Linear Prediction Filters
(Chapter 11)

The term "model" is used to describe the hidden laws that are supposed to govern a physical data. The representation of a stochastic process by a model dates back to an idea by Yule (1927). The idea is that a time-series *x(n)* consisting of highly correlated observations may be generated by applying a series of statistically in-depth "shocks" (impulses) to a linear filter. The "shocks" or impulses are random variables drawn from a fixed distribution that is normally Gaussian with zero mean and constant variance. Such a series is a pure random process and is called "White Gaussian Noise", and it satisfies:

$$E\{w(n)\} = 0 \text{ for all } n$$

$$E\{w(n)w^*(k)\} = \begin{cases} s_w^2 & if\ k = n \\ o & else \end{cases},$$

where s_w^2 is the variance of the process.

In general, the time-domain of the linear stochastic model may be described as the following:

Present values of the model output	+	linear combination of past values of the model output	=	linear combination of present and past values of input
$x(n)$	+	$\sum_{k=1}^{N} a_k x(n-k)$	=	$\sum_{k=o}^{M} b_k w(n-k)$

There are three popular types of Stochastic Models:

1) *Autoregressive (AR) Model*, in which no past values of input are used. In other words, the output is defined based on the present value of input and a linear combination of past values of output. (Note that here, w(n) is the input and x(n) is the output of the system model.)

$$x(n) = w(n) - \sum_{k=1}^{P} a_k x(n-k) \quad \text{or} \quad \sum_{k=o}^{P} a_k x(n-k) = w(n), \quad a_o = 1$$

(Note that here, *w(n)* is the input and *x(n)* is the output of the system model.)

2) *Moving Average (MA) Model*, in which no past values of output are used. In other words, the output is defined based on present and past values of the input.

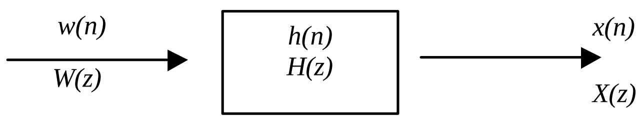
$$x(n) = \sum_{k=o}^{q} b_k w(n-k)$$

3) *Mixed ARMA Model*

$$\sum_{k=o}^{M} a_k x(n-k) = \sum_{k=o}^{N} b_k w(n-k), a_o = 1$$

Notes:

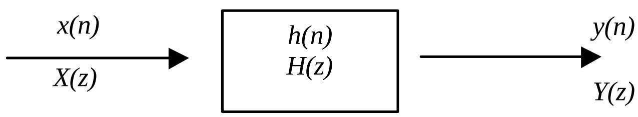
1. An AR Model is equivalent to an LTI System:



$$H_{AR}(z) = \frac{X(z)}{W(z)} = \frac{1}{1 + \sum_{k=1}^{p} a_k z^{-k}}$$

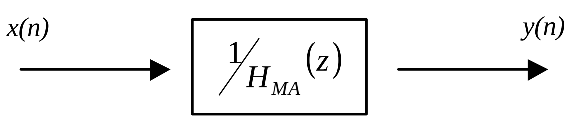
AR-process generator is an all-pole filter.

2. MA model is equivalent to LTI system:



$$H_{MA}(z) = \frac{Y(z)}{X(z)} = \sum_{k=1}^{q} b_k z^{-k}$$

This is an all-zero system. A solution of particular interest for the coefficients of this filter is a Min-Phase Filter. Then, equivalently we can represent the system by its inverse:



where $\dfrac{1}{H_{MA}(z)}$ is an all-pole filter.

This filter is called "whitening" filter and *y(n)* is called the "innovation process".

<u>Wold Decompostion</u>

In 1938, Wold proved a fundamental theorem, which states that any stationary discrete-time stochastic process may be decomposed into the sum of a general linear process and predictable process, with these two being uncorrelated with each other. More precisely, Wold proved the following:

Theorem: Any stationary discrete-time stochastic process *x(n)* may be expressed as $x(n) = u(n) + s(n)$ where

1) *s(n)* and *u(n)* are uncorrelated;

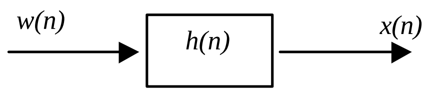
2) *u(n)* is a general linear process presented by the MA model:

$u(n) = \sum_{k=o}^{\infty} b_k^* w(n-k)$ (1), with $b_o = 1$ and $\sum_{k=o}^{\infty} |b_k|^2 < \infty$ and *w(n)* is a white-noise uncorrelated with *s(n)*; that is $E\{w(n) s^*(k)\} = 0$ or all *n* and *k*.

3) *s(n)* is a predictable process; that is, the process can be predicted from its own past values with zero prediction variance error.

The above is known as “Wold Decompostion Theorem”. According to Equation (1), *u(n)* may be generated by feeding an all-zeros filter with the white noise *w(n).* The zeros of this filter transfer function are the roots at the equation: $B(z)=\sum_{k=o}^{\infty} b_k^* z^{-k} = o$. If we solve it such that it becomes min-phase, then we can also produce *u(n)* by an all-pole filter and therefore by an AR model. The basic difference between MA and AR models is that *B(z)* operates on the input *w(n)* in the MA model, whereas the $B^{-1}(z)$ operates on the output *u(n)* of the AR model.

Now lets get back to the ARMA model and try to find the model parameters.



ARMA Model: $x(n)+\sum_{k=1}^{p} a_k x(n-k)=\sum_{k=0}^{q} b_k w(n-k)$. Multiply both sides by *x*(n-m)* and take the expected value. Then:

$$E\{x(n)x^*(n-m)\}=-\sum_{k=1}^{p} a_k E\{x(n-k)x^*(n-m)\}+\sum_{k=0}^{q} b_k E\{w(n-k)x^*(n-m)\}$$

$$g_{xx}(m)=-\sum_{k=1}^{p} a_k g_{xx}(m-k)+\sum_{k=o}^{q} b_k g_{wx}(m-k), \text{ but}$$

$$\begin{aligned} g_{wx}(m) &= E\{x^*(n)w(n+m)\} \\ &= E\left\{\sum_{k=0}^{\infty} h(k)w^*(n-k)\cdot w(n+m)\right\} \\ &= \sum_{k=0}^{\infty} h(k)E\{w^*(n-k)w(n+m)\} \end{aligned}$$

but *w(n)* is a white noise $\rightarrow$ $E\{w^*(n-k)w(n+m)\}=\begin{cases} s_w^2 & k=-m \\ 0 & else \end{cases}$ $\rightarrow$ $g_{wx}(m)=h(-m)s_w^2$ but the system is casual, meaning *h(negative values) = 0* $\Rightarrow g_{wx}(m)=\begin{cases} 0 & m>0 \\ s_w^2 h(-m) & m\le 0 \end{cases}$

$$\therefore \underbrace{g_{xx}(m)}_{ARMA} = \begin{cases} -\sum_{k=1}^{p} a_k g_{xx}(m-k) & m > q \\ -\sum_{k=1}^{p} a_k g_{xx}(m-k) + s_w^2 \sum_{k=o}^{q-m} h(k) b_{k+m} & 0 \le m \le q, \\ g_{xx}^*(-m) & m < 0 \end{cases} \qquad \text{Non linear equations!}$$

In the case of an AR model, it simplifies to:

$$\underbrace{g_{xx}(m)}_{AR} = \begin{cases} -\sum_{k=1}^{p} a_k g_{xx}(m-k) & m > 0 \\ -\sum_{k=1}^{p} a_k g_{xx}(m-k) + s_w^2 & m = 0 \\ g_{xx}^*(-m) & m < 0 \end{cases}$$

This can be written in Matrix form as:

$$\underbrace{\begin{bmatrix} g_{xx}(o) & g_{xx}(-1) & & g_{xx}(-P) \\ g_{xx}(1) & g_{xx}(0) & g_{xx}(-1) & g_{xx}(-P+1) \\ & & & \\ g_{xx}(P) & g_{xx}(P-1) & & g_{xx}\quad(0) \end{bmatrix}}_{(p+1)\times(p+1)\text{ Teoplitz matrix}} \begin{bmatrix} 1 \\ a_1 \\ \\ a_p \end{bmatrix} = \begin{bmatrix} s_w^2 \\ o \\ \\ o \end{bmatrix}$$

or in short: $\underline{\Gamma_{xx}}(n) \cdot \underline{a} = \begin{bmatrix} s_w^2 \\ \underline{0} \end{bmatrix}$ $g_{xx}(-k) = g_{xx}^*(k)$. These are called ***"Yull-Walker"*** Equations. Note that $\underline{\Gamma_{xx}}(w) = s_w^2 |H(w)|^2$.