Least-squares lattice predictor.

Consider the predictor of the form

$$\hat{x}(n) = \sum_{i=1}^{N} a_{N,i} \; x(n-i)$$

where the coefficients of the predictor are selected to minimize the cost function

$$J = \frac{1}{P} \sum_{k=1}^{P} \left(x(n) - \hat{x}(n) \right)^2$$

Clearly, the optimal solution is given by

$$A = \begin{bmatrix} a_{N,1} \\ a_{N,2} \\ \vdots \\ a_{N,N} \end{bmatrix} = \hat{R}_{XX}^{-1} \hat{P}_{xX}$$

Where

$$\hat{R}_{XX} = \frac{1}{P} \sum_{k=1}^{P} X(k) X^T(k)$$

and

$$\hat{P}_{xX} = \frac{1}{P} \sum_{k=1}^{P} x(k) X(k)$$

with $X(k) = \begin{bmatrix} x(k-1) \\ \vdots \\ x(k-N) \end{bmatrix}$ $\Leftarrow$ note that the first element is $x(k-1)$

The orthogonality is in the sense that

$$\frac{1}{P} \sum_{k=1}^{P} X(k) \, x(k) = 0$$

To derive the lattice predictor, we recognize that the orthogonal basis set we seek ~~has~~ are the backward prediction errors of the form

$$b_j(n) = x(n-j) - \sum_{i=0}^{j-1} g_{j,i}\, x(n-i)$$

$$; j = 0, 1, 2, \ldots, N-1$$

where the optimal backward predictor coefficients are selected as

$$G_j = \begin{bmatrix} g_{j,0} \\ g_{j,1} \\ \vdots \\ g_{j,j-1} \end{bmatrix} = \hat{R}_j^{-1} \hat{P}_j$$

where

$$\hat{R}_j = \frac{1}{P} \sum_{k=1}^{P} X_j(k) X_j^T(k)$$

and

$$\hat{P}_j = \frac{1}{P} \sum_{k=1}^{P} X_j(k)\, x(k)$$

with $X_j(k) = \left[\, x(k-1) \;\; x(k-2) \;\; \cdots \;\; x(k-j) \,\right]^T$

Note that the elements of the ~~cross~~ autocorrelation matrix are of the form

$$\hat{r}_{xx}(i,j) = \frac{1}{P} \sum_{k=1}^{P} x(k-i)\, x(k-j)$$

If we change the definitions of the cost function to sum upto $k = P+N$ instead of $k = P$ as above, the resulting

least-squares autocorrelation matrix will be Toeplitz. We will assume that this is the case. Therefore, only the lag value matters; i.e., $\underbrace{\hat{r}_{xx}(i,j)}_{\text{(we can define)}} = \hat{r}_{xx}(i-j)$.
We will update the backward prediction errors and the coefficients of the backward predictor filters in an order-recursive manner.

For order 0

$$b_0(n) = x(n)$$

For order 1

$$b_1(n) = x(n-1) - g_{1,0}\, x(n)$$

with

$$g_{1,0} = \frac{\hat{r}_{xx}(1)}{\hat{r}_{xx}(0)}$$

We know from prior work that we need the forward prediction errors to update the backward prediction errors.

Direct evaluation of the forward predictor and backward predictor coefficients, and making use of the Toeplitz nature of the autocorrelation matrices involved in each problem will show that the following symmetry holds: (The steps are identical to what we did for the MMSE case).

$$a_{m,i} = g_{m,m-i} \quad ; \; i = 1,2,\ldots,m.$$

~~For the derivation of~~ Furthermore, the least-squares value of forward and backward errors of each order can be shown to be identical.

For the derivation of the m th ~~order~~ stage of the lattice predictor, we assume that we have the following available to us:

(1) $\hat{r}_{xx}(j)$; $j = 0, 1, \cdots, m$

(2) $a_{m-1,j}$; $j = 1, 2, 3, \cdots, j$ (Note that $g_{m-1,j} = a_{m-1,m-1-j}$)

(3) σ_{j-1}^2, the least-squares value of the $j-1$ th order forward or backward prediction error signal.

The m-th order forward prediction error is given by

$$f_m(n) = x(n) - \sum_{j=1}^{m} a_{m,j}\, x(n-j)$$

$$= \underbrace{x(n) - \sum_{j=1}^{m-1} a_{m-1,j}\, x(n-j)}_{\text{error in estimate using } m-1 \text{ samples}} - \text{estimate of } x(n) \text{ using new information in } x(n-m)$$

The new information in $x(n-m)$ is the error in estimating $x(n-m)$ using $x(n-1), x(n-2), \cdots, x(n-m+1)$, i.e., the $m-1$ th order backward prediction error delayed by one sample, i.e., $b_{m-1}(n-1)$

The coefficient of the new information is

$$\rho_m^{(f)} = \frac{\frac{1}{P}\sum_{k=1}^{P} b_{m-1}(k-1)\, x(k)}{\sigma_{m-1}^2}$$

$$= \frac{\frac{1}{P}\sum_{k=1}^{P}\left(x(k-m) - \sum_{j=0}^{m-2} g_{m-1,j}\, x(k-1-j)\right)x(k)}{\sigma_{m-1}^2}$$

$$= \frac{\hat{r}_{xx}(m) - \sum_{j=0}^{m-2} g_{m-1,j}\, \hat{r}_{xx}(j+1)}{\sigma_{m-1}^2}$$

Use Symmetry $\Longrightarrow$

$$= \frac{\hat{r}_{xx}(m) - \sum_{j=0}^{m-2} a_{m-1,m-1-j}\, \hat{r}_{xx}(j+1)}{\sigma_{m-1}^2}$$

$$= \frac{\hat{r}_{xx}(m) - \sum_{j=1}^{m-1} a_{m-1,m-j}\, \hat{r}_{xx}(j)}{\sigma_{m-1}^2}$$

For the next stage, we need to update σ_m^2. We will do this later. Before that, let us update the backward prediction error also. Working in a manner similar to before,

$$b_m(n) = x(n-m) - \sum_{j=0}^{m-1} g_{m,j}\, x(n-j)$$

$$= \underbrace{x(n-m) - \sum_{j=0}^{m-1} g_{m-1,j}\, x(n-1-j)}$$

error in estimate of $x(n-m)$
using $x(n-1), x(n-2), \cdots x(n-m+1)$; i.e., $b_{m-1}(n-1)$

— error in estimating $x(n-m)$
using new information in $x(n)$.

The new information is the forward prediction error of order $m-1$.

$$\rho_m^{(b)} = \frac{\frac{1}{P}\sum_{k=1}^{P}\left(x(k-m)\right)\left(x(k) - \sum_{j=1}^{m-1} a_{m-1,j}\, x(k-j)\right)}{\sigma_{m-1}^2}$$

$$= \frac{\hat{r}_{xx}(m) - \sum_{j=1}^{m-1} a_{m-1,j}\, \hat{r}_{xx}(m-j)}{\sigma_{m-1}^2}$$

$$= \frac{\hat{r}_{xx}(m) - \sum_{l=1}^{m-1} a_{m-1,m-l}\, \hat{r}_{xx}(l)}{\sigma_{m-1}^2}$$

$$= \rho_m^{(f)}$$

Since both reflection coefficients are identical, we will just use ρ_m to denote them both.

We will now update the direct form predictor coefficients $a_{m,j}$; $j = 1, 2, \cdots, m$.

From earlier derivations,

$$f_m(n) = x(n) - \sum_{j=1}^{m} a_{m,j}\, x(n-m) \qquad \Leftarrow \text{LINE 1}$$

$$= x(n) - \sum_{j=1}^{m-1} a_{m-1,j}\, x(n-m) - \rho_m\, b_{m-1}(n-1)$$

$$= x(n) - \sum_{j=1}^{m-1} a_{m-1,j}\, x(n-m) - \rho_m \left(x(n-m) - \sum_{j=0}^{m-2} g_{m-1,j}\, x(n-1-j) \right)$$

$$= x(n) - \sum_{j=1}^{m-1} a_{m-1,j}\, x(n-m) - \rho_m \left\{ x(n-m) - \sum_{j=0}^{m-2} a_{m-1,m-1-j}\, x(n-1-j) \right\}$$

$$= x(n) - \sum_{j=1}^{m-1} a_{m-1,j}\, x(n-j) - \rho_m\, x(n-m) + \sum_{j=1}^{m-1} \rho_m\, a_{m-1,m-j}\, x(n-j)$$

$\Leftarrow$ LAST LINE

Since the equalities are true for any signal, we can equate the coefficients of identical samples on both sides. This gives

$$a_{m,m} = \rho_m$$

$$a_{m,j} = a_{m-1,j} \cancel{+} - \rho_m \, a_{m-1,m-j} \; ; \; j = 1, 2, \cdots, m-1$$

Now we have to update the least-squares error values.

$$\sigma_m^2 = \frac{1}{P} \sum_{k=1}^{P} f_m^2(k)$$

$$= \frac{1}{P} \sum_{k=1}^{P} \Bigg(\overbrace{x(k) - \sum_{j=1}^{m-1} a_{m-1,j} \, x(k-j)}^{f_{m-1}(n)} - \rho_m \, b_{m-1}(k-1) \Bigg)^2$$

$$= \frac{1}{P} \sum_{k=1}^{P} \Bigg(\overbrace{f_{m-1}^2(n)}^{\text{Will add upto } \sigma_m^2} + \rho_m^2 \, \overbrace{b_{m-1}^2(n-1)}^{\text{Will add upto } \sigma_{m-1}^2} - 2\rho_m \, \overbrace{x(k) \, b_{m-1}(k-1)} - 2\rho_m \sum_{j=1}^{m-1} a_{m-1,j} \, x(k-j) \, b_{m-1}(k-1) \Bigg)$$

Cross-correlation between $x(n)$ and $b_{m-1}(n-1)$!!
$= \rho_m \cdot \sigma_{m-1}^2$

Orthogonal signals; Will add up to zero.

Thus
$$\sigma_m^2 = \sigma_{m-1}^2 + \rho_m^2 \sigma_{m-1}^2 - 2\rho_m^2 \sigma_{m-1}^2$$
$$= \sigma_{m-1}^2 \left(1 - \rho_m^2\right)$$

We now have all the required results. The resulting Levinson-Durbin algorithm for the least-squares cost function is

Initialization.

$$\rho_1 = \frac{\hat{r}_{xx}(1)}{\hat{r}_{xx}(0)}$$

$$a_{1,1} = \rho_1$$

$$\sigma_1^2 = \hat{r}_{xx}(0)\left(1-\rho_1^2\right)$$

for the m th stage

$$\rho_m = \frac{\hat{r}_{xx}(m) - \sum_{\ell=1}^{m-1} a_{m-1,\ell}\, \hat{r}_{xx}(m-\ell)}{\sigma_{m-1}^2}$$

$$a_{m,m} = \rho_m$$

$$a_{m,j} = a_{m-1,j-1} - \rho_m a_{m-1,m-j}$$

$$; j = 1, 2, \ldots, m-1$$

$$\sigma_m^2 = \sigma_{m-1}^2\left(1-\rho_m^2\right)$$

(I just noticed that the indices of the "a" coefficients are reversed from those in the book!).