**Autocorrelation and related questions**

(Vaidyanathan P. 2008. The Theory of Linear Prediction)

**Part I. Linear prediction basic definition**

Let x(n) be a wide sense stationary (WSS) random process. Suppose we wish to predict the value of the sample x(n) using a linear combination of N most recent past samples. The estimate has the form

The integer N is called the prediction order.

Error is,

We denote the *mean squared error* as eps*forward*:

eps*forward* = *E*[|e*forward*(n)|2].

***Note***: The optimum predictor (i.e., the optimum set of coefficients ai) is the one that minimizes this mean squared value.

***Note***: Linear prediction essentially converts the signal x(n) into the set of N numbers { ai} and the error signal e*forward*(n).

***Note***: The optimal value of a*i* should be such that the error e*forward*(n) is orthogonal to x(n – i), that is, *E*[e(n)x\*(*n – i*)] = 0.

Define *R*(*k*) to be the autocorrelation sequence of the WSS process x(n), that is,

*R*(*k*) = *E*[x(n)x\*(*n – k*)].

Using the fact that *R*(*k*) = *R*(*–k*), expression

transform to

In matrix notation:

**Ra** = – **r**

where **R** is the correlation matrix of the random vector **x**; **a** is prediction coefficients; **r** is cross-correlation vector (its *i*-th element represents the correlation between the random variable x and observations x*i*. If the correlation matrix **R** is nonsingular, we can solve for **a** and obtain:

**a** = –**R**–**1r** (optimal linear estimator)

Thus, we can find a unique set of optimal predictor coefficients a*i*, as long as the NxN matrix **R** is nonsingular. Note that the matrix **R** is Toeplitz.

**Part II. Properties of autocorrelation matrix**

*A filtering interpretation of eigenvalues*

Consider an FIR filter, *V*(z) = *v*0 + *v*1z-1 + … + *v*N-1z-(N-1), with input x(n).

Its output can be expressed as, y(n) = *v*0x(n)+ *v*1x(n-1)+ … + *v*N-1x(n-N+1) = **v**Tx(n).

The mean square value of y(n) is *E*[|**v**T**x**(n)|2] = **v**T*E*[|**x**(n) **x**T (n)|] **v** = **v**T**Rv**.

That is, given the Toeplitz autocorrelation matrix **R**, the quadratic form **v**T**Rv** is the mean square value of the output of the FIR filter *V*(z), with input signal x(n).

**v**T**Rv** = λ

Thus, any eigenvalue of **R** can be regarded as the mean square value of the output y(n) for appropriate choice of the unit-energy filter *V*(z) driven by x(n).

***Note***: If the autocorrelation matrix is singular, then the corresponding random variables are linearly dependent.

***Note***: In terms of frequency domain, for a *white process*, the power spectrum is constant, whereas for a *fully predictable process*, the power spectrum can only have impulses - it cannot have any smooth component.

Note: Determinant of autocorrelation matrix det(**R**) = epsN-1\*epsN-2\*…\*eps0.

**Part III. Estimation of autocorrelation matrix**

*The autocorrelation method*

In this method, we define the truncated version of measured data, xL(n) = x(n), 0 ≤ n ≤ L-1; xL(n) = 0, *outside*.

If we have L = 5 and wish to estimate the 3x3 autocorrelation matrix **R3**, the computation is equivalent to defining the data matrix

and forming the estimate of **R3** using, **R3** = (**X**T**X**)\*.

*The covariance method*

In a variation called the “covariance method” the data matrix **X** is formed differently:

and the autocorrelation estimated as **R3** = (**X**T**X**)\*.

In the *covariance method*, each *R*(*k*) is an average of N possibly nonzero samples of data, unlike the *autocorrelation method*, where the number of samples used in the estimate of *R*(*k*) decreases as *k* increases. So, the estimates, in general, tend to be better than in the autocorrelation method, but the matrix **RN** in this case is not necessarily Toeplitz. So, Levinson’s recursion cannot be applied for solving the normal equations, and we have to solve the normal equations directly.

**Part IV. Backward prediction**

Let x(n) represent a WSS random process as usual. A backward linear predictor for this process estimates the sample x(n – N – 1) based on the “future values” x(n – 1), …, x(n – N). The predicted value is a linear combination of the form

and the prediction error is

***Note***: both x*forward* (n) and x*backward* (n – N – 1) are based on the same set of N measurements.

Figure 1 demonstrates the difference between the forward and backward predictors.

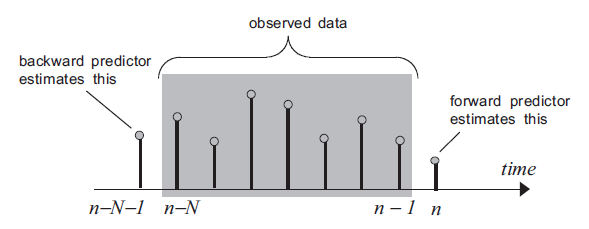


Fig. 1. Comparison of the forward and backward predictors.

**Part V. Autoregressive process**

A WSS random process w(n) is said to be autoregressive (AR) if it can be generated by using the recursive difference equation

Where

1. e(n) is a zero-mean white WSS process, and
2. the polynomial *D*(z)=1+sum(d*i*\*z–1); *i* = 1,N has all zeros inside the unit circle.

In other words, we can generate w(n) as the output of a causal, stable, all-pole IIR filter 1/*D*(z) in response to *white input* (Fig. 2).

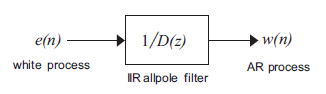


Fig. 2. Generation of an AR process from white noise and an all-pole IIR filter.

If the error *stalls*, that is, eps*mforward* does not decrease anymore as *m* increases beyond some value *N*, then e(n) is white (assuming x(n) has zero mean). Thus, the stalling phenomenon implies that x(n) is AR(N). {Stalling i.e. that increasing of prediction order does not lead to decreasing the error}.

*Shortly*: eps*1* ≥ eps*2* ≥ … ≥ eps*N-1* = e*m*, m >N; then x(n) is AR(N).

Figure 3 shows difference between AR and non-AR processes.

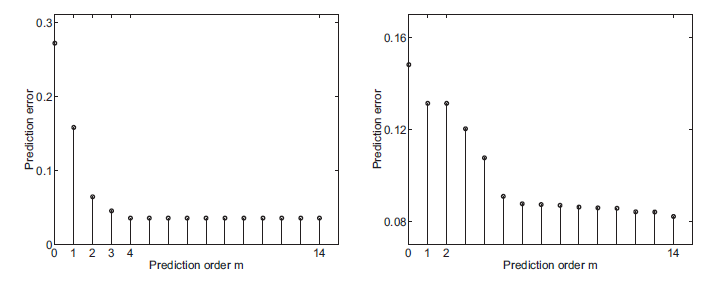


Fig. 3. AR process (*left*) and non-AR process (*right*).

***Note***: If a process is AR, then LPC (linear prediction coding) will reveal It.

**Part VI. MA and ARMA processes**

1. We know that a WSS random process x(n) is said to be **AR** if it satisfies a recursive (IIR) difference

equation of the form

where e(n) is a zero-mean white WSS process, and the polynomial *D*(z)=1+sum(d*i*\*z–1); *i* = 1,*N* has all zeros inside the unit circle.

1. We say that a WSS process x(n) is a moving average (**MA**) process if it satisfies a non-recursive (FIR) difference equation of the form

where e(n) is a zero-mean white WSS process.

1. Finally, we say that a WSS process x(n) is an **ARMA** process if

where e(n) is a zero-mean white WSS process. Defining the polynomials

We see that the above processes can be represented as in Fig. 4. In each of the three cases, x(n) is the output of a rational discrete time filter, driven by zero-mean white noise. For the AR process, the filter is an all-pole filter. For the MA process, the filter is FIR. For the ARMA process, the filter is IIR with both poles and zeros.

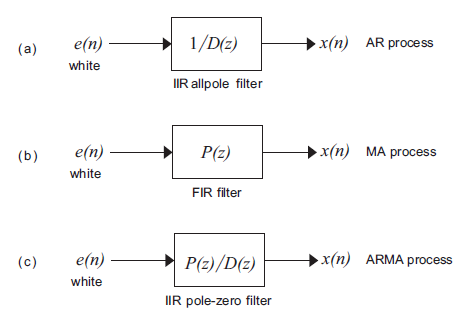


Fig. 4. (a) AR, (b) MA, and (c) ARMA processes.