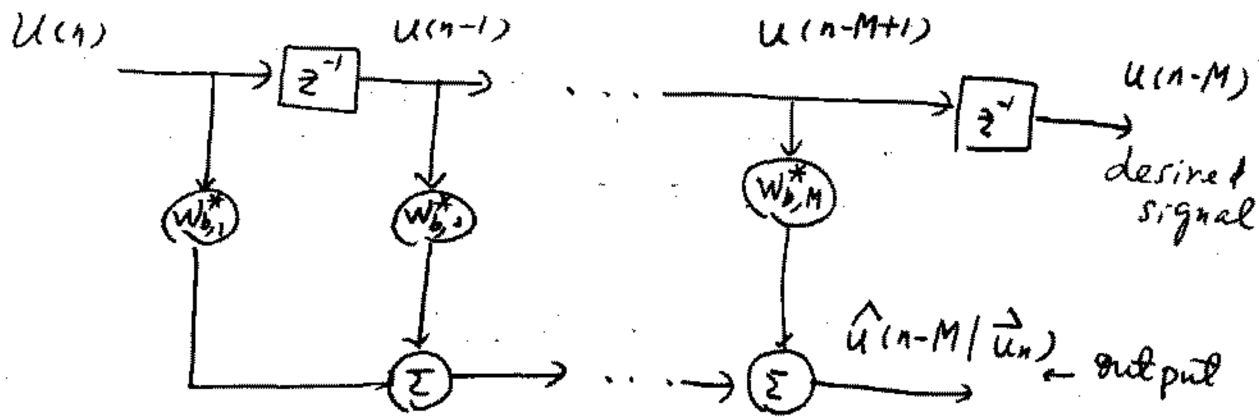
where $\Delta_r = \det\left(\begin{bmatrix} r(0) & r(1) & r(2) \\ r^*(1) & r(0) & r(1) \\ r^*(2) & r^*(1) & r(0) \end{bmatrix}\right)$

Since $a_{2,0} = 1$, $\quad P_2 = \dfrac{\Delta_r}{r^2(0) - |r(1)|^2}$

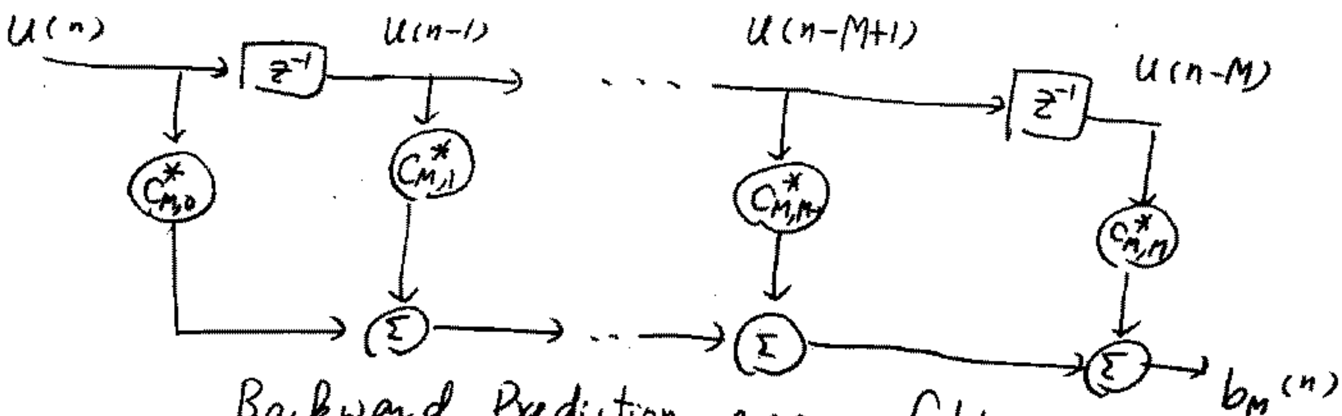
$$a_{2,1} = -\frac{r^*(1)\,r(0) - r(1)\,r^*(2)}{r^2(0) - |r(1)|^2}$$

$$a_{2,2} = \frac{(r^*(1))^2 - r(0)\,r^*(2)}{r^2(0) - |r(1)|^2}$$

3.2 Backward Linear Prediction



One-step backward predictor



Backward Prediction-error filter

The output of a backward predictor can be written as

$$\hat{u}(n-M \mid \vec{u}_n) = \sum_{k=1}^{M} w_{b,k}^{*}\, u(n-k+1)$$

$$d(n) = u(n-M),$$

and the prediction error is

$$b_M(n) = u(n-M) - \hat{u}(n-M \mid \vec{u}_n),$$

Thus the mean-square error is

$$P_M = E\{|b_M(n)|^2\}, \quad \forall n$$

Let $\vec{w}_b$ denote the M-by-1 optimum tap weight vector of the backward predictor in Fig 3-2 (a), such that

$$\vec{w}_b = [w_{b,1}, w_{b,2}, \ldots, w_{b,M}]^T.$$

To find the $\vec{w}_b$ through Wiener-Hopf Equations we need to construct the following stuff:

1. The input vector of this backward predictor is $\vec{u}(n) = [u(n), u(n-1), \cdots, u(n-M+1)]^T$.

Thus, the correlation matrix is

$$\tilde{R} = E[\vec{u}(n)\,\vec{u}^H(n)]$$

2. The cross-correlation vector is

$$\vec{p} = \vec{r}^{B*} = E[\vec{u}(n)\,u^*(n-M)]$$

3. The desired signal energy is

$$E[|d(n)|^2] = E[|u(n-M)|^2] = r(0)$$

Therefore the Wiener-Hopf Equations can be formulated as

$$\tilde{R}\,\vec{w_b} = \vec{r}^{B*},$$

and the minimum mean-square backward prediction error is

$$P_M = r(0) - \vec{r}^{B^T}\vec{w_b}$$

X. Relations between backward and Forward Predictors:

Compare Eq. (3.9) and Eq. (3.24), the Wiener-Hopf Equations for foward predictor and backward predictor, respectively,

$$\begin{cases} \tilde{R}\,\vec{w}_f = \vec{r} \\ \tilde{R}\,\vec{w}_b = \vec{r}^{B*} \Rightarrow \tilde{R}^T \vec{w}_b^{B} = \vec{r}^{*} \end{cases}$$

$$\Rightarrow \tilde{R}^H \vec{w}_b^{B*} = \vec{r}$$

$$\Rightarrow \tilde{R}\,\vec{w}_b^{B*} = \vec{r}$$

Therefore, $\vec{w}_b^{B*} = \vec{w}_f$

Hence the forward prediction-error in Eq. (3.10)

$$P_M = r(0) - \vec{r}^H \vec{w}_f$$

$$= r(0) - \vec{r}^H \vec{w}_b^{B*}$$

$$P_M^B = P_M = r(0) - \vec{r}^{B^H} \vec{w}_b^{*}$$

$$(P_M^B)^* = P_M = r(0) - \vec{r}^{B^T} \vec{w}_b \dashrightarrow \text{Eq. (3.25)}$$

Therefore, we find that the mean-square backward prediction-error has exactly the same value as the mean-square forward prediction-error. P_M can denote both kinds of prediction errors quantitively.

※ Backward Prediction-Error Filter

The backward prediction error $b_M(n)$ equals

$$b_m(n) = u(n-M) - \sum_{k=1}^{M} w_{b,k}^{*}\, u(n-k+1)$$

If we define the coefficients in such a backward prediction error filter as follows:

$$C_{M,k} = \begin{cases} -w_{b,k+1} & , k=0,1,\cdots,M-1 \\ 1 & , k=M. \end{cases}$$

Hence, $b_M(n) = \sum_{k=0}^{M} C_{M,k}^{*}\, u(n-k)$

From Eq.(3.28), we have

$$w_{b,M-k+1}^{*} = w_{f,k} \quad , k=1,2,\cdots,M$$

or $$w_{b,k} = w_{f,M-k+1}^{*} \quad , k=1,2,\cdots,M$$

$$\therefore \quad c_{M,k} = \begin{cases} -w_{f,M-k}^{*}, & k = 0, 1, \cdots, M-1 \\ 1 & , k = 0 \end{cases}$$

Compare it with Eq. (3.12), we have

$$c_{M,k} = a_{M,M-k}^{*}, \quad k = 0, 1, 2, \cdots, M$$

Hence, the backward prediction error filter output can be written as

$$b_M(n) = \sum_{k=0}^{M} a_{M,M-k}\, u(n-k)$$

※ Augmented Wiener-Hopf Equations for Backward Prediction

Combine Eq. (3.24) and Eq. (3.25), we have

$$\begin{bmatrix} \tilde{R} & \vec{r}^{B*} \\ \vec{r}^{B^T} & r(0) \end{bmatrix} \begin{bmatrix} -\vec{w}_b \\ 1 \end{bmatrix} = \begin{bmatrix} \vec{0} \\ P_M \end{bmatrix}$$

Or $\sum_{l=0}^{M} a_{M,M-l}^{*}\, r(l-i) = \begin{cases} 0, & i = 0, \ldots, M-1 \\ P_M, & i = M \end{cases}$

3.3 Levinson-Durbin Algorithm

We can recursively compute the prediction error filter coefficients from model order $m = 1, 2, 3, \ldots, M$, where M is the final order of the filter. Let the $(m+1)$-by-1 vector $\vec{a}_m$ denote the tap-weight vector of a forward prediction error filter of order m; then $\vec{a}_m^{B*}$ denotes the tap-weight vector corresponding to backward prediction error filter coefficients.

The Levinson-Durbin recursion may be stated in one of the following two equivalent ways.

1. The tap-weight vector of a forward prediction-error filter may be order-updated by the equation:

$$\vec{a}_m = \begin{bmatrix} \vec{a}_{m-1} \\ 0 \end{bmatrix} + K_m \begin{bmatrix} 0 \\ \vec{a}_{m-1}^{B*} \end{bmatrix},$$

   where K_m is a constant.

The scalar version of this order update is

$$a_{m,l} = a_{m-1,l} + k_m a^*_{m-1,m-l}, \quad l = 0, 1, \cdots, m,$$

where $a_{m,l}$ is the l^{th} tap-weight of a forward prediction-error filter of order m, and $a^*_{m-1,m-l}$ is the l^{th} tap-weight of a backward prediction-error filter of order $m-1$. Note that $a_{m-1,0} = 1$, and $a_{m-1,m} = 0$

2 The tap-weight vector of a backward prediction-error filter may be order-updated by the equation:

$$\vec{a}_m^{B*} = \begin{bmatrix} 0 \\ \vec{a}^{B*}_{m-1} \end{bmatrix} + k_m^* \begin{bmatrix} \vec{a}_{m-1} \\ 0 \end{bmatrix}.$$

The scalar version of this order update is

$$a^*_{m,m-l} = a^*_{m-1,m-l} + k^*_m a_{m-1,l}, \quad l = 0, 1, \ldots, m,$$

where $a^*_{m,m-l}$ is the l^{th} tap-weight of the backward prediction-error filter of order m.

The Levinson-Durbin recursion is usually formulated in the context of forward prediction given by Eq. (3.40) or Eq. (3.41). The formulation of the recursion in the context of backward prediction, in Eq. (3.42) and (3.43) can follow from Eq. (3.40) & (3.41).

To derive Lenvinson-Durbin algorithm, we proceed in four stages:

1. We premultiply both sides of Eq. (3.40) by $\tilde{R}_{m+1}$ such that

$$\tilde{R}_{m+1}\vec{a}_m = \begin{bmatrix} P_m \\ \vec{0} \end{bmatrix},$$

where P_m is the mean-square forward prediction error.

2. Since $\tilde{R}_{m+1} = \begin{bmatrix} \tilde{R}_m & \vec{r}^{B^*} \\ \vec{r}_m^{B^T} & r(0) \end{bmatrix}$,

$$\tilde{R}_{m+1}\begin{bmatrix} \vec{a}_{m-1} \\ 0 \end{bmatrix} = \begin{bmatrix} \tilde{R}_m & \vec{r}_m^{B^*} \\ \vec{r}_m^{B^T} & r(0) \end{bmatrix} \begin{bmatrix} \vec{a}_{m-1} \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} \tilde{R}_m \vec{a}_{m-1} \\ \vec{r}_m^{B^T} \vec{a}_{m-1} \end{bmatrix}$$

Iteratively, the set of augmented Wiener-Hopf equations for the forward prediction-error filter of order $m-1$ is

$$\tilde{R}_m \vec{a}_{m-1} = \begin{bmatrix} P_{m-1} \\ \vec{0}_{m-1} \end{bmatrix},$$

where P_{m-1} is the mean-square predition error for this filter of order $m-1$. Next, we define a scalar

$$\Delta_{m-1} = \vec{r}_m^{\,B^T} \vec{a}_{m-1}$$

$$= \sum_{l=0}^{m-1} r(l-m)\, a_{m-1,l}.$$

Substituting Eq. (3.46) and (3.47) into Eq. (3.45), we have

$$\tilde{R}_{m+1} \begin{bmatrix} \vec{a}_{m-1} \\ 0 \end{bmatrix} = \begin{bmatrix} P_{m-1} \\ \vec{0}_{m-1} \\ \Delta_{m-1} \end{bmatrix}$$

3. Since $\tilde{R}_{m+1} = \begin{bmatrix} r(0) & \vec{r}_m^{\,H} \\ \vec{r}_m & \tilde{R}_m \end{bmatrix}$,

$$\tilde{R}_{m+1} \begin{bmatrix} 0 \\ \vec{a}_{m-1}^{\,B^*} \end{bmatrix} = \begin{bmatrix} r(0) & \vec{r}_m^{\,H} \\ \vec{r}_m & \tilde{R}_m \end{bmatrix} \begin{bmatrix} 0 \\ \vec{a}_{m-1}^{\,B^*} \end{bmatrix}$$

$$= \begin{bmatrix} \vec{r}_m^{\,H} \vec{a}_{m-1}^{\,B^*} \\ \tilde{R}_m \vec{a}_{m-1}^{\,B^*} \end{bmatrix}$$

Consequently,

$$\vec{r}_m^{\,H} \vec{a}_{m-1}^{\,B*} = \sum_{k=1}^{m} r^*(-k)\, a^*_{m-1,\,m-k}$$

$$= \sum_{l=0}^{m-1} r^*(l-m)\, a^*_{m-1,\,l}$$

$$= \Delta^*_{m-1}.$$

Therefore, the set of augmented Wiener-Hopf equations for the backward prediction-error filter of order $m-1$ is

$$\tilde{R}_m \vec{a}_{m-1}^{\,B*} = \begin{bmatrix} \vec{0}_{m-1} \\ P_{m-1} \end{bmatrix}.$$

Substituting Eq. (3.50), (3.51) into Eq. (3.49) we may write

$$\tilde{R}_{m+1} \begin{bmatrix} 0 \\ \vec{a}^{\,B*}_{m-1} \end{bmatrix} = \begin{bmatrix} \Delta^*_{m-1} \\ \vec{0}_{m-1} \\ P_{m-1} \end{bmatrix}.$$

4. In summary, combining (3.44), (3.48) and (3.52), we will have

$$\begin{bmatrix} P_m \\ \vec{0}_m \end{bmatrix} = \begin{bmatrix} P_{m-1} \\ \vec{0}_{m-1} \\ \Delta_{m-1} \end{bmatrix} + k_m \begin{bmatrix} \Delta^*_{m-1} \\ \vec{0}_{m-1} \\ P_{m-1} \end{bmatrix}$$

Eq. (3.40) and Eq. (3.53) can provide the complete recursion for order-update from $m-1$ to m, in the application of a prediction-error filter. $P_m = P_{m-1} + K_m \Delta_{m-1}^*$

From Eq. (3.53), two important deductions can be achieved:

1. $P_m = P_{m-1} + K_m \Delta_{m-1}^*$

2. $0 = \Delta_{m-1} + K_m P_{m-1}$

$$\Rightarrow K_m = -\frac{\Delta_{m-1}}{P_{m-1}}$$

$$\Rightarrow P_m = P_{m-1}(1-|K_m|^2)$$

Hence, $0 \le P_m \le P_{m-1}$, $m \ge 1$, $P_0 = r(0)$,

deductively, $P_M = P_0 \prod_{m=1}^{M} (1-|K_m|^2)$

※ Interpretation of the Parameters K_m and Δ_{m-1}

The parameters K_m, $1 \le m \le M$, in the Levinson-Durbin recursion are called reflection coefficients.

According to Eq. (3.58),

$$|K_m| \leq 1, \quad \forall m$$

According to Eq. (3.41),

$$K_m = a_{m,m}$$

Δ_{m-1} is interpreted as a cross-correlation between the forward prediction error $f_{m-1}(n)$ and the delayed backward prediction error $b_{m-1}(n-1)$.

$$\Delta_{m-1} = E\left\{ b_{m-1}(n-1)\, f_{m-1}^{*}(n) \right\},$$

where $f_{m-1}(n)$ is produced at the output of a forward prediction-error filter of order $m-1$ in response to the tap inputs $u(n), u(n-1), \ldots, u(n-m+1)$ and $b_{m-1}(n-1)$ is the output of a backward prediction-error filter of order $m-1$ in response to the tap inputs $u(n-1), u(n-2), \ldots, u(n-m)$.

Note that $f_0(n) = b_0(n) = u(n)$

$$\Delta_0 = E\left[b_0(n-1) f_0^*(n) \right]$$
$$= E\left[u(n-1) u^*(n) \right] = r^*(1).$$

※ Relationship between reflection coefficients and partial correlation coefficients:

The partial correlation (PARCOR) coefficient between the forward prediction error $f_{m-1}(n)$ and the delayed backward prediction error $b_{m-1}(n-1)$ is defined as

$$\rho_m = \frac{E\left[b_{m-1}(n-1) f_{m-1}^*(n) \right]}{\sqrt{\left(E\left[|b_{m-1}(n-1)|^2 \right] E\left[|f_{m-1}(n)|^2 \right] \right)}}$$

According to the definition, we have

$$|\rho_m| \le 1, \quad \forall m,$$

since
$$\left| E\left[b_{m-1}(n-1) f_{m-1}^*(n) \right] \right|^2 \le E\left[|b_{m-1}(n-1)|^2 \right] \times E\left[|f_{m-1}(n)|^2 \right].$$

From Eq. (3.56) and (3.60),
$$K_m = - \frac{E\left[b_{m-1}(n-1) f_{m-1}^*(n) \right]}{P_{m-1}}$$

and $P_{m-1} = E[|f_{m-1}(n)|^2] = E[|b_{m-1}(n-1)|^2]$

From Eq. (3.61) and Eq. (3.62), we may obtain $\kappa_m = -\rho_m$.

Example: Use the Levinson-Durbin recursion to determine the corresponding tap weights $a_{3,1}$, $a_{3,2}$, $a_{3,3}$ and the prediction-error power P_3 for a prediction-error filter of order 3.

$m = 1, 2, 3$

1. For a prediction-error filter of order $m=1$

$$a_{1,0} = 1$$
$$a_{1,1} = \kappa_1, \quad \text{and} \quad P_1 = P_0(1 - |\kappa_1|^2)$$

2. For a prediction-error filter of order $m=2$

$$a_{2,0} = 1$$
$$a_{2,1} = \kappa_1 + \kappa_2 \kappa_1^*$$
$$a_{2,2} = \kappa_2 \quad \text{and} \quad P_2 = P_1(1 - |\kappa_2|^2)$$

3. For a prediction-error filter of order $m=3$

$$a_{3,0} = 1$$
$$a_{3,1} = a_{2,1} + \kappa_3 \kappa_2^*$$
$$a_{3,2} = \kappa_2 + \kappa_3 a_{2,1}^*.$$

$a_{3,3} = K_3$ and $P_3 = P_2(1-|K_3|^2)$

※ Derive the Reflection Coefficients

Define the inner-product of two complex random variables with zero-mean, as

$$\langle x, y \rangle = E\{x^* y\},$$

Therefore $\|x\|^2 = \langle x, x \rangle = E\{|x|^2\} = Var(x).$

The predicted value can be written as

$$\hat{u}_m(n) = \hat{u}(n \mid \vec{u}_{n-1}) = \sum_{k=1}^{m} w_{f,k}^* \, u(n-k)$$

$$= - \sum_{k=1}^{m} a_{m,k}^* \, u(n-k)$$

The mean-square error is

$$\mathcal{E}_m = E\{|u(n) - \hat{u}_m(n)|^2\}$$

$$= \|u(n) - \hat{u}_m(n)\|^2.$$

$\mathcal{E}_m$ will be minimized when the principle of the orthogonality is satisfied, such that $\langle u(n-k), u(n) - \hat{u}_m(n) \rangle = 0$, $k = 1, 2, \cdots, m$

Thus, $\langle u(n-k), u(n) + \sum_{l=1}^{m} a_{m,l}^{*} u(n-l) \rangle = 0$

$$\Rightarrow \sum_{l=1}^{m} a_{m,l}^{*} \langle u(n-l), u(n-l) \rangle = - \langle u(n-k), u(n) \rangle$$

$$\Rightarrow \sum_{l=1}^{m} a_{m,l}^{*} E\{u^{*}(n-k)\, u(n-l)\} = - E\{u^{*}(n-k)\, u(n)\}$$

$$\Rightarrow \sum_{l=1}^{m} a_{m,l}^{*}\, r_u(k-l) = - r_u(k), \quad k = 1, 2, \dots, m$$

$\ldots\ldots (*)$

$$P_m = \min\{\mathcal{E}_m\} = \langle u(n) - \hat{u}_m(n), u(n) - \hat{u}_m(n) \rangle$$

$$= \langle u(n), u(n) - \hat{u}_m(n) \rangle - \underbrace{\langle \hat{u}_m(n), u(n) - \hat{u}_m(n) \rangle}_{=0}$$

$$= \langle u(n), u(n) \rangle - \langle u(n), \hat{u}_m(n) \rangle$$

$$= \langle u(n), u(n) \rangle + \sum_{k=1}^{m} a_{m,k}^{*} \langle u(n), u(n-k) \rangle$$

$$= r_u(0) + \sum_{k=1}^{m} a_{m,k}^{*}\, r_u(-k) \quad \ldots (**)$$

Eq $(*)$ is the well-known Yule-Walker Equations, or in the matrix form, it can be stated as follows:

$$\underbrace{\begin{bmatrix} r_u(0) & r_u(-1) & \cdots & r_u(-m+1) \\ r_u(1) & r_u(0) & \cdots & r_u(-m+2) \\ \vdots & \vdots & & \vdots \\ r_u(m-1) & r_u(m-2) & \cdots & r_u(0) \end{bmatrix}}_{\tilde{R}_u = \tilde{R}_u^H} \begin{bmatrix} a_{m,1}^* \\ a_{m,2}^* \\ \vdots \\ a_{m,m}^* \end{bmatrix} = - \begin{bmatrix} r_u(1) \\ r_u(2) \\ \vdots \\ r_u(m) \end{bmatrix}$$

Alternatively, from Eq (**), the augmented Yule-Walker Equations can be derived in the matrix form as follows:

$$\begin{bmatrix} r_u(0) & r_u(-1) & \cdots & r_u(-m) \\ r_u(1) & r_u(0) & \cdots & r_u(-m+1) \\ \vdots & \vdots & & \vdots \\ r_u(m) & r_u(m-1) & \cdots & r_u(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_{m,1}^* \\ \vdots \\ a_{m,m}^* \end{bmatrix} = \begin{bmatrix} P_m \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Take $m=1$,

$$\hat{u}_1(n) = -a_{1,1}\, u(n-1)$$

$$\langle u(n-1),\, u(n) - \hat{u}_1(n) \rangle = 0$$

$$\Rightarrow a_{1,1}^* = - \frac{\langle u(n-1),\, u(n) \rangle}{\langle u(n-1),\, u(n-1) \rangle}$$

$$\therefore \quad \hat{u}_1(n) = \frac{\langle u(n-1), u(n)\rangle}{\langle u(n-1), u(n-1)\rangle}\, u(n-1)$$

Denote $\quad \bar{e}_0^{\,b}(n-1) = \dfrac{u(n-1)}{\|u(n-1)\|}$, where

$$\|u(n-1)\| = \sqrt{\|u(n-1)\|^2} = \sqrt{\langle u(n-1), u(n-1)\rangle}$$

Therefore, $\quad \hat{u}_1(n) = \dfrac{\langle u(n-1), u(n)\rangle}{\|u(n-1)\|}\, \dfrac{u(n-1)}{\|u(n-1)\|}$

$$= \langle \bar{e}_0^{\,b}(n-1), u(n)\rangle\, \bar{e}_0^{\,b}(n-1).$$

It can be seen that the first-order optimal linear predictor is achieved by projecting $u(n)$ along the $u(n-1)$ "direction" where the "unit vector" along the $u(n-1)$ direction is $\bar{e}_0^{\,b}(n-1)$.

Now take $m = 2$,

$$\hat{u}_2(n) = -a_{2,1}^{*}\, u(n-1) - a_{2,2}^{*}\, u(n-2)$$

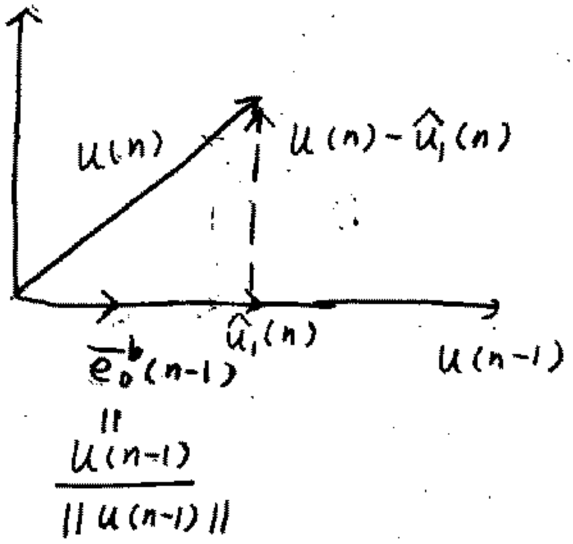
$= \hat{u}_1(n) +$ best prediction of $u(n)$
based on part of $u(n-2)$
in new orthogonal direction.

$\quad\quad \|$

$\langle \bar{e}_1^{\,b}(n-1), u(n)\rangle\, \bar{e}_1^{\,b}(n-1)$

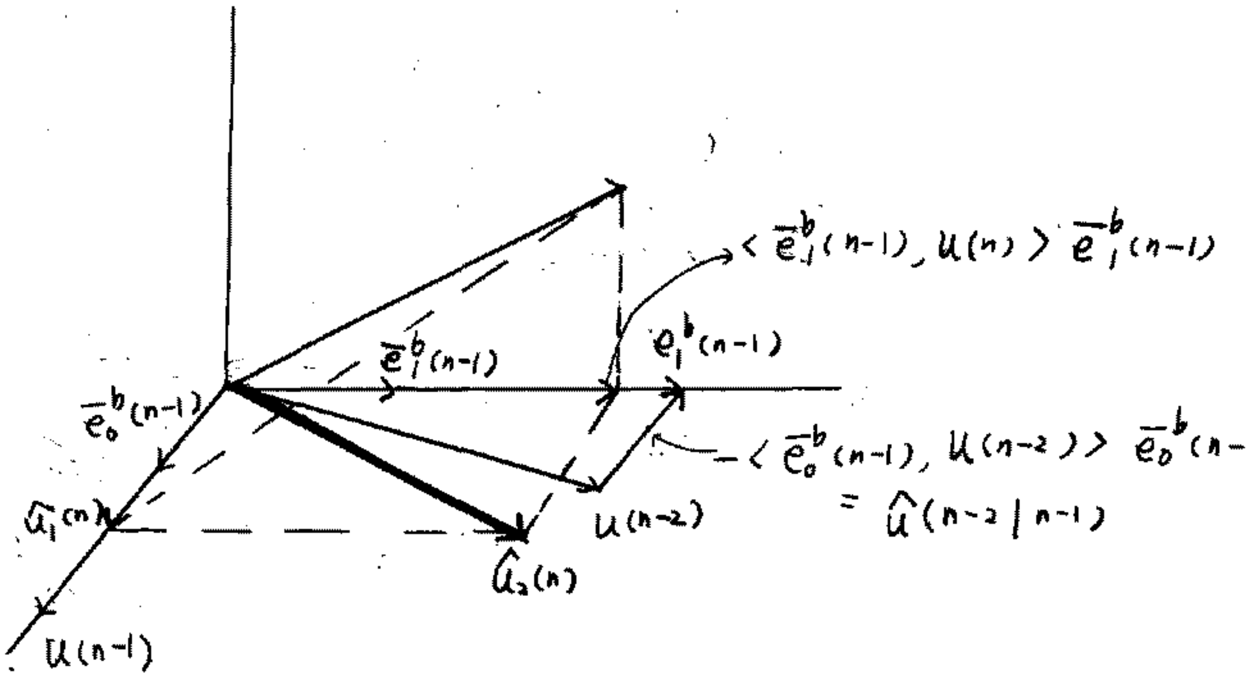
Graphically, the linear prediction can be depicted as follows:

$m = 1$



The optimal $\hat{u}_1(n)$ is the projected vector from $u(n)$ onto $\overline{e}_0^b(n-1)$

$m = 2$



From the figure above, it clearly shows that $\hat{u}_2(n)$ consists of two components, namely $\hat{u}_1(n)$, and $\langle \bar{e}_1^b(n-1), u(n) \rangle \, \bar{e}_1^b(n-1)$,

where $e_1^b(n-1) = u(n-2) - \hat{u}(n-2 \mid n-1)$

and $\bar{e}_1^b(n-1) = \dfrac{e_1^b(n-1)}{\| e_1^b(n-1) \|}$.

Thus,
$$\hat{u}_2(n) = \hat{u}_1(n) + \langle \bar{e}_1^b(n-1), u(n) \rangle \, \bar{e}_1^b(n-1)$$
$$= -a_{2,1}^* u(n-1) - a_{2,2}^* u(n-2)$$

By induction,
$$\hat{u}_m(n) = \hat{u}_{m-1}(n) + \langle \bar{e}_{m-1}^b(n-1), u(n) \rangle \times \bar{e}_{m-1}^b(n-1),$$

where $e_{m-1}^b(n-1)$ is the backward prediction error if $u(n-m)$ is predicted on the basis of $u(n-(m-1)), u(n-(m-2)), \ldots, u(n-1)$.

$$\hat{u}_m(n) = -\sum_{i=1}^{m-1} a_{m-1,i}^* u(n-i) + \frac{\langle e_{m-1}^b(n-1), u(n) \rangle}{\| e_{m-1}^b(n-1) \|^2} e_{m-1}^b($$

Let $$K_m \equiv - \frac{\left(\langle e^b_{m-1}(n-1), u(n)\rangle\right)^*}{\| e^b_{m-1}(n-1)\|^2}$$

Note: $e^b_{m-1}(n-1) \equiv b_{m-1}(n-1)$ as before !!

Since $\| b_{m-1}(n-1)\|^2 = \| f_{m-1}(n-1)\|^2 = \| f_{m-1}(n)\|^2$,

and $\langle e^b_{m-1}(n-1), u(n)\rangle = \langle b_{m-1}(n-1), u(n) + \sum_{i=1}^{m-1} a_{m-1,i}\, u(n-i)\rangle$

$$= \langle b_{m-1}(n-1), f_{m-1}(n)\rangle \quad (\langle b_{m-1}(n-1), u(n-i)\rangle = 0, \ \forall\, i = 1, 2, \cdots, m-1),$$

then $$K_m = - \frac{\left(\langle b_{m-1}(n-1), f_{m-1}(n)\rangle\right)^*}{\| f_{m-1}(n)\| \, \| b_{m-1}(n-1)\|}$$

$$= \frac{E\left[b_{m-1}(n-1)\, f^*_{m-1}(n)\right]}{\sqrt{E\left[|b_{m-1}(n-1)|^2\right] E\left[|f_{m-1}(n)|^2\right]}}$$

as defined before.

The m^{th} order predictor becomes

$$\hat{u}_m(n) = - \sum_{i=1}^{m-1} a^*_{m-1,i}\, u(n-i) + K^*_m\, b_{m-1}(n-1),$$

where $b_{m-1}(n-1) = u(n-m) - \left[- \sum_{i=0}^{K-2} c^*_{m-1,i}\, u(n-1-i)\right]$

Since $C_{m-1,m-1} = 1$, [illegible],

$$\hat{u}_m(n) = -\sum_{i=1}^{m-1} a^*_{m-1,i}\, u(n-i) - K^*_m \sum_{i=0}^{m-1} C^*_{m-1,i}\, u(n-1-i)$$

$$= -\sum_{i=1}^{m} a^*_{m,i}\, u(n-i).$$

$$\Rightarrow \quad C_{m-1,i} = a^*_{m-1,m-1-i}, \qquad i = 0, 1, 2, \cdots, m-1$$

According to the principle of the orthogonality,

$$\langle u(n-j), \sum_{i=0}^{m-1} C^*_{m-1,i}\, u(n-1-i) \rangle = 0, \quad j = 1, 2, \cdots, m-1$$

or $$\sum_{i=0}^{m-1} C^*_{m-1,i}\, r_u(j-i-1) = 0.$$

Let $l = m-1-i$, we obtain

$$\sum_{l=0}^{m-1} C^*_{m-1,m-1-l}\, r_u(j-m+l) = 0$$

Take the complex-conjugate, we have

$$\sum_{l=0}^{m-1} C_{m-1,m-1-l}\, r_u(m-j-l) = 0, \quad j = 1, 2, \cdots, m-1.$$

$$\Rightarrow \sum_{l=0}^{m-1} C_{m-1,m-1-l}\, r_u(m-l) = 0, \quad m = 1, 2, \cdots, m-1$$

$$\Rightarrow a^*_{m-1,l} = C_{m-1,m-1-l}, \quad l = 0, 1, 2, \cdots, m-1$$

$$\hat{u}_m(n) = -\sum_{i=1}^{m-1} a^*_{m-1,i}\, u(n-i) - k^*_m \Big(u(n-m) + \sum_{i=0}^{m-2} a_{m-1,\,m-1-i}\, u(n-1-i) \Big)$$

$$= -\sum_{i=1}^{m-1} \left(a^*_{m-1,i} + k^*_m a_{m-1,\,m-i} \right) u(n-i) - k^*_m u(n-m).$$

Since $\hat{u}_m(n) = -\sum_{i=1}^{m} a^*_{m,i}\, u(n-i)$,

therefore $a^*_{m,i} = a^*_{m-1,i} + k^*_m a_{m-1,\,m-i}$

or $a_{m,i} = \begin{cases} a_{m-1,i} + k_m a^*_{m-1,\,m-i}, & i=1,2,\cdots,m \\ k_m, & i = m \end{cases}$

3.4. Properties of Prediction-error Filters

Property 1. The corresponding set of reflection coefficients can be obtained based on the set of autocorrelation function $r(1), \ldots, r(M)$, such that $K_m = -\frac{1}{P_{m-1}} \sum_{k=0}^{m-1} a_{m-1}\, r(k-m)$.

Proof: From Eq. (3.47),

$$\Delta_{m-1} = \sum_{\ell=0}^{m-1} r(\ell-m)\, a_{m-1,\ell}$$

From Eq. (3.55),

$$\Delta_{m-1} = -k_m P_{m-1}.$$

Hence $$\sum_{\ell=0}^{m-1} a_{m-1,\ell}\, r(\ell-m) = -k_m P_{m-1}$$

Since $r(m) = r^*(-m)$ and $a_{m-1,\phi} = 1$,

we get $$k_m = -\frac{1}{P_{m-1}} \sum_{\ell=0}^{m-1} a_{m-1,\ell}\, r(\ell-m),$$

$$m = 1, 2, \cdots, M$$

Example: Provided P_ϕ, k_1, k_2 and k_3,

Compute $r(0)$, $r(1)$, $r(2)$ and $r(3)$.

Solution:

$$m=1 \Rightarrow k_1 = -\frac{1}{P_\phi} \underbrace{a_{\phi,\phi}}_{=1} r^*(1)$$

$$\Rightarrow r(1) = -P_0 k_1^*,$$

where $P_\phi = r(0)$

$$m=2 \Rightarrow \quad K_2 = -\frac{1}{P_1}\left(\underbrace{a_{1,0}}_{=1}\, r^*(2) + \underbrace{a_{1,1}}_{=K_1}\, r^*(1)\right).$$

$$\text{where} \quad P_1 = P_0\,(1-|K_1|^2)$$

$$\Rightarrow \quad r(2) = -P_1 K_2^* - r(1)\, K_1^*$$

$$m=3 \Rightarrow \quad K_3 = -\frac{1}{P_2}\left(\underbrace{a_{2,0}}_{=1}\, r^*(3) + a_{2,1}\, r^*(2) + \underbrace{a_{2,2}}_{=K_2}\, r^*(1)\right)$$

$$\text{where} \quad P_2 = P_1\,(1-|K_2|^2)$$

$$\Rightarrow \quad r(3) = -P_2 K_3^* - a_{2,1}^*\, r(2) - K_2^*\, r(1),$$

$$\text{where} \quad a_{2,1} = a_{1,1} + K_2\, a_{1,1}^*$$

$$= K_1 + K_2 K_1^*$$

Property 2. The transfer function of a foward prediction-error filter can be stated as follows:

$$H_{f,m}(z) = H_{f,m-1}(z) + K_m^*\, z^{-1} H_{b,m-1}(z),$$

$$\text{where} \quad H_{f,m}(z) = \sum_{k=0}^{m} a_{m,k}^*\, z^{-k} \quad \text{and}$$

$$H_{b,m-1}(z) = \sum_{k=0}^{m-1} c_{m-1,k}^*\, z^{-k}$$

Property 3. A forward prediction-error filter is minimum phase.

Proof: The mean-square forward prediction error is

$$\rho \equiv E\{|u(n)-\hat{u}(n)|^2\}$$

$$= E\left\{\left|\sum_{k=0}^{m} a_{m,k}\, u(n-k)\right|^2\right\}$$

The minimum prediction error power can be written as

$$P_m = \int_{-\frac{1}{2}}^{\frac{1}{2}} S_{uu}(f)\left|1+\sum_{k=1}^{m} a_{m,k}\, e^{-j2\pi fk}\right|^2 df$$

Let $S_a(e^{j2\pi f}) \equiv 1+\sum_{k=1}^{m} a_{m,k}\, e^{-j2\pi fk}$

$$S_a(z) = \prod_{j=1}^{m}(1-z_j z^{-1}), \text{ where } z_j$$

is the j^{th} zero of $S_a(z)$

If some z_i is outside the unit circle,

$$S_a(z) = (1-z_i z^{-1}) \underbrace{\prod_{\substack{j=1\\ j\neq i}}^{m}(1-z_j z^{-1})}_{S_a'(z)},$$

where $|z_i|>1$

Then $$P_m = \int_{-\frac{1}{2}}^{\frac{1}{2}} |1-z_i e^{-j2\pi f}|^2 |S_a'(e^{j2\pi f})|^2 \times S_{uu}(f)\, df$$

$$> \int_{-\frac{1}{2}}^{\frac{1}{2}} \left|1-\frac{1}{z_i^*}e^{-j2\pi f}\right|^2 |S_a'(e^{j2\pi f})|^2 S_{uu}(f)\, df$$

$\Rightarrow P_m$ is not minimum prediction error!

Therefore no such z_i, where $|z_i|>1$ can exist! Since $\tilde{R}_{m+1}$ is positive definite, no z_j can be on the unit circle either.

$\Rightarrow$ z_j must be within the unit-circle $\Rightarrow$ minimum phase.

Property 4. A backward prediction-error filter is maximum-phase.

Proof:

$$H_{f,m}(z) = \sum_{k=0}^{m} a_{m,k} z^{-k}$$

$$H^*_{f,m}\left(\frac{1}{z^*}\right) = z^m \sum_{k=0}^{m} a_{m,m-k} z^{-k}$$

$$= z^m H_{b,m}(z)$$

The zeros in $H_{f,m}(z)$ are assumed to be z_i, $i=1,2,\ldots,m$

Hence $H_{f,m}(z_i)=0$ or $H^*_{f,m}(z_i)=0$

Then, $H_{b,m}\left(\frac{1}{z_i^*}\right) = H^*_{f,m}(z_i) = 0$

$\therefore$ $H_{b,m}(z)$ has m zeros as $\frac{1}{z_i^*}$, $i=1,2,\ldots,m$.

Since $H_{f,m}(z)$ has minimum-phase zeros z_i, $H_{b,m}(z)$ will have all maximum phase zeros $\frac{1}{z_i^*}$.

Property 5. A forward prediction error filter is a whitening filter.

Proof: By orthogonality principle, it can be easily shown that

$$E\{f_m^*(n) f_m(n-k)\}$$

$$= \begin{cases} P_m, & k = 0 \\ 0, & k \neq 0 \end{cases}$$

3.6 Autoregressive Modeling of a Stationary random Process

We may view the operation of prediction-error filtering applied to a random process $u(n)$ as one of the autoregressive output response and the white process $f_M(n)$

is the excitation input. The AR modeling can be implemented in two stages:

1. Given $u(n), u(n-1) \cdots u(n-M)$, then estimate the optimal $a_{M,k}$, $k=0,1,\cdots M$, such that $f_M(n)$ is white. (Analysis)

2. Excite the AR filter with the determined $a_{M,k}$, $k=0,1,2,\cdots M$, obtained from stage 1, with arbitrary new white process $v(n)$ to replace $f_M(n)$ to generate the artificial AR outputs (Synthesis)

3.7. Cholesky Factorization

Since the optimal solution based on the Wiener-Hopf Equations would involve with the inverse of a large matrix $\tilde{R}$, an efficient computational algorithm can be stated here as the Cholesky Factorization

The backward prediction-error sequences can be expressed as:

$$b_0(n) = u(n)$$
$$b_1(n) = a_{1,1}\, u(n) + a_{1,0}\, u(n-1)$$
$$\vdots$$
$$\underline{b_M(n)} = a_{M,M}\, u(n) + a_{M,M-1}\, u(n-1) + \cdots + a_{M,0}\, u(\cdot$$

Hence,

$$\vec{b}(n) = \tilde{L}\, \vec{u}(n),$$

where $\vec{u}(n) = [u(n), u(n-1), \cdots u(n-M)]^T$,

$\vec{b}(n) = [b_0(n), b_1(n), \ldots, b_M(n)]^T$,

and

$$\tilde{L} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ a_{1,1} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{M,M} & a_{M,M-1} & \cdots & 1 \end{bmatrix}$$

Orthogonality of Backward Prediction Errors

It can be easily derived such that

$$E\{b_m(n)\, b_i^*(n)\} = \begin{cases} P_m, & i = m \\ 0, & i \neq m \end{cases}$$

※ Factorization of the inverse Correlation Matrix

$$E\left[\vec{b}(n)\,\vec{b}^H(n)\right] = E\left[\tilde{L}\,\vec{u}(n)\,\vec{u}^H(n)\,\tilde{L}^H\right]$$

$$= \tilde{L}\,E\left[\vec{u}(n)\,\vec{u}^H(n)\right]\tilde{L}^H$$

Denote $\tilde{D} = E\left[\vec{b}(n)\,\vec{b}^H(n)\right]$

$$= \tilde{L}\,\tilde{R}\,\tilde{L}^H = \begin{bmatrix} P_0 & 0 & \cdots & 0 \\ 0 & P_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & P_M \end{bmatrix}$$

$$= \mathrm{diag}\{P_0, P_1 \cdots P_M\}$$