Linear Stochastic Models

Special Types of Random Processes: AR, MA, and ARMA

Digital Signal Processing

Department of Electrical and Electronic Engineering, Imperial College

d.mandic@imperial.ac.uk

Motivation:- Wold Decomposition Theorem

The most fundamental justification for time series analysis is due to Wold's decomposition theorem, where it is explicitly proved that any (stationary) time series can be decomposed into two different parts.

Therefore, a general random process can be written a sum of two processes

$$x[n] = x_p[n] + x_r[n]$$

$\Rightarrow x_r[n]$ – regular random process
$\Rightarrow x_p[n]$ – predictable process, with $x_r[n] \quad \perp x_p[n]$,

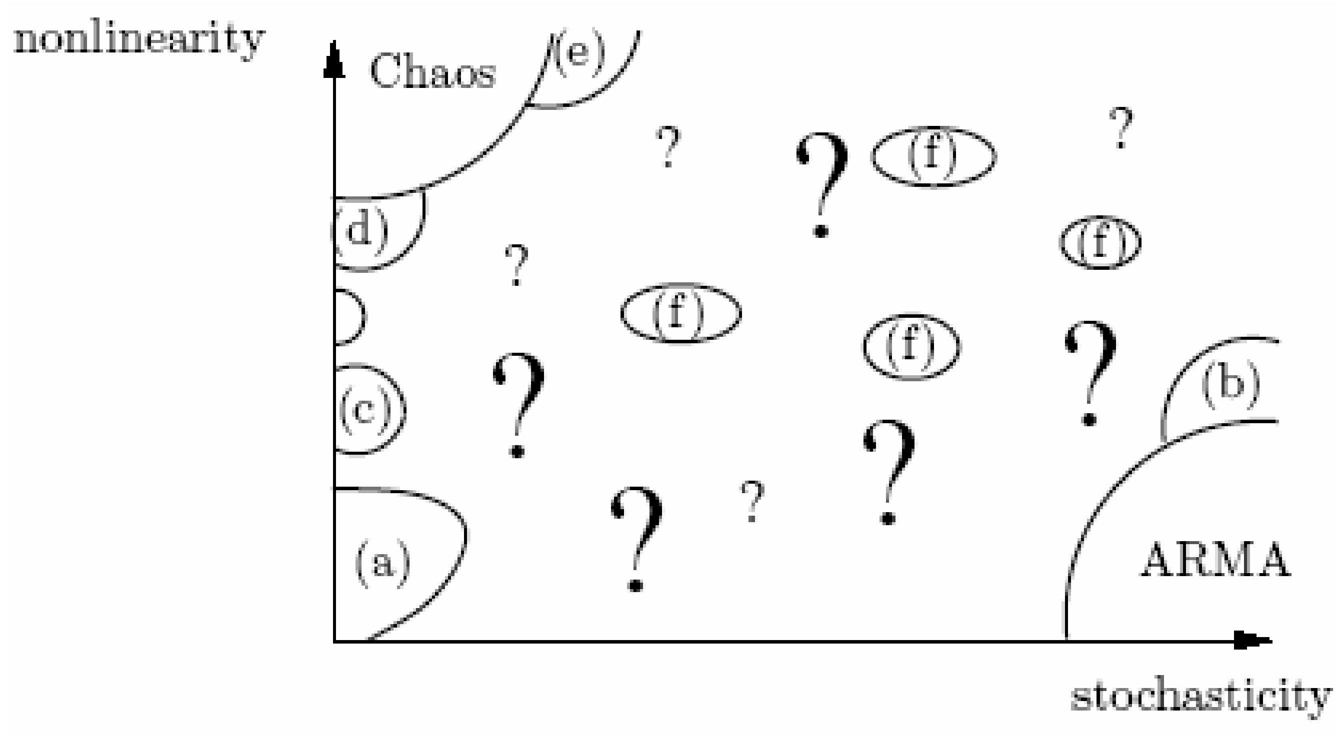
$$E\{x_r[m]x_p[n]\} = 0$$

that is we can **separately** treat the **predictable** process (i.e. a deterministic signal) and a **random** signal.

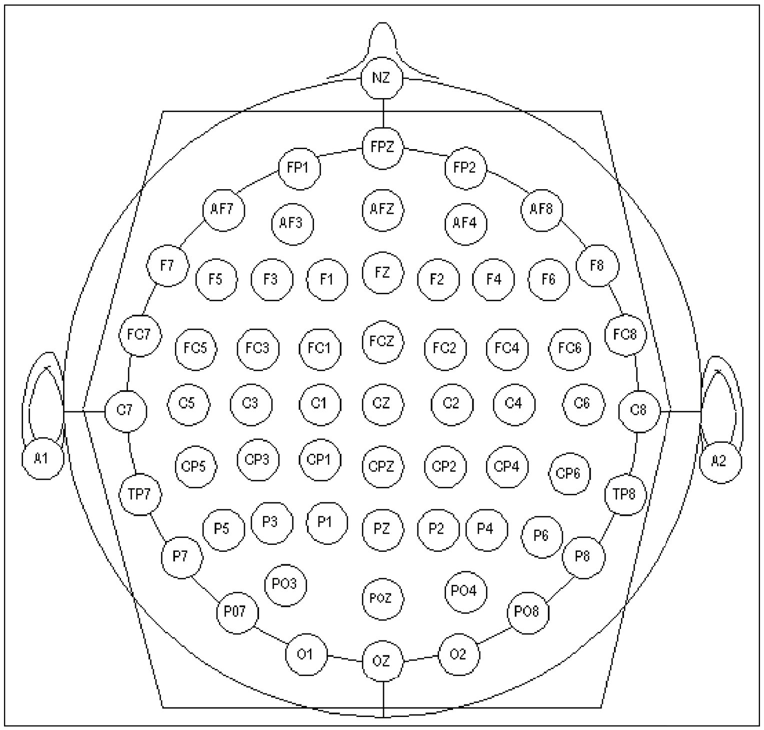
What do we actually mean?

a) Periodic oscillations b) Small nonlinearity c) Route to chaos

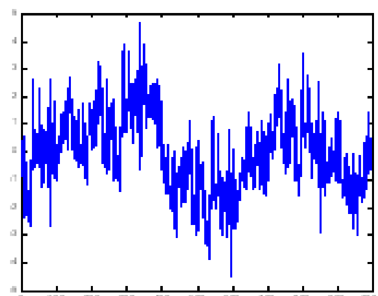
d) Route to chaos e) small noise f) HMM and others



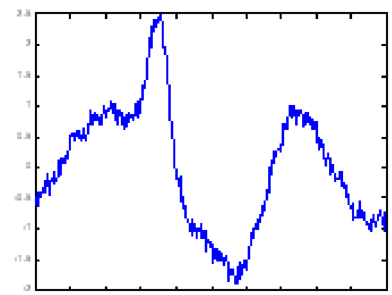
Example from brain science



Electrode positions



Raw EEG



Useful signal

Linear Stochastic Processes

It therefore follows that the general form for the power spectrum of a WSS process is

$$P_x(e^{\jmath\omega}) = P_{x_r}(e^{\jmath\omega}) + \sum_{k=1}^{N} \alpha_k u_0(\omega - \omega_k)$$

We look at processes generated by filtering white noise with a linear shift–invariant filter that has a rational system function. These include the

- Autoregressive (AR) $\rightarrow$ all pole system
- Moving Average (MA) $\rightarrow$ all zero system
- Autoregressive Moving Average (ARMA) $\rightarrow$ poles and zeros

Notice the difference between shift–invariance and time–invariance

ACF and Spectrum of ARMA models

Much of interest are the autocorrelation function and power spectrum of these processes. (**Recall that $ACF \quad \equiv PSD$ in terms of the available information**)

Suppose that we filter white noise $w[n]$ with a causal linear shift–invariant filter having a rational system function with p poles and q zeros

$$H(z) = \frac{B_q(z)}{A_p(z)} = \frac{\sum_{k=0}^{q} b_q(k) z^{-k}}{1 + \sum_{k=1}^{p} a_p(k) z^{-k}}$$

Assuming that the filter is stable, the output process $x[n]$ will be wide–sense stationary and with $P_w = \sigma_w^2$, the power spectrum of $x[n]$ will be

$$P_x(z) = \sigma_w^2 \frac{B_q(z) B_q(z^{-1})}{A_p(z) A_p(z^{-1})}$$

Recall that "$(\cdot)^*$" in analogue frequency corresponds to "z^{-1}" in "digital freq."

Frequency Domain

In terms of "digital" frequency θ (unit circle – $e^{-\jmath\theta} = e^{-\jmath\omega T}$)

- $B_q(z)B_q(z^{-1}) \looparrowright$ "quadratic form" and **real valued**
- $A_p(z)A_p(z^{-1}) \looparrowright$ "quadratic form" and **real valued**

$$P_z(e^{\jmath\theta}) = \sigma_w^2 \frac{\left|B_q(e^{\jmath\theta})\right|^2}{\left|A_p(e^{\jmath\theta})\right|^2}$$

We are therefore using $H(z)$ to shape the spectrum of white noise.

A process having a power spectrum of this form is known as an **autoregressive moving average** process of order (p, q) and is referred to as an

ARMA(p,q) process

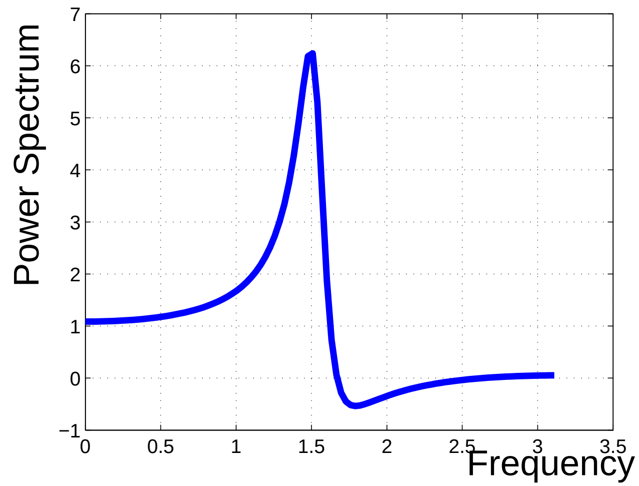
Example

Plot the power spectrum of an ARMA(2,2) process for which

- the zeros of $H(z)$ are $z = 0.95e^{\pm \jmath\pi/2}$
- poles are at $z = 0.9e^{\pm \jmath 2\pi/5}$

Solution: The system function is (poles and zeros – resonance & sink)

$$H(z) = \frac{1 + 0.9025z^{-2}}{1 - 0.5562z^{-1} + 0.81z^{-2}}$$



Difference Equation Representation

Random processes $x[n]$ and $w[n]$ are related by the linear constant coefficient equation

$$x[n] - \sum_{l=1}^{p} a_p(l)x[n-l] = \sum_{l=0}^{q} b_q(l)w[n-l]$$

Notice that the autocorrelation function of $x[n]$ and crosscorrelation between $x[n]$ and $w[n]$ follow the same difference equation, i.e. if we multiply both sides of the above equation by $x[n-k]$ and take the expected value, we have

$$r_{xx}(k) - \sum_{l=1}^{p} a_p(l)r_{xx}(k-l) = \sum_{l=0}^{q} b_q(l)r_{xw}(k-l)$$

Since x is WSS, it follows that $x[n]$ and $w[n]$ are jointly WSS.

General Linear Processes: Stationarity and Invertibility

Consider a linear stochastic process $\nrightarrow$ output from a linear filter, driven by WGN $w[n]$

$$x[n] = w[n] + b_1 w[n-1] + b_2 w[n-2] + \cdots = w[n] + \sum_{j=1}^{\infty} b_j w[n-j]$$

that is, **a weighted sum** of past inputs $w[n]$.

For this process to be a valid stationary process, the coefficients must be absolutely summable, that is $\sum_{j=0}^{\infty} |b_j| < \infty$.

The model implies that under suitable condition, $x[n]$ is also a weighted sum of past values of x, plus an added shock $w[n]$, that is

$$x[n] = a_1 x[n-1] + a_2 x[n-2] + \cdots + w[n]$$

- Linear Process is *stationary* if $\sum_{j=0}^{\infty} |b_j| < \infty$
- Linear Process is *invertible* if $\sum_{j=0}^{\infty} |a_j| < \infty$

Are these ARMA(p,q) processes?

- **Unit response** $u[n] = \begin{cases} 0, & n < 0 \\ 1, & n \geq 0 \end{cases}$
  - If $w[n] = \delta[n]$ then

$$u[n] = u[n-1] + w[n], \qquad n \geq 0$$

- **Ramp function** $r[n] = \begin{cases} 0, & n < 0 \\ n, & n \geq 0 \end{cases}$
  - If $w[n] = u[n]$ then

$$r[n] = r[n-1] + w[n], \qquad n \geq 0$$

Autoregressive Processes

A general $AR(p)$ process (autoregressive of order p) is given by

$$x[n] = a_1 x[n-1] + \cdots + a_p x[n-p] + w[n] = \sum_{i=1}^{p} a_i x[n-i] + w[n]$$

Observe the auto–regression above

Duality between AR and MA processes:

For instance the first order autoregressive process

$$x[n] = a_1 x[n-1] + w[n] \quad \Leftrightarrow \quad \sum_{j=0}^{\infty} b_j w[n-j]$$

Due to its “all–pole“ nature follows the duality between IIR and FIR filters.

ACF and Spectrum of AR Processes

To obtain the autocorrelation function of an AR process, multiply the above equation by $x[n-k]$ to obtain

$$x[n-k]x[n] = a_1 x[n-k]x[n-1] + a_2 x[n-k]x[n-2] + \cdots$$
$$+ a_p x[n-k]x[n-p] + x[n-k]w[n]$$

Notice that $E\{x[n-k]w[n]\}$ vanishes when $k > 0$. Therefore we have

$$r_{xx}(k) = a_1 r_{xx}(k-1) + a_2 r_{xx}(k-2) + \cdots + a_p r_{xx}(k-p) \quad k > 0$$

On dividing throughout by $r_{xx}(0)$ we obtain

$$\rho(k) = a_1 \rho(k-1) + a_2 \rho(k-2) + \cdots + a_p \rho(k-p) \quad k > 0$$

Parameters $\rho(k)$ are correlation coefficients

Variance and Spectrum of AR Processes

Variance:

When $k = 0$ the contribution from the term $E\{x[n-k]w[n]\}$ is σ_w^2, and

$$r_{xx}(0) = a_1 r_{xx}(-1) + a_2 r_{xx}(-2) + \cdots + a_p r_{xx}(-p) + \sigma_w^2$$

Divide by $r_{xx}(0) = \sigma_x^2$ to obtain

$$\sigma_x^2 = \frac{\sigma_w^2}{1 - \rho_1 a_1 - \rho_2 a_2 - \cdots - \rho_p a_p}$$

Spectrum:

$$P_{xx}(f) = \frac{2\sigma_w^2}{\left|1 - a_1 e^{-\jmath 2\pi f} - \cdots - a_p e^{-\jmath 2\pi p f}\right|^2} \qquad 0 \le f \le 1/2$$

Recall *Spectrum of linear systems* from the Course Introduction

Yule–Walker Equations

For $k = 1, 2, \ldots, p$ from the general autocorrelation function, we obtain a set of equations:-

$$\begin{aligned} r_{xx}(1) &= a_1 r_{xx}(0) + a_2 r_{xx}(1) + \cdots + a_p r_{xx}(p-1) \\ r_{xx}(2) &= a_1 r_{xx}(1) + a_2 r_{xx}(0) + \cdots + a_p r_{xx}(p-2) \\ \vdots &= \vdots \\ r_{xx}(p) &= a_1 r_{xx}(p-1) + a_2 r_{xx}(p-2) + \cdots + a_p r_{xx}(0) \end{aligned}$$

These equations are called the Yule–Walker or normal equations.

Their solution gives us the set of autoregressive parameters $\mathbf{a} = [a_1, \ldots, a_p]^T$. This can be expressed in a vector–matrix form as

$$\mathbf{a} = \mathbf{R}_{xx}^{-1} \mathbf{r}_{xx}$$

Due to Toeplitz structure of $\mathbf{R}_{xx}$, its positive definitness enables matrix inversion

ACF Coefficients

For the autocorrelation coefficients

$$\rho_k = r_{xx}(k)/r_{xx}(0)$$

we have

$$\begin{aligned} \rho_1 &= a_1 + a_2\rho_1 + \cdots + a_p\rho_{p-1} \\ \rho_2 &= a_1\rho_1 + a_2 + \cdots + a_p\rho_{p-2} \\ \vdots &= \vdots \\ \rho_p &= a_1\rho_{p-1} + a_2\rho_{p-2} + \cdots + a_p \end{aligned}$$

When does the sequence $\{\rho_0, \rho_1, \rho_2, \ldots\}$ vanish?

Homework:- Try command **xcorr** in Matlab

Example:- Yule–Walker modelling in Matlab

In Matlab – Power spectral density using Y–W method *pyulear*

```
Pxx = pyulear(x,p)
[Pxx,w] = pyulear(x,p,nfft)
[Pxx,f] = pyulear(x,p,nfft,fs)
[Pxx,f] = pyulear(x,p,nfft,fs,'range')
[Pxx,w] = pyulear(x,p,nfft,'range')
```

Description:-

```
Pxx = pyulear(x,p)
```

implements the Yule-Walker algorithm, and returns Pxx, an estimate of the power spectral density (PSD) of the vector x.

To remember for later → This estimate is also an estimate of the maximum entropy.

Se also **aryule, lpc, pburg, pcov, peig, periodogram**

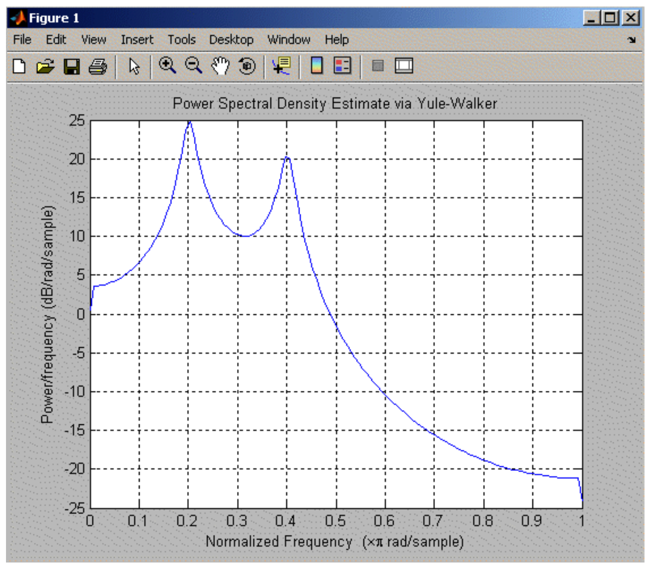
Example:- $AR(p)$ signal generation

- Generate the input signal **x** by filtering white noise through the AR filter

- Estimate the PSD of **x** based on a fourth-order AR model

Solution:-

```
randn('state',1);
x = filter(1,a,randn(256,1));     % AR system output
pyulear(x,4)                      % Fourth-order estimate
```

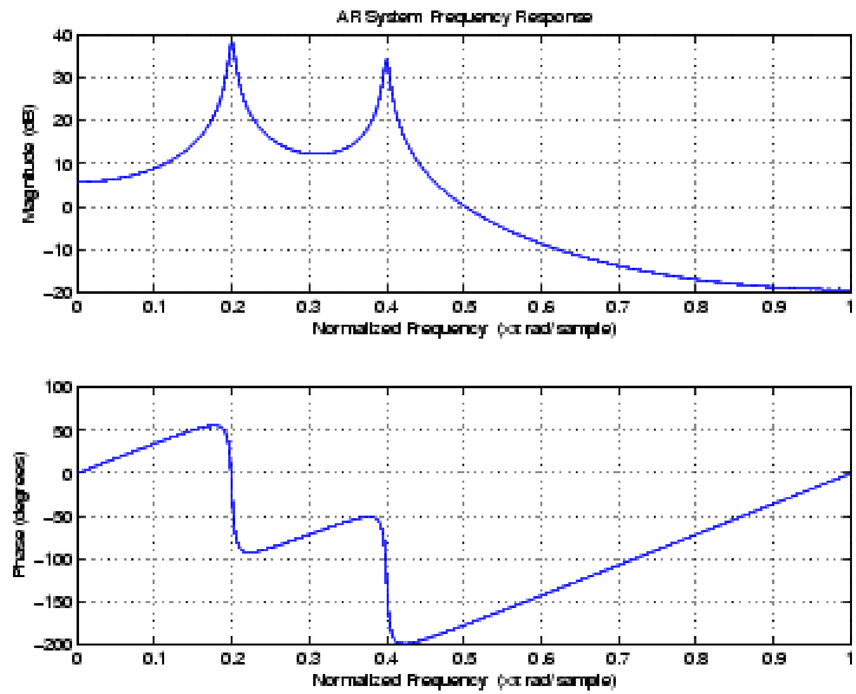


Alternatively:- Yule–Walker modelling

AR(4) system given by

$$y[n] = 2.2137y[n-1] - 2.9403y[n-2] + 2.1697y[n-3] - 0.9606y[n-4] + w[n]$$

```
a = [1 -2.2137 2.9403 -2.1697 0.9606];  % AR filter coefficients
freqz(1,a)  % AR filter frequency response
title('AR System Frequency Response')
```



From Data to $AR(p)$ Model

So far, we assumed the model (AR, MA, or ARMA) and analysed the ACF and PSD based on known model coefficients.

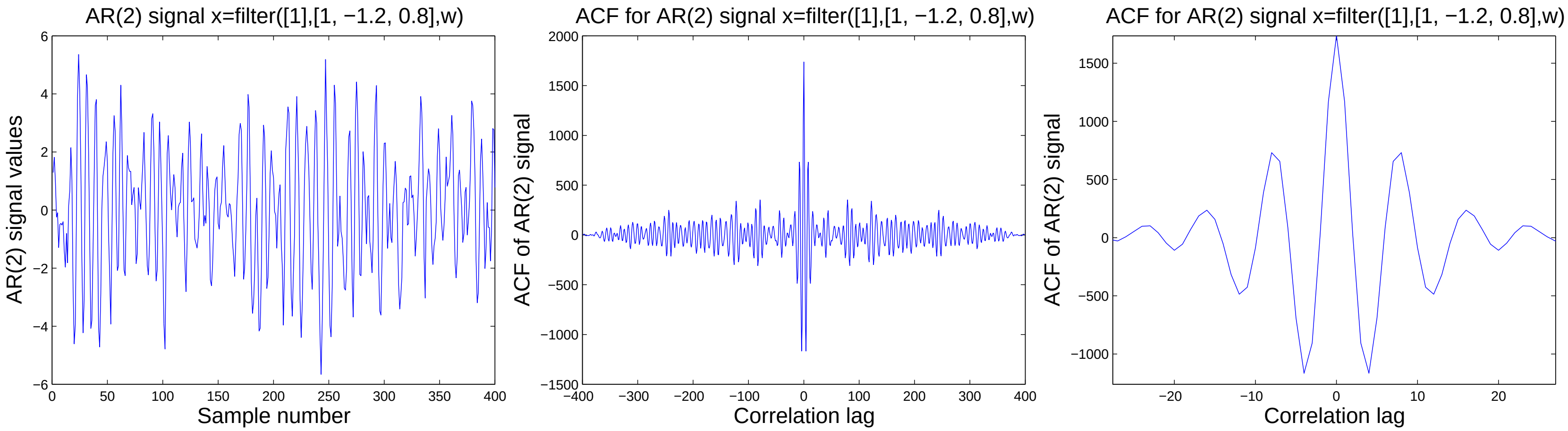
In practice:- **DATA** $\looparrowright$ **MODEL**

This procedure is as follows:-

```
* record data x(k)
* find the autocorrelation of the data ACF(x)
* divide by r_xx(0) to obtain correlation coefficients \rho(k)
* write down Yule-Walker equations
* solve for the vector of AR paramters
```

The problem is that we do not know the model order p beforehand; we will deal with this problem later in Lecture 2.

Example:- Finding parameters of $x[n] = 1.2x[n-1] - 0.8x[n-2] + w[n]$



Apply:- `for i=1:6; [a,e]=aryule(x,i); display(a);end`

$\mathbf{a}^{(1)} = [0.6689]$ $\quad\quad$ $\mathbf{a}^{(2)} = [1.2046, -0.8008]$

$\mathbf{a}^{(3)} = [1.1759, -0.7576, -0.0358]$

$\mathbf{a}^{(4)} = [1.1762, -0.7513, -0.0456, 0.0083]$

$\mathbf{a}^{(5)} = [1.1763, -0.7520, -0.0562, 0.0248, -0.0140]$

$\mathbf{a}^{(6)} = [1.1762, -0.7518, -0.0565, 0.0198, -0.0062, -0.0067]$

Special case:- AR(1) Process (Markov)

Given below (Recall $p(x[n], x[n-1], \ldots, x[0]) = p(x[n] \,|\, x[n-1])$)

$$x[n] = a_1 x[n-1] + w[n] = w[n] + a_1 x[n-1] + a_1^2 w[n-2] + \cdots$$

i) for the process to be **stationary** $-1 < a_1 < 1$.

ii) **Autocorrelation Function:-** from Yule-Walker equations

$$r_{xx}(k) = a_1 r_{xx}(k-1), \quad k > 0$$

or for the correlation coefficients, with $\rho_0 = 1$

$$\rho_k = a_1^k, \quad k > 0$$

Notice the difference in the behaviour of the ACF for a_1 positive and negative

Variance and Spectrum of AR(1) process

Can be calculated directly from a general expression of the variance and spectrum of $AR(p)$ processes.

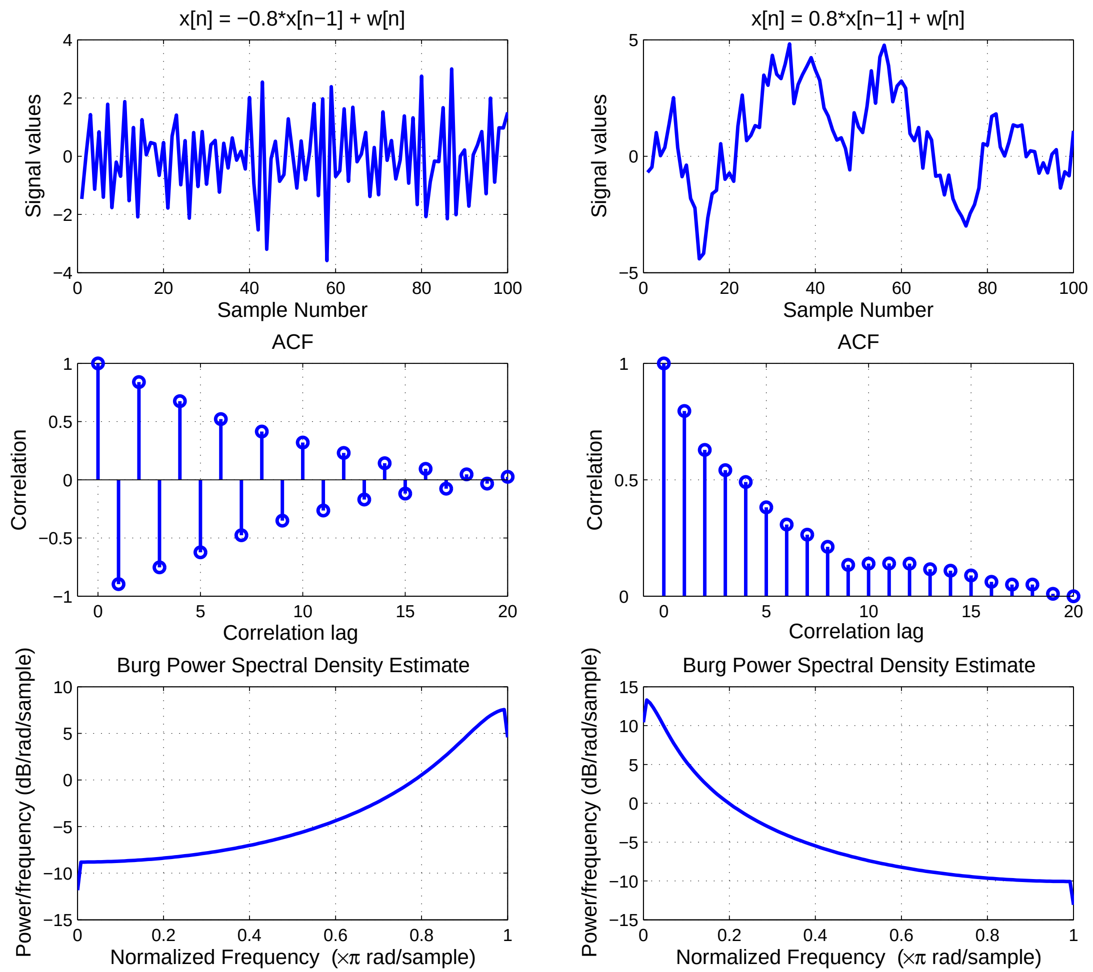
- **Variance:**- Also from a general expression for the variance of linear processes from Lecture 1

$$\sigma_x^2 = \frac{\sigma_w^2}{1 - \rho_1 a_1} = \frac{\sigma_w^2}{1 - a_1^2}$$

- **Spectrum:**- Notice how the flat PSD of WGN is shaped according to the position of the pole of $AR(1)$ model (LP or HP)

$$P_{xx}(f) = \frac{2\sigma_w^2}{|1 - a_1 e^{-\jmath 2\pi f}|^2} = \frac{2\sigma_w^2}{1 + a_1^2 - 2a_1 cos(2\pi f)}$$

Example: ACF and Spectrum of AR(1) for $a = \pm 0.8$



$a < 0 \rightarrow$ **High Pass** $a > 0 \rightarrow$ **Low Pass**

Special Case:- Second Order Autoregressive Processes AR(2)

The input–output functional relationship is given by

$$x[n] = a_1 x[n-1] + a_2 x[n-2] + w[n]$$

For **stationarity**- (to be proven later)

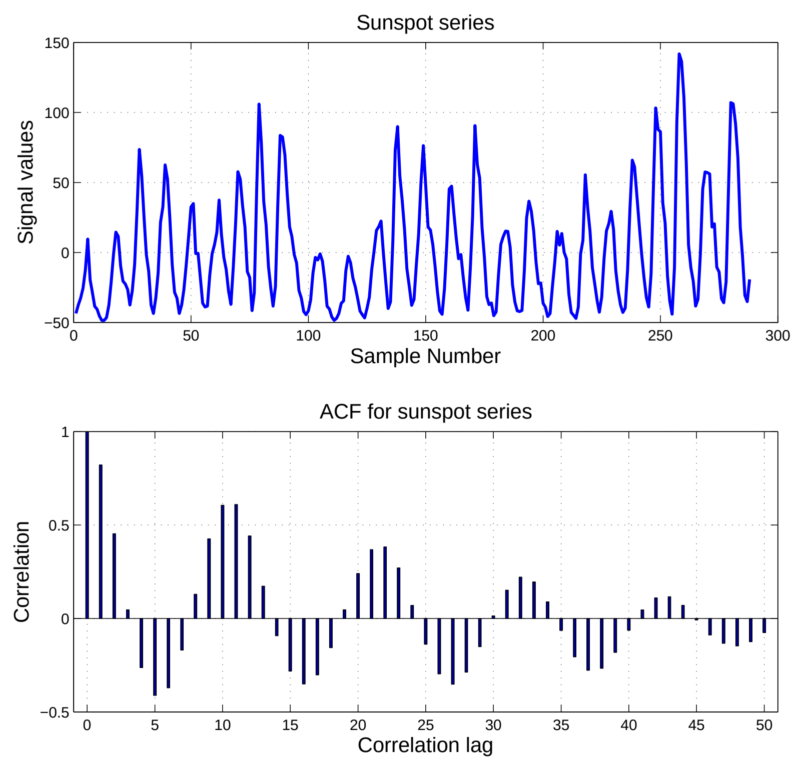
$$\begin{aligned} a_1 + a_2 &< 1 \\ a_2 - a_1 &< 1 \\ -1 < a_2 &< 1 \end{aligned}$$

This will be shown within the so–called "stability triangle"

Work by Yule – Modelling of sunspot numbers

Recorded for more than 300 years.

In 1927, Yule modelled them and invented $AR(2)$ model



Sunspot numbers and its autocorrelation function

Autocorrelation function of AR(2) processes

The ACF

$$\rho_k = a_1 \rho_{k-1} + a_2 \rho_{k-2} \qquad k > 0$$

- **Real roots:** $\Rightarrow$ $(a_1^2 + 4a_2 > 0)$ ACF $=$ mixture of damped exponentials
- **Complex roots:** $\Rightarrow$ $(a_1^2 + 4a_2 < 0)$ $\Rightarrow$ ACF exhibits a pseudo–periodic behaviour

$$\rho_k = \frac{D^k \sin(2\pi f_0 k + F)}{\sin F}$$

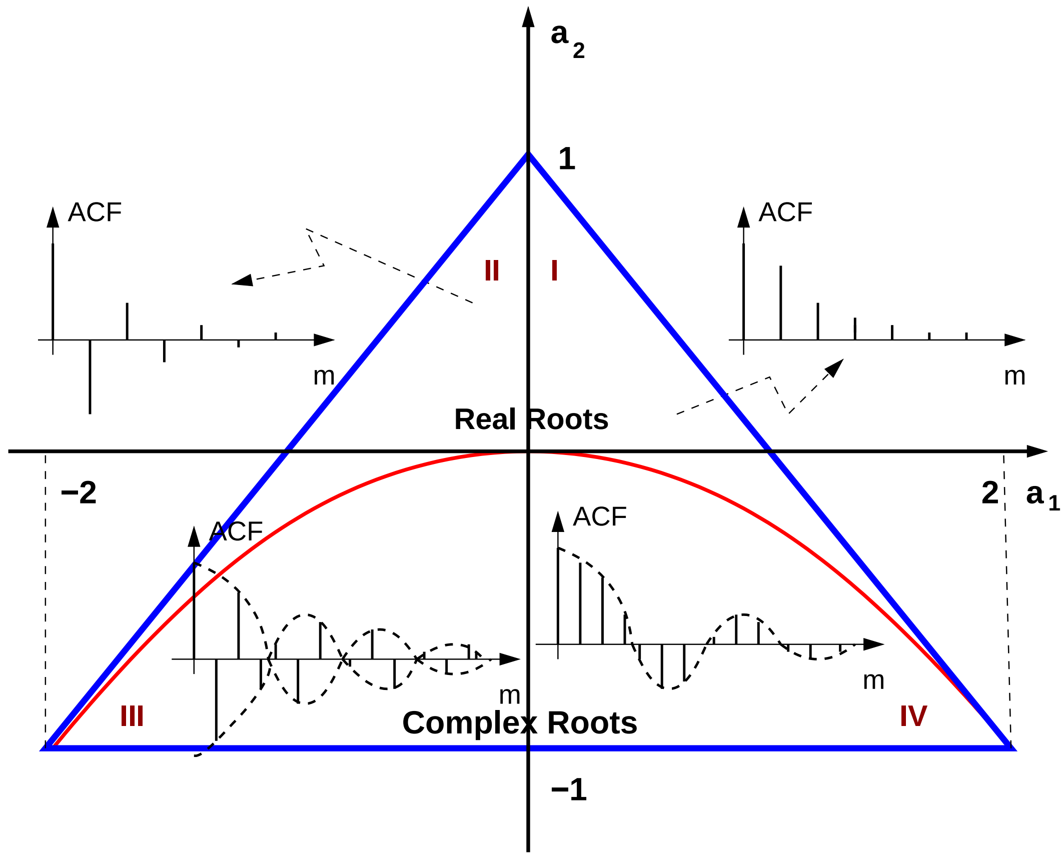
D - damping factor, of a sine wave with frequency f_0 and phase F.

$$D = \sqrt{-a_2}$$

$$\cos(2\pi f_0) = \frac{a_1}{2\sqrt{-a_2}}$$

$$\tan(F) = \frac{1 + D^2}{1 - D^2} \tan(2\pi f_0)$$

Stability Triangle



i) **Real roots** Region 1: Monotonically decaying ACF
ii) **Real roots** Region 2: Decaying oscillating ACF
iii) **Complex roots** Region 3: Oscilating pseudoperiodic ACF
iv) **Complex roots** Region 4: Pseudoperiodic ACF

Yule–Walker Equations

Substituting $p = 2$ into Y-W equations we have

$$\rho_1 = a_1 + a_2\rho_1$$

$$\rho_2 = a_1\rho_1 + a_2$$

which when solved for a_1 and a_2 gives

$$a_1 = \frac{\rho_1(1-\rho_2)}{1-\rho_1^2} \qquad a_2 = \frac{\rho_2-\rho_1^2}{1-\rho_1^2}$$

or substituting in the equation for ρ

$$\rho_1 = \frac{a_1}{1-a_2}$$

$$\rho_2 = a_2 + \frac{a_1^2}{1-a_2}$$

Variance and Spectrum

More specifically, for the $AR(2)$ process, we have:-

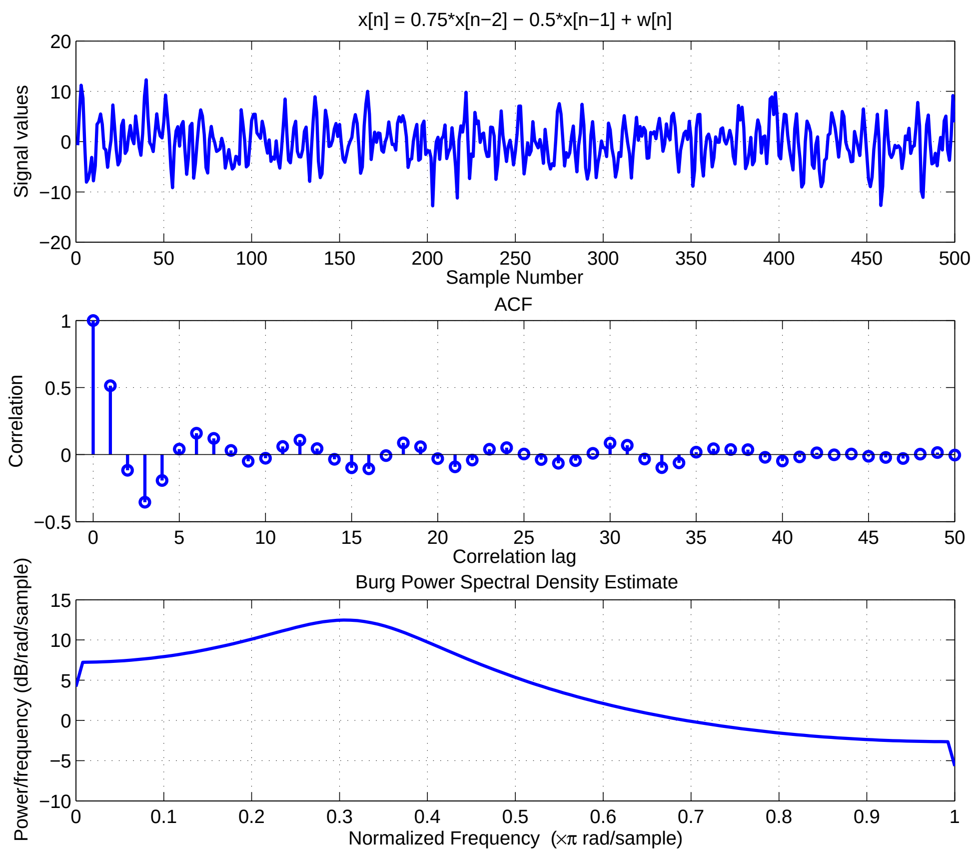
Variance

$$\sigma_x^2 = \frac{\sigma_w^2}{1 - \rho_1 a_1 - \rho_2 a_2} = \left(\frac{1 - a_2}{1 + a_2}\right) \frac{\sigma_w^2}{(1 - a_2)^2 - a_1^2}$$

Spectrum

$$\begin{aligned} P_{xx}(f) &= \frac{2\sigma_w^2}{\left|1 - a_1 e^{-\jmath 2\pi f} - a_2 e^{-\jmath 4\pi f}\right|^2} \\ &= \frac{2\sigma_w^2}{1 + a_1^2 + a_2^2 - 2a_1(1 - a_2\cos(2\pi f) - 2a_2\cos(4\pi f))}, \qquad 0 \le f \le 1/2 \end{aligned}$$

Example AR(2): $x[n] = 0.75x[n-1] - 0.5x[n-2] + w[n]$

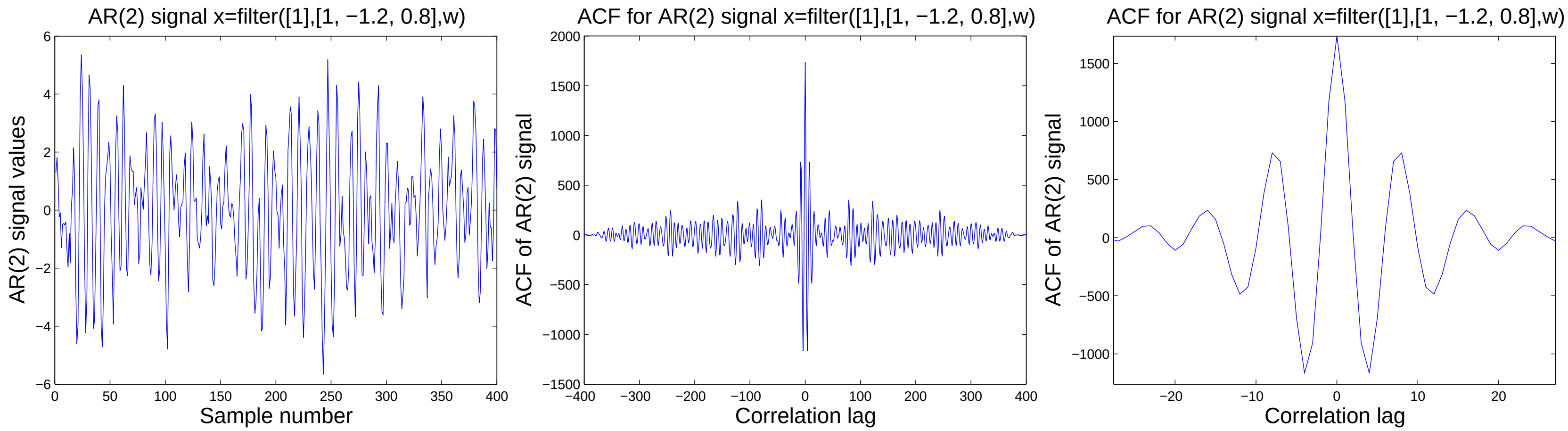


The damping factor $D = \sqrt{0.5} = 0.71$, frequency $f_0 = \frac{\cos^{-1}(0.5303)}{2\pi} = \frac{1}{6.2}$

The fundamental period of the autocorrelation function is 6.2.

Partial Autocorrelation Function:- Motivation

Let us revisit example from page 21 of Lecture Slides.



`We do not know p, let us re-write the coefficients as [a_1p,...,a_p`

$$\mathbf{p=1} \looparrowright [0.6689] = a_{11} \qquad \mathbf{p=2} \looparrowright [1.2046, -0.8008] = [a_{21}, a_{22}]$$

$$\mathbf{p=3} \looparrowright [1.1759, -0.07576, -0.0358] = [a_{31}, a_{32}, a_{33}]$$

$$\mathbf{p=4} \looparrowright [1.1762, -0.7513, -0.0456, 0.0083] = [a_{41}, a_{42}, a_{43}, a_{44}]$$

$$\mathbf{p=5} \looparrowright [1.1763, -0.7520, -0.0562, 0.0248, -0.0140] = [a_{51}, \ldots, a_{55}]$$

$$\mathbf{p=6} \looparrowright [1.1762, -0.7518, -0.0565, 0.0198, -0.0062, -0.0067] = [a_{61}, \ldots, a_{66}]$$

Partial Autocorrelation Function

Notice: ACF of $AR(p)$ infinite in duration, **but** can by be described in terms of p nonzero functions ACFs.

Denote by a_{kj} the jth coefficient in an autoregressive representation of order k, so that a_{kk} is the last coefficient. Then

$$\rho_j = a_{kj}\rho_{j-1} + \cdots + a_{k(k-1)}\rho_{j-k+1} + a_{kk}\rho_{j-k} \qquad j = 1, 2, \ldots, k$$

leading to the Yule–Walker equation, which can be written as

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & 1 \end{bmatrix} \begin{bmatrix} a_{k1} \\ a_{k2} \\ \vdots \\ a_{kk} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix}$$

Partial ACF Coefficients:

Solving these equations for $k = 1, 2, \ldots$ successively, we obtain

$$a_{11} = \rho_1, \qquad a_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}, \qquad a_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}, \qquad \text{etc}$$

- The quantity a_{kk}, regarded as a function of lag k, is called the **partial autocorrelation function**.

- For an AR(p) process, the PAC a_{kk} will be nonzero for $k \leq p$ and zero for $k > p \Rightarrow$ tells us the order of an $AR(p)$ process.

Importance of Partial ACF

For a zero mean process $x[n]$, the best **linear predictor** in the **mean square error** sense of $x[n]$ based on $x[n-1], x[n-2], \ldots$ is

$$\hat{x}[n] = a_{k-1,1}x[n-1] + a_{k-1,2}x[n-2] + \cdots + a_{k-1,k-1}x[n-k+1]$$

(apply the $E\{\cdot\}$ operator to the general $AR(p)$ model expression, and recall that $E\{w[n]\} = 0$)

(Hint:
$E\{x[n]\} = \hat{x}[n] = E\left\{a_{k-1,1}x[n-1] + \cdots + a_{k-1,k-1}x[n-k+1] + w[n]\right\} = a_{k-1,1}x[n-1] + \cdots + a_{k-1,k-1}x[n-k+1]$))

whether the process is an AR or not

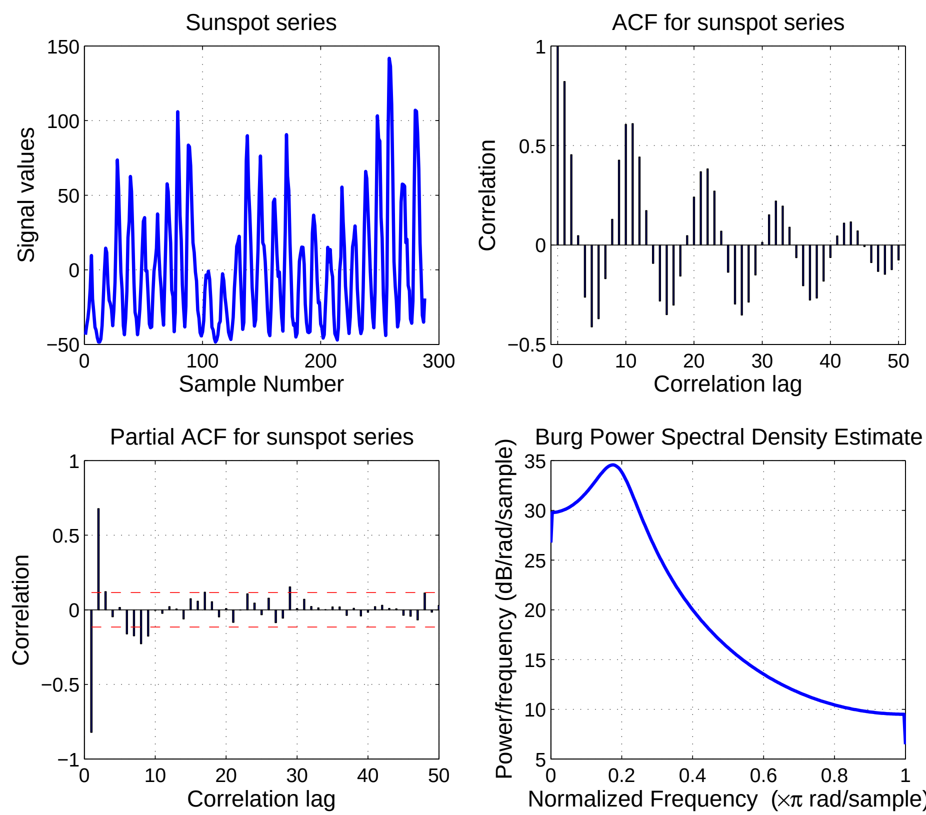
In MATLAB, check the function:

`ARYULE`

and functions

`PYULEAR, ARMCOV, ARBURG, ARCOV, LPC, PRONY`

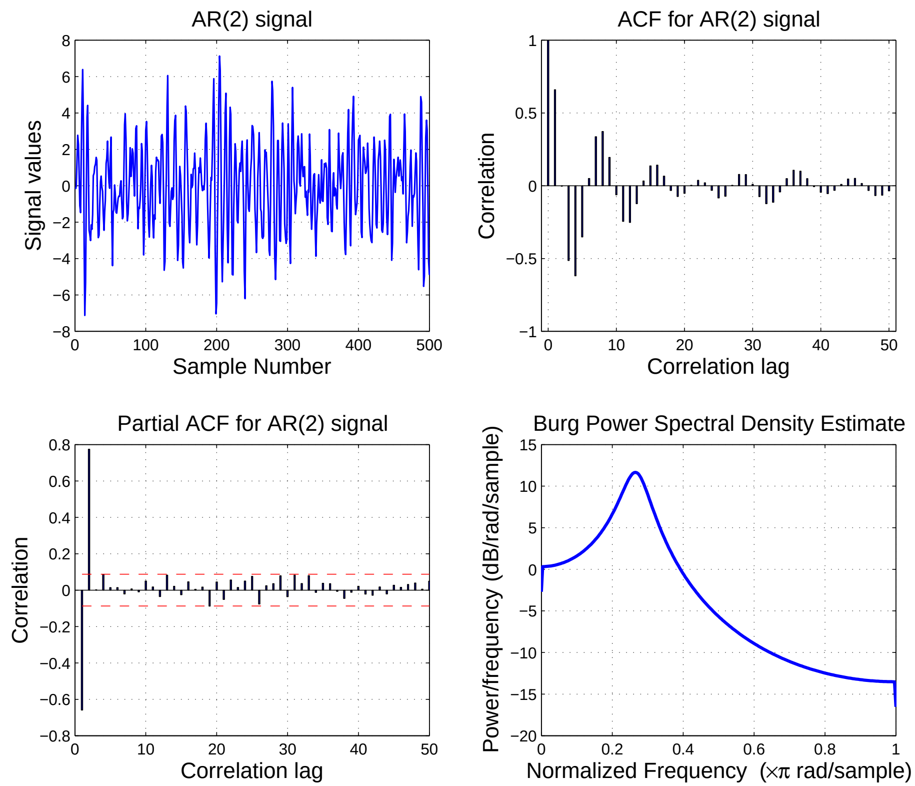
Model order for Sunspot numbers



Sunspot numbers, their ACF and partial autocorrelation (PAC)

After lag $k = 2$, the PAC becomes very small

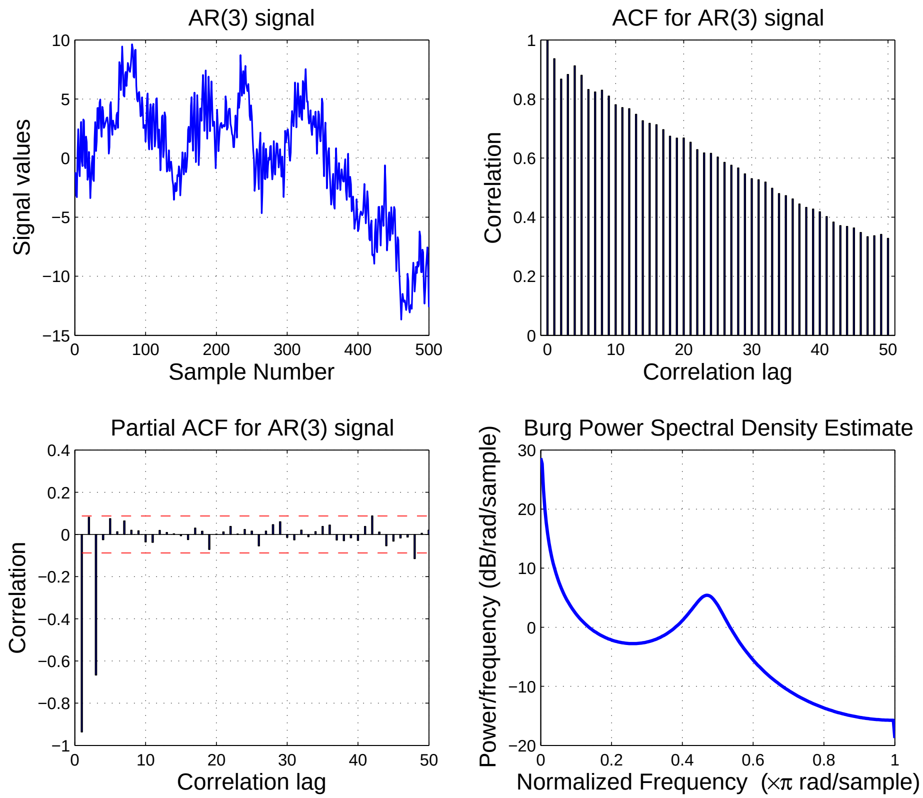
Model order for AR(2) generated process



AR(2) signal, its ACF and partial autocorrelation (PAC)

After lag $k = 2$, the PAC becomes very small

Model order for AR(3) generated process



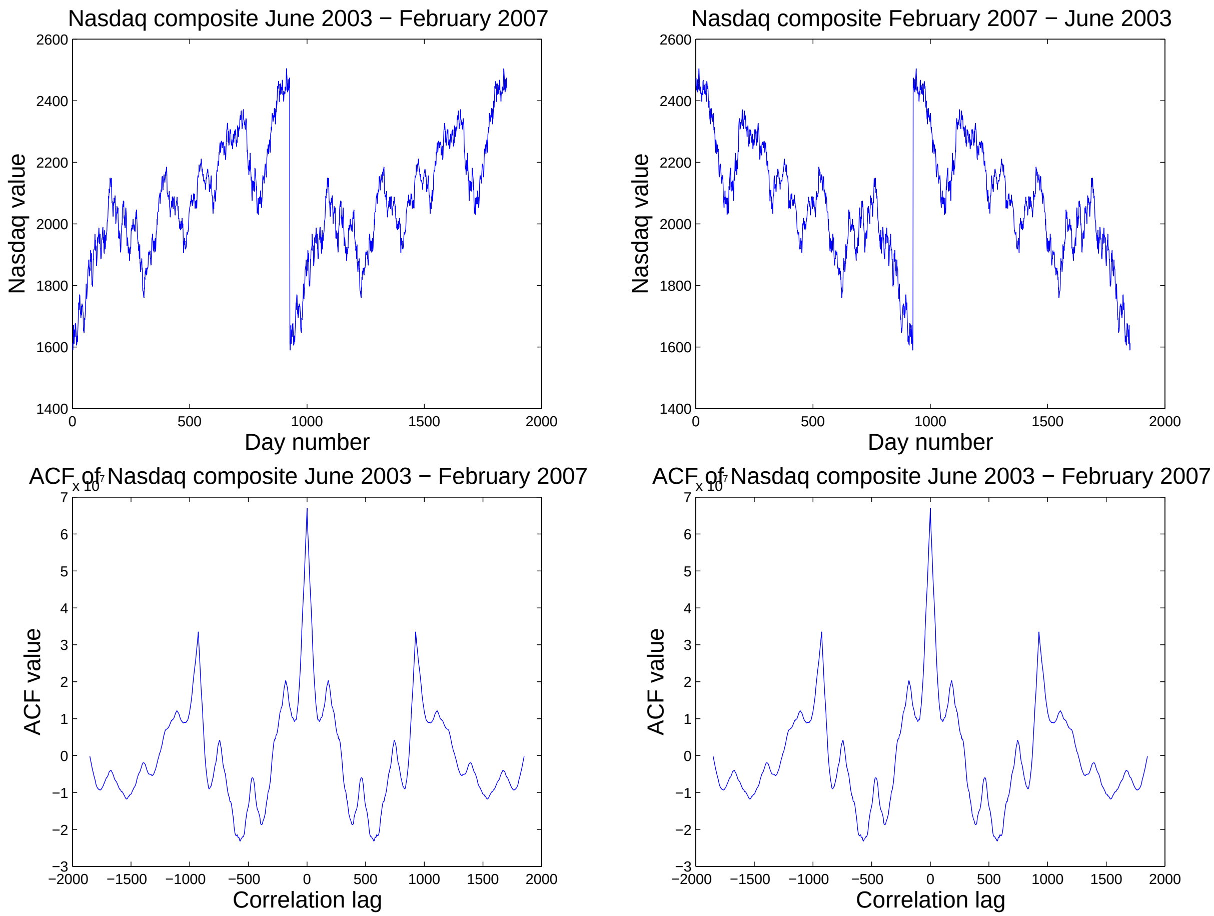
AR(3) signal, its ACF and partial autocorrelation (PAC)

After lag $k = 3$, the PAC becomes very small

Model order for a financial time series

From:- http://finance.yahoo.com/q/ta?s=%5EIXIC&t=1d&l=on&z=m&q=b&p=v&a=&c=

Nasdaq ascending | **Nasdaq descending**



Partial ACF for financial time series

```
a = 1.0000   -0.9994

a = 1.0000   -0.9982   -0.0011

a = 1.0000   -0.9982    0.0086   -0.0097

a = 1.0000   -0.9983    0.0086   -0.0128    0.0030

a = 1.0000   -0.9983    0.0086   -0.0128    0.0026    0.0005

a = 1.0000   -0.9983    0.0086   -0.0127    0.0026    0.0017   -0.0
```

Model Order Selection – Practical issues

In practice – the greater the model order the higher the accuracy

$\Rightarrow$ **When do we stop?**

To save on computational complexity, we introduce "penalty" for a high model order. The criteria for model order selection are, for instance MDL (minimum description length - Rissanen), AIC (Akaike Information criterion), given by

$$MDL = log(E) + \frac{p * log(N)}{N}$$

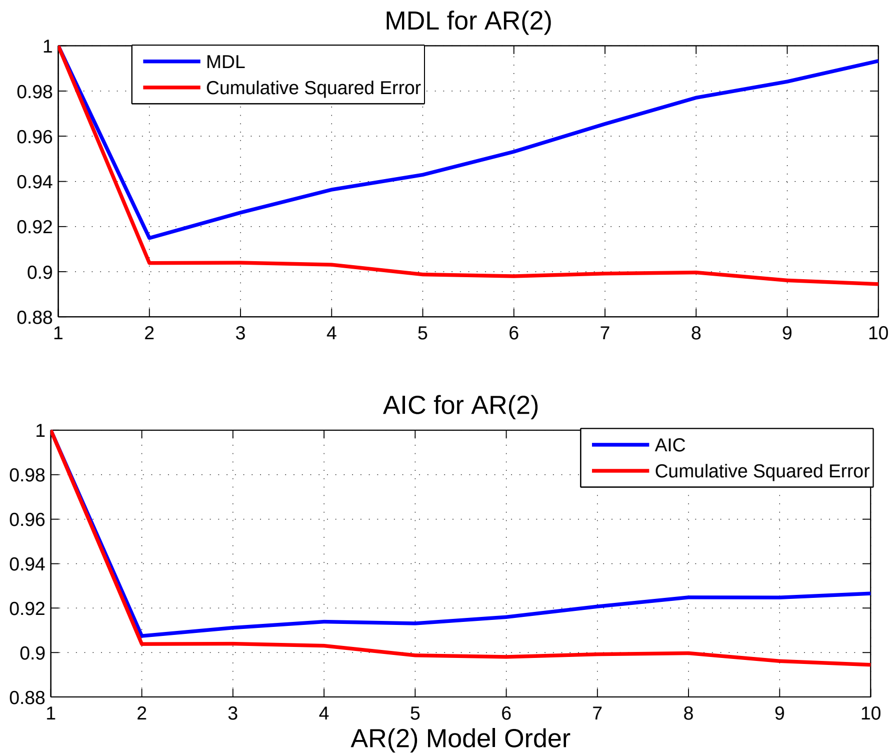
$$AIC = log(E) + 2p/N$$

E = the loss function (typically cumulative squared error,
p = the number of estimated parameters
N = the number of estimated data.

Example:- Model order selection – MDL vs AIC

Let us have a look at the squared error and the MDL and AIC criteria for an AR(2) model with

$$a_1 = 0.5 \qquad a_2 = -0.3$$



(Model error)^2 versus the model order p

Moving Average Processes

A general MA(q) process is given by

$$x[n] = w[n] + b_1 w[n-1] + \cdots + b_q w[n-q]$$

Autocorrelation function: The autocovariance function of MA(q)

$$c_k = E[(w[n] + b_1 w[n-1] + \cdots + b_q w[n-q])(w[n-k] \\ + b_1 w[n-k-1] + \cdots + b_q w[n-k-q])]$$

Hence the variance of the process

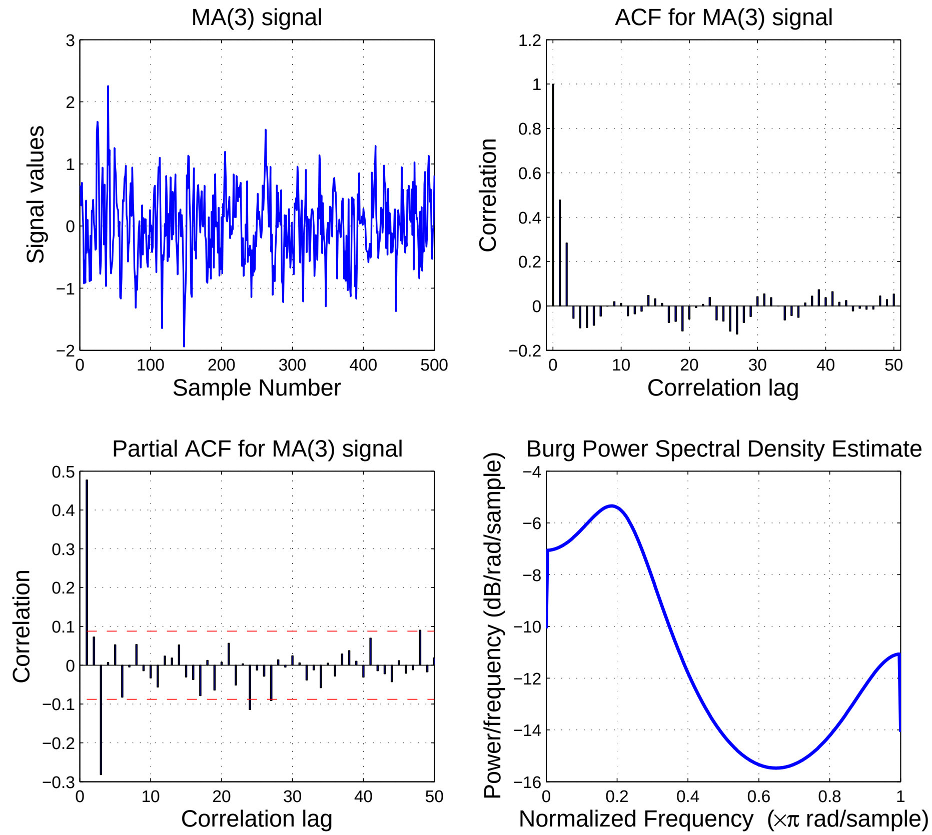
$$c_0 = (1 + b_1^2 + \cdots + b_q^2)\sigma_w^2$$

The ACF of an MA process has a cutoff after lag q.

Spectrum: All zeros $\Rightarrow$ struggles to model PSD with peaks

$$P(f) = 2\sigma_w^2 \left|1 + b_1 e^{-\jmath 2\pi f} + b_2 e^{-\jmath 4\pi f} + \cdots + b_q e^{-\jmath 2\pi q f}\right|$$

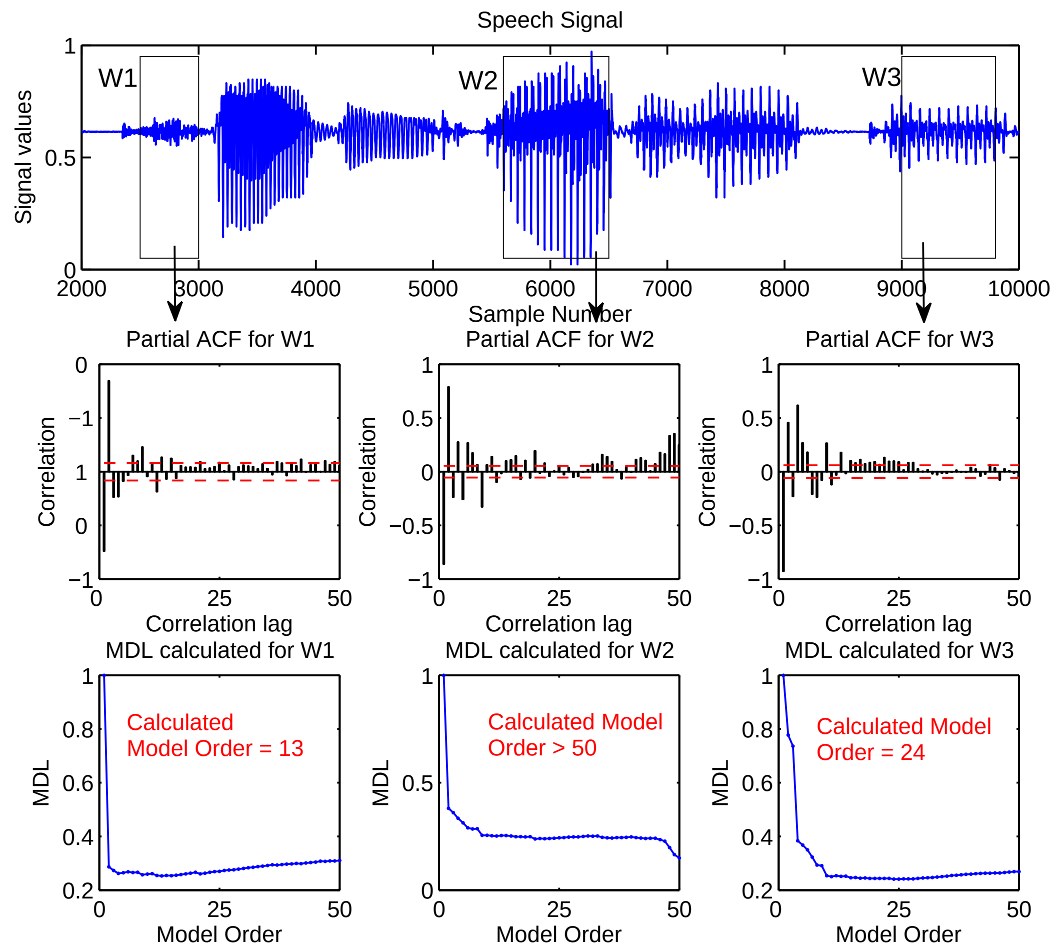
Example:- MA(3) process



MA(3) model, its ACF and partial autocorrelation (PAC)

After lag $k = 3$, the ACF becomes very small

Analysis of Nonstationary Signals



Different AR models for different segments of speech

To deal with nonstationarity we need short sliding windows

Duality Between AR and MA Processes

i) A stationary finite AR(p) process can be represented as an infinite order MA process. A finite MA process can be represented as an infinite AR process.

ii) The finite MA(q) process has an ACF that is zero beyond q. For an AR process, the ACF is infinite in extent and consits of mixture of damped exponentials and/or damped sine waves.

iii) Parameters of finite MA process are not required to satisfy any condition for stationarity. However, for invertibility, the roots of the characteristic equation must lie inside the unit circle.

ARMA modelling is a classic technique which has found a tremendous number of applications