Statistical Signal Processing
5. The Levinson–Durbin Recursion

Dr. Chau-Wai Wong

Electrical & Computer Engineering
North Carolina State University

Readings: Hayes §5.2; Haykin 4th Ed. §3.3

Contact: *chauwai.wong@ncsu.edu*. Updated: November 4, 2020.

Acknowledgment: ECE792-41 slides were adapted from ENEE630 slides developed by Profs. K.J. Ray Liu and Min Wu at the University of Maryland.

Complexity in Solving Linear Prediction

Recall Augmented Normal Equations for linear prediction:

$$\underline{\text{FLP}} \quad \mathbf{R}_{M+1}\underline{a}_M = \begin{bmatrix} P_M \\ \underline{0} \end{bmatrix} \qquad\qquad \underline{\text{BLP}} \quad \mathbf{R}_{M+1}\underline{a}_M^{B*} = \begin{bmatrix} \underline{0} \\ P_M \end{bmatrix}$$

As $\mathbf{R}_{M+1}$ is usually non-singular, $\underline{a}_M$ may be obtained by inverting $\mathbf{R}_{M+1}$, or Gaussian elimination for solving equation array:

$\Rightarrow$ Computational complexity $O(M^3)$.

Note that these two equations are equivalent. Why?

Exploiting Structures in Matrix and LP Problem

Complexity in solving a general linear equation array:

- Method 1: invert the matrix, e.g., compute determinant of $\mathbf{R}_{M+1}$ matrix and the adjacency matrices

  $\Rightarrow$ matrix inversion has $O(M^3)$ complexity

- Method 2: use Gaussian elimination

  $\Rightarrow$ approximately $M^3/3$ multiplication and division

By exploring the Toeplitz structure of the matrix, Levinson–Durbin recursion can reduce complexity to $O(M^2)$

- M steps of order recursion, each step has a linear complexity w.r.t. intermediate order
- Memory use: Gaussian elimination $O(M^2)$ for the matrix, vs. Levinson-Durbin $O(M)$ for the autocorrelation vector and model parameter vector.

Levinson–Durbin Recursion

The **Levinson–Durbin recursion** is an order-recursion to efficiently solve linear systems with Toeplitz matrices, e.g., Augmented N.E. For M steps of order recursion, each step has a linear complexity w.r.t. intermediate order.

Goal: To solve $\underline{a}_m$ from the Augmented N.E., $\mathbf{R}_{m+1}\underline{a}_m = \begin{bmatrix} P_m \\ \underline{0} \end{bmatrix}$, where $\mathbf{R}_{m+1}$ is Toeplitz.

Plan: For N.E. at order $m+1$, we target to create an order recursion from order m.

Creating Order Update

First, create auxiliary vectors $\begin{bmatrix} \underline{a}_{m-1} \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ \underline{a}_{m-1}^{B*} \end{bmatrix}$ using order-m vectors. Second, multiply from left using order-$(m+1)$ correlation matrix.

$$\text{\underline{FLP}} \quad \mathbf{R}_{m+1} \begin{bmatrix} \underline{a}_{m-1} \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{R}_m & \underline{r}_m^{B^*} \\ \underline{r}_m^{BT} & r(0) \end{bmatrix} \begin{bmatrix} \underline{a}_{m-1} \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{R}_m \underline{a}_{m-1} \\ \underline{r}_m^{BT} \underline{a}_{m-1} \end{bmatrix} = \begin{bmatrix} P_{m-1} \\ \underline{0}_{m-1} \\ \Delta_{m-1} \end{bmatrix}, \text{ where } \Delta_{m-1} \triangleq \underline{r}_m^{BT} \underline{a}_{m-1}. \quad (1a)$$

$$\text{\underline{BLP}} \quad \mathbf{R}_{m+1} \begin{bmatrix} 0 \\ \underline{a}_{m-1}^{B*} \end{bmatrix} = \begin{bmatrix} r(0) & \underline{r}^H \\ \underline{r} & \mathbf{R}_m \end{bmatrix} \begin{bmatrix} 0 \\ \underline{a}_{m-1}^{B*} \end{bmatrix}$$

$$= \begin{bmatrix} \underline{r}^H \underline{a}_{m-1}^{B*} \\ \mathbf{R}_m \underline{a}_{m-1}^{B*} \end{bmatrix} = \begin{bmatrix} \Delta_{m-1}^* \\ \underline{0}_{m-1} \\ P_{m-1} \end{bmatrix}. \quad (1b)$$

Third, poll these two equations together by (1a) + $\Gamma_m \times$ (1b):

Creating Order Update

$$\mathbf{R}_{m+1}\left(\begin{bmatrix} \underline{a}_{m-1} \\ 0 \end{bmatrix} + \Gamma_m \begin{bmatrix} 0 \\ \underline{a}_{m-1}^{B*} \end{bmatrix}\right) = \begin{bmatrix} P_{m-1} \\ \underline{0}_{m-1} \\ \Delta_{m-1} \end{bmatrix} + \Gamma_m \begin{bmatrix} \Delta_{m-1}^* \\ \underline{0}_{m-1} \\ P_{m-1} \end{bmatrix}.$$

Fourth, compare it to the order-$(m+1)$ N.E., $\mathbf{R}_{m+1}\underline{a}_m = \begin{bmatrix} P_m \\ \underline{0} \end{bmatrix}$. To obtain order update relationship, we need:

$$\begin{cases} \underline{a}_m \overset{\text{set}}{=} \begin{bmatrix} \underline{a}_{m-1} \\ 0 \end{bmatrix} + \Gamma_m \begin{bmatrix} 0 \\ \underline{a}_{m-1}^{B*} \end{bmatrix}, \\ \begin{bmatrix} P_m \\ \underline{0}_m \end{bmatrix} \overset{\text{set}}{=} \begin{bmatrix} P_{m-1} \\ \underline{0}_{m-1} \\ \Delta_{m-1} \end{bmatrix} + \Gamma_m \begin{bmatrix} \Delta_{m-1}^* \\ \underline{0}_{m-1} \\ P_{m-1} \end{bmatrix}. \end{cases}$$

Solving for Γ_m That Allows Order Update

$$\Rightarrow \begin{cases} P_m = P_{m-1} + \Gamma_m \Delta_{m-1}^* \\ 0 = \Delta_{m-1} + \Gamma_m P_{m-1} \end{cases}$$

$$\Gamma_m = -\frac{\Delta_{m-1}}{P_{m-1}} (= a_{m,m})$$

$$P_m = P_{m-1}\left(1 - |\Gamma_m|^2\right)$$

To ensure the prediction MSE $P_m \geq 0$, we require $|\Gamma_m|^2 \leq 1$.

P_m is non-increasing as we increase the order of the predictor, i.e., $P_m \leq P_{m-1}$, $\forall m > 0$.

Order Update Summary: Two Viewpoints of LD Recursion

Denote $\underline{a}_m \in \mathbb{C}^{(m+1)\times 1}$ as the tap weight vector of a forward-prediction-error filter of order $m = 0, ..., M$.

$a_{m-1,0} = 1$, $a_{m-1,m} \triangleq 0$, $a_{m,m} = \Gamma_m$ (*reflection coefficient*)

Order Update—Forward Prediction Viewpoint

$$a_{m,k} = a_{m-1,k} + \Gamma_m \, a^*_{m-1,m-k}, \; k = 0, 1, \ldots, m$$

Vector form: $$\underline{a}_m = \begin{bmatrix} \underline{a}_{m-1} \\ 0 \end{bmatrix} + \Gamma_m \begin{bmatrix} 0 \\ \underline{a}^{B*}_{m-1} \end{bmatrix} \quad (**)$$

Order Update—Backward Prediction Viewpoint

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

Vector form: $$\underline{a}^{B*}_m = \begin{bmatrix} 0 \\ \underline{a}^{B*}_{m-1} \end{bmatrix} + \Gamma^*_m \begin{bmatrix} \underline{a}_{m-1} \\ 0 \end{bmatrix}$$

(4) Reflection Coefficient Γ_m

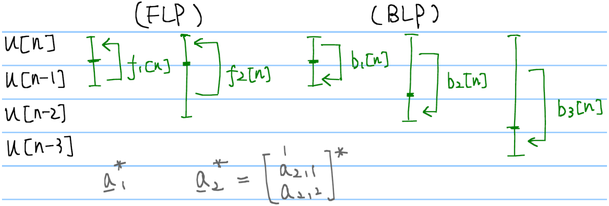
Let $P_0 = r(0)$ as the initial estimation error has power equal to the signal power, i.e., no regression is applied, we have

$$P_M = P_0 \cdot \prod_{m=1}^{M} (1 - |\Gamma_m|^2).$$

Question: Under what situation is $\Gamma_m = 0$?
i.e., increasing order won't reduce error.

Consider a process with Markovian-like property in 2nd order statistic sense (e.g. AR process) s.t. info of further past is contained in k recent samples.

Recall: Forward and Backward Prediction Errors



- $f_m[n] = u[n] - \hat{u}[n] = \underline{a}_m^H \underbrace{\underline{u}[n]}_{(m+1)\times 1}$
- $b_m[n] = u[n-m] - \hat{u}[n-m] = \underline{a}_m^{B,T} \underline{u}[n]$

(5) About Δ_m

One can show that the cross-correlation of <u>BLP error</u> and <u>FLP error</u> $\mathbb{E}\left[b_{m-1}[n-1]f^*_{m-1}[n]\right]$ is equal to Δ_{m-1}.

(Derive from the definition $\Delta_{m-1} \triangleq \underline{r}_m^{BT}\underline{a}_{m-1}$, and use definitions of $b_{m-1}[n-1], f^*_{m-1}[n]$ and orthogonality principle.)

Thus the reflection coefficient can be written as

$$\Gamma_m = -\frac{\Delta_{m-1}}{P_{m-1}} = -\frac{\mathbb{E}\left[b_{m-1}[n-1]f^*_{m-1}[n]\right]}{\mathbb{E}\left[|f_{m-1}[n]|^2\right]}$$

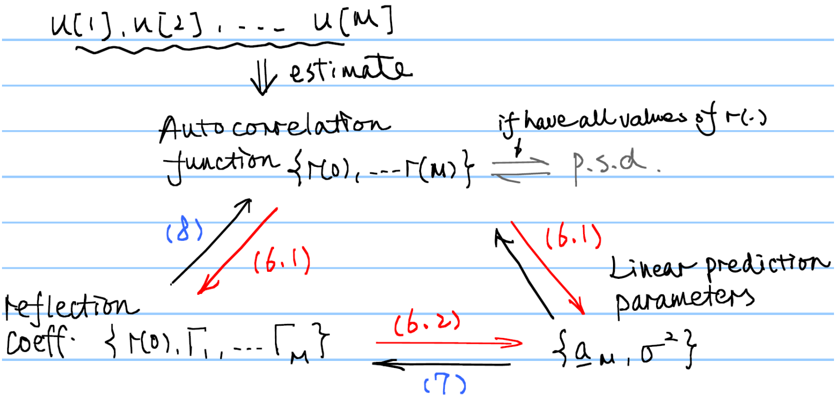
which is also the negative *partial correlation coefficient*.

Note: for the 0th order predictor, use the mean value, i.e., zero, as the estimate, s.t. $f_0[n] = u[n] = b_0[n]$,

$\therefore \Delta_0 = \mathbb{E}\left[b_0[n-1]f^*_0[n]\right] = \mathbb{E}\left[u[n-1]u^*[n]\right] = r(-1) = r^*(1)$

Preview: Relations of w.s.s and LP Parameters

For any w.s.s. process $\{u[n]\}$:



(6) Computing $\underline{a}_M$ and P_M by Forward Recursion

Case 1 : If we know the autocorrelation function $r(\cdot)$:

① $\Delta_0 = r(-1)$, $P_0 = r(0)$

② for $m = 1, \ldots M$ (order recursion)

$\quad \Gamma_m = -\frac{\Delta_{m-1}}{P_{m-1}}$

$\quad$ for $k = 1, \ldots m$ (diff predictor parameters for order-m)

$\quad\quad a_{m,k} = a_{m-1,k} + \Gamma_m a^*_{m-1,m-k}$

$\quad\quad$ (where $a_{m-1,0} = 1$; $a_{m-1,m} = 0$)

$\quad \Delta_m = r^{BT}_{m+1} \underline{a}_m$

$\quad P_m = P_{m-1}(1 - |\Gamma_m|^2)$

- # of iterations $= \sum_{m=1}^{M} m = \frac{M(M+1)}{2}$, comp. complexity is $O(M^2)$
- $r(k)$ may be estimated from time average of one realization of $\{u[n]\}$:

  $\hat{r}(k) = \frac{1}{N-k} \sum_{n=k+1}^{N} u[n]u^*[n-k], \; k = 0, 1, \ldots, M$

  (recall correlation ergodicity)

(6) Computing $\underline{a}_M$ and P_M by Forward Recursion

Case 2 : If we know Γ_1, Γ_2, ..., Γ_M and $P_0 = r(0)$, we can carry out the recursion for $m = 1, 2, \ldots, M$:

$$\begin{cases} a_{m,k} = a_{m-1,k} + \Gamma_m a^*_{m-1,m-k}, \ k = 1, \ldots, m \\ P_m = P_{m-1}\left(1 - |\Gamma_m|^2\right) \end{cases}$$

Note: $a_{m,m} = a_{m-1,m} + \Gamma_m a^*_{m-1,0} = 0 + \Gamma_m \cdot 1 = \Gamma_m$

(7) Inverse Form of Levinson-Durbin Recursion

Given the tap-weights $\underline{a}_M$, find the reflection coefficients $\Gamma_1, \Gamma_2, \ldots, \Gamma_M$:

Recall: $\begin{cases} \text{(FP) } a_{m,k} = a_{m-1,k} + \Gamma_m\, a^*_{m-1,m-k}, \ \ k = 0, \ldots, m \\ \text{(BP) } a^*_{m,m-k} = a^*_{m-1,m-k} + \Gamma^*_m\, a_{m-1,k}, \ \ a_{m,m} = \Gamma_m \end{cases}$

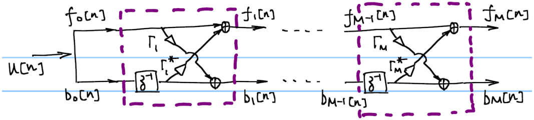
Multiply (BP) by Γ_m and subtract from (FP):

$$a_{m-1,k} = \frac{a_{m,k} - \Gamma_m a^*_{m,m-k}}{1 - |\Gamma_m|^2} = \frac{a_{m,k} - a_{m,m} a^*_{m,m-k}}{1 - |a_{m,m}|^2}, k = 0, \ldots, m-1$$

$\Rightarrow \Gamma_m = a_{m,m}, \ \Gamma_{m-1} = a_{m-1,m-1}, \ldots,$ i.e., From $\underline{a}_M \Rightarrow \underline{a}_m \Rightarrow \Gamma_m$

iterate with $m = M-1, M-2, \ldots$ *to lower order*

Lattice structure:



(8) Autocorrelation Function & Reflection Coefficients

Recall: The 2nd-order statistics of a stationary time series can be represented in terms of autocorrelation function $r(k)$, or equivalently the power spectral density by taking DTFT.

Another way is to use $\{r(0), \Gamma_1, \Gamma_2, \ldots, \Gamma_M\}$.

To find the relation between them, recall:

$\Delta_{m-1} \triangleq \underline{r}_m^{BT} \underline{a}_{m-1} = \sum_{k=0}^{M-1} a_{m-1,k} r(-m+k)$ and $\Gamma_m = -\frac{\Delta_{m-1}}{P_{m-1}}$

$\Rightarrow -\Gamma_m P_{m-1} = \sum_{k=0}^{m-1} a_{m-1,k} r(k-m)$, where $a_{m-1,0} = 1$.

(8) Autocorrelation Function & Reflection Coefficients

1. $r(m) = r^*(-m) = -\Gamma_m^* P_{m-1} - \sum_{k=1}^{m-1} a_{m-1,k}^* r(m-k)$

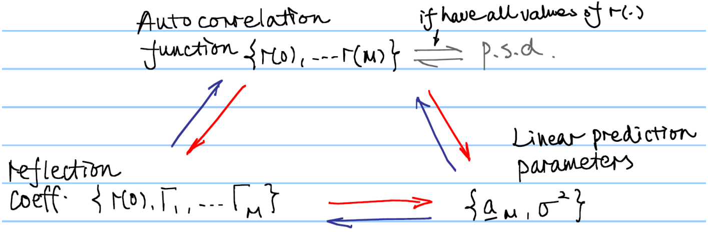
   $r(1), \ldots, r(M)$ can be generated iteratively. Note that $\underline{a}_m$ can be found using $r(0), \Gamma_1, \Gamma_2, \ldots, \Gamma_M$ by (6.2).

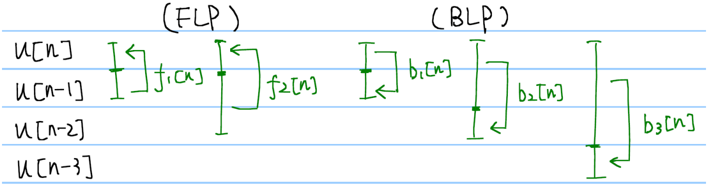
2. Recall if $r(0), \ldots, r(M)$ are given, we can get $\underline{a}_m$ by (6.1). So $\Gamma_1, \ldots, \Gamma_M$ can be obtained iteratively: $\Gamma_m = a_{m,m}$.

3. These facts imply that the reflection coefficients $\{\Gamma_k\}$ can uniquely represent the 2nd-order statistics of a w.s.s. process.

Summary

Statistical representation of w.s.s. process





Detailed Derivations/Examples

Example of Forward Recursion Case 2

e.g. (case 2). Given $\Gamma_1, \Gamma_2, \Gamma_3$ and $P(0)$, find $\underline{a}_3$ and P_3 of a prediction-error filter of order 3.

(0) $P_0 = r(0)$

(1) $m=1$: $a_{1,0} = 1$; $a_{1,1} = \Gamma_1$; $a_{1,2} = 0$; $P_1 = P_0(1 - |\Gamma_1|^2)$

(2) $m=2$: $a_{2,0} = 1$; $a_{2,1} = a_{1,1} + \Gamma_2 a_{1,1}^* = \Gamma_1 + \Gamma_2 \cdot \Gamma_1^*$ (used in §2.5.4. for inverse filtering)

$a_{2,2} = \Gamma_2$

$P_2 = P_1(1 - |\Gamma_2|^2)$

(3) $m=3$: $a_{3,0} = 1$; $a_{3,1} = a_{2,1} + \Gamma_3 a_{2,2}^* = \Gamma_1 + \Gamma_2 \Gamma_1^* + \Gamma_3 \cdot \Gamma_2^*$

$a_{3,2} = a_{2,2} + \Gamma_3 a_{2,1}^* = \Gamma_2 + \Gamma_3 \Gamma_1^* + \Gamma_1 \Gamma_2^* \Gamma_3$

$a_{3,3} = \Gamma_3$

$P_3 = P_2(1 - |\Gamma_3|^2)$

Proof for Δ_{m-1} Property

Proof:

$$\Delta_{m-1} = \Gamma_m^{BT} a_{m-1} = [r(-m), \dots r(-1)]\, a_{m-1}$$
$$= E[u^*[n]\, u_m^{BT}[n-1]]\, a_{m-1}$$
$$= E[u^*[n]\, u_m^{BT}[n-1]\, a_{m-1}]$$
$$= E[u^*[n]\, (u_m^{T}[n-1]\, a_{m-1}^{B})]$$
$$= E[u^*[n]\, b_{m-1}[n-1]]$$
$$= E[f^*_{m-1}[n]\, b_{m-1}[n-1]]$$

Recall

① $\Gamma_M = \begin{bmatrix} r(-1) \\ \vdots \\ r(-M) \end{bmatrix}$

② $r(-k) = E[u[n-k]\, u^*[n]] = (E[u[n]\, u^*[n-k]])^*$

③ $u_m[n-1] = \begin{bmatrix} u[n-1] \\ \vdots \\ u[n-m] \end{bmatrix}$

④ $b_{m-1}[n-1] = \sum_{k=0}^{m-1} a_{m-1,m-1-k}\, u[n-1-k] = a_{m-1}^{B}\, u_m[n-1]$

⑤ $a_{m-1,0} = 1$

$$f_{m-1}[n] = \sum_{k=0}^{m-1} a^*_{m-1,k}\, u[n-k] = u[n] + \sum_{k=1}^{m-1} a^*_{m-1,k}\, u[n-k]$$

⑥ $b_{m-1}[n-1] \perp \{u[n-1], \dots u[n-m+1]\}$

Haykin's 4th Ed. (P152)

* partial correlation (PARCOR) coeff. between $f_{m-1}[n]$ and $b_{m-1}[n-1]$:

$$\rho_m \triangleq \frac{E[b_{m-1}[n-1]\, f^*_{m-1}[n]]}{\left(E[|b_{m-1}[n-1]|^2]\, E[|f_{m-1}[n]|^2]\right)^{1/2}} \overset{\text{for w.s.s.}}{=} \frac{\Delta_{m-1}}{P_{m-1}} = -\Gamma_m$$

Recall $E[|f_m[n]|^2] = E[|b_m[n]|^2] = P_m$