Fundamental components of a time series

A time series may contain deterministic component(s) and stochastic component.

Deterministic components are non-random in nature and are of following types:

m_t : trend or long term movement / tendency characterizing a time series

s_t : seasonal components are distinguishable patterns of regular annual variations in a time series.

c_t : cyclical components are more or less regular long range swings above and below some equilibrium level or trend line
stages of cyclical component : upswing, peak downswing, trough

Stochastic random component of a time series is referred to as the irregular component. This component accounts for the random nature of any time series sequence.

Models of time series

- Additive model

$$Y_t = m_t + s_t + c_t + e_t$$

- Multiplicative model

$$Y_t = m_t \, s_t \, c_t \, e_t$$

Note: e_t is the irregular random component

Remark: A particular time may have one or more deterministic components present in it; e.g. suppose a time series has trend and seasonal components, we call it is trend-seasonal model. Such a trend-seasonal model will also contain the irregular random component.

Preliminary tests of a time series

(I) Testing for existence of trend

(a) Relative ordering test

This is a non-parametric test procedure used for testing existence of trend component

Null hypothesis ag Alternate hypothesis

H_0 : no trend ag H_A : trend is present

Let the time series be denoted by $\{Y_1, \ldots, Y_n\}$ (at n time points)

Define

$$q_{ij} = \begin{cases} 1, & \text{If } Y_i > Y_j \text{ when } i<j \\ 0, & o/w \end{cases}$$

$$Q = \sum_{i}\sum_{\substack{j \\ i<j}} q_{ij}$$

Note that Q counts the # of decreasing points in the time series and is also the # discordances.

If there is no trend (increasing or decreasing) in the time series,

$$P(q_{ij}=0) = P(q_{ij}=1) = \frac{1}{2}$$

(i.e equally likely to be concordant or discordant)

$\Rightarrow$ under no trend (i.e under H_0),

$$E(Q) = \sum_{i<j}\sum E(q_{ij}) = \frac{n(n-1)}{4}$$

If observed $Q \ll E(Q)$ then it would be an indication of rising trend and If observed $Q \gg E(Q)$ then It would be an indication of a falling trend.

If obsd Q does not differ "significantly" from $E(Q)$ (under H_0) then it would indicate no trend.

Q is related with Kendall's τ, the rank correlation coefficient, through the relationship

$$\tau = 1 - \frac{4Q}{n(n-1)}$$

using the standard results of Kendall's τ, we have that, under the null hypothesis of no trend

$$E(\tau) = 0 \quad \& \quad V(\tau) = \frac{2(2n+5)}{9n(n-1)}.$$

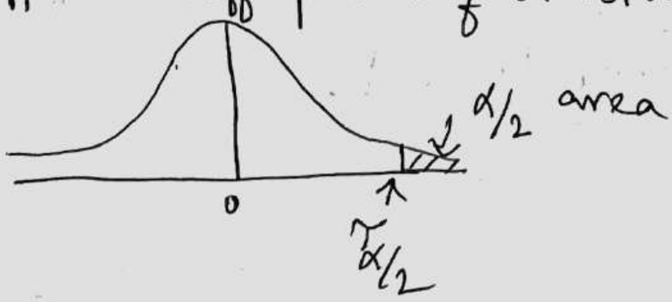
Asymptotic test for H_0: no trend is based on the statistic

$$Z = \frac{\tau - E(\tau)}{\sqrt{V(\tau)}} \overset{asym}{\sim} N(0,1) \text{ under } H_0.$$

We would reject the null hypothesis of no trend at level of significance α if

$$\text{observed } |Z| > \tau_{\alpha/2}$$

($\tau_{\alpha/2}$ is the $\alpha/2$ th upper cut off point of a standard normal dist^n, i.e



)

$P(Z > \tau_{\alpha/2}) = \alpha/2$

$Z \sim N(0,1)$

(b) Parametric test for existence of trend

This is a parametric test for existence of trend using significance of regression approach.

underlying assumption : normality

Testing for existence of linear trend

$$Y_t = \alpha + \beta t + \epsilon_t; \quad t = 1(1)n$$

$\{\epsilon_t\}$ seq of i.i.d. $N(0, \sigma^2)$

We set $H_0: \beta = 0$ ag $H_A: \beta \neq 0$

no linear trend　　　　linear trend

usual t/F test is used for the testing (Ref: Linear regression - Montgomery).

Testing for existence of quadratic ~~for~~ trend

$$Y_t = \alpha + \beta t + \gamma t^2 + \epsilon_t; \quad \epsilon_t \overset{iid}{\sim} N(0, \sigma^2)$$

We can set the following type of hypotheses testing

$H_{01}: \gamma = 0$ ag $H_{A1}: \gamma \neq 0$ (testing for 2^{nd} order polynomial time trend given that 1^{st} order time trend is accepted)

OR

$H_{02}: \beta = 0, \gamma = 0$ ag H_{A2}: not H_{02}

(joint testing for β and γ)

Joint testing is done using the standard F testing with restricted and unrestricted sum of squares

(Ref: Linear regression - Montgomery)

<u>Remark</u>: If test for existence of trend indicates significant trend component is present, then the next step would be to estimate the trend, $\hat{m}_t$ (say)

<u>Remark</u>: After trend estimation, detrend the data as $y_t - \hat{m}_t$ (for additive model) and check for existence of trend in the detrended data to ensure trend is captured properly through trend estimation.

II Testing for existence of seasonality

Friedman's test (Friedman, Journal of American Statistical Association; 1937)

This is once again a non-parametric test procedure.

Null hypothesis of "no seasonality" is tested against the alternate hypothesis of "presence of seasonality".

For testing seasonality the underlying data is either monthly or quarterly.

Steps for Friedman's test (for a monthly data)

Step 1 : Remove trend, if necessary, from the time series (This will involve estimation of trend component, if it is present - we will shortly consider the topic of estimation of trend)

Step 2 Rank the values obtained from step 1 within each year from smallest (1) to largest (12).

Let M_{ij} denote the rank corresponding to the i^{th} month for the j^{th} year

Let c denote the number of years of data

Let M_i denote the total rank for the month i (total over the c years)

$i = 1(1)12$

i.e. $M_i = \sum_j M_{ij}$

we thus obtain following rank table from the date

month \ yr	1	2	- - - -	c	M_i
1	$M_{1,1}$	$M_{1,2}$		$M_{1,c}$	M_1
2	$M_{2,1}$	$M_{2,2}$		$M_{2,c}$	M_2
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
$r=12$	$M_{12,1}$	$M_{12,2}$		$M_{12,c}$	M_{12}

Note that each column $(M_{1j}, M_{2j}, \ldots, M_{12,j})'$ for a particular j (year) is a permutation of $\{1, 2, \ldots 12\}$.

Under the null hypothesis of no seasonality, each permutation of $\{1, 2, \ldots, 12\}$ is equally likely.

In such a situation

M_{ij} can take any of $1, 2, \ldots r$ with equal probability $\frac{1}{r}$ and $E(M_{ij}) = \frac{1}{r}(1+2+\ldots+r)$

$$= \frac{r+1}{2}$$

$$\& \quad E(M_i) = E\left(\sum_{j=1}^{c} M_{ij}\right) = \frac{c(r+1)}{2}$$

Step 3 : Compute the asymptotic test statistic

$$X = 12 \sum_{i=1}^{r} \left(M_i - \frac{c(r+1)}{2}\right)^2 \Big/ cr(r+1) \quad ; \; r = 12$$

Using asymptotic theory (I would avoid deriving that), it can be shown that

$$X \overset{asym}{\sim} \chi^2_{r-1} \text{ (a central } \chi^2 \text{ on } r-1 \text{ d.f)}$$

under the null hypothesis of no seasonality

The ~~asy~~ asymptotic test would reject null hypothesis of no seasonality at level of significance α if

$$\text{obsd } (X) > \chi^2_{r-1}(\alpha)$$

where $\chi^2_{r-1}(\alpha)$ is the upper α cutoff point of a central χ^2 dist^n on $r-1$ degrees of freedom.

i.e $\quad P\left(\chi^2 > \chi^2_{r-1}(\alpha)\right) = \alpha$

$\chi^2 \sim$ central chi-square on $r-1$ d.f.

III Test for randomness of a time series

Turning point test

This is a non-parametric test procedure for testing randomness of a time series

Null hypothesis

H_0 : Series is purely random (does not contain any deterministic component)

against the alternate hypothesis

H_A : not H_0 (i.e the series is not purely random)

A turning point is defined as either a 'peak' when a value is greater than its 2 neighboring values or a 'trough' when a value is less than its 2 neighboring values

i.e. y_i is a turning point if

$$y_i > y_{i-1} \ \& \ y_i > y_{i+1} \quad - \text{peak}$$

$$\text{or} \quad y_i < y_{i-1} \ \& \ y_i < y_{i+1} \quad - \text{trough}$$

All $n-2$ time points $(2, 3, \ldots, n-1)$ are checked for being declared as a turning point

Define

$$U_i = \begin{cases} 1, & \text{If } y_i \text{ is a turning pt} \\ 0, & \text{o/w} \end{cases}$$

$$P = \sum_{i=2}^{n-1} U_i \;:\; \text{total \# of turning points}$$

To see the expected value of U_i and hence the expected value of P when the series is purely random, let us consider 3 values (y_{i-1}, y_i, y_{i+1}) leading to one such U_i

Let $(y_{(1)}, y_{(2)}, y_{(3)})$ denote the ordered values derived from (y_{i-1}, y_i, y_{i+1}). $\begin{pmatrix} y_{(1)} : \text{smallest} \\ y_{(3)} : \text{largest} \end{pmatrix}$

Now (y_{i-1}, y_i, y_{i+1}) can be any of the following 6 possible orders

$(y_{(1)}, y_{(2)}, y_{(3)})$, $(y_{(1)}, y_{(3)}, y_{(2)})$, $(y_{(2)}, y_{(1)}, y_{(3)})$, $(y_{(2)}, y_{(3)}, y_{(1)})$, $(y_{(3)}, y_{(1)}, y_{(2)})$ & $(y_{(3)}, y_{(2)}, y_{(1)})$.

Under the assumption that the series is purely random (i.e. H_0), all 6 possible outcomes are equally likely

$\Rightarrow$ in such a case $E(U_i) = 1 \times \frac{4}{6} + 0 \times \frac{2}{6} = \frac{2}{3}$

(Note that 4 out of the above 6 possible outcomes have turning points) under the null hypothesis that the series is purely random

$$E(P) = \frac{2}{3}(n-2) \quad \& \quad V(P) = \frac{16n-29}{90}$$

(Ref: Kendall, Stuart & Ord; Adv Th of Statistics)

Asymptotic test for H_0 is based on the statistic

$$Z = \frac{P - E(P)}{\sqrt{V(P)}} = \frac{P - \frac{2}{3}(n-2)}{\sqrt{\frac{16n-29}{90}}}$$

$$Z \overset{asym}{\sim} N(0,1) \text{ under } H_0.$$

We would reject H_0 at level of significance α

If observed $|Z| > \tau_{\alpha/2}$

$\left(\tau_{\alpha/2} : \text{upper } \alpha/2 \text{ cutoff point of } N(0,1) \right)$

Estimation / elimination of trend & seasonality

(I) Estimation of trend in the absence of seasonality

Consider a time series model

$$Y_t = m_t + e_t; \quad e_t \text{ is } \ni \; E(e_t) = 0$$

$$Cov(e_t, e_s) = \begin{cases} \sigma^2, & t = s \\ 0, & o/w \end{cases}$$

Method 1 : Least squares estimation of m_t

We assume that trend is polynomial trend of a particular order

e.g. $m_t = a_0 + a_1 t$ — linear time trend

$m_t = a_0 + a_1 t + a_2 t^2$ — quadratic time trend

$\vdots$

$m_t = a_0 + a_1 t + \cdots + a_k t^k$; k^{th} order polynomial trend

We obtain estimates of $(a_0, a_1, \ldots, a_k)$ by minimizing the f^n

$$g(a_0, a_1, \ldots, a_k) = \sum_{t=1}^{n} \left(y_t - \sum_{i=0}^{k} a_i t^i \right)^2$$

$$\hat{\underset{\sim}{a}}_{LS} = \underset{\underset{\sim}{a}}{\arg\min} \sum_{t=1}^{n} \left(y_t - \sum_{i=0}^{k} a_i t^i \right)^2$$

Note that the above is a simple linear model LS estimation problem and hence

$$\hat{\underset{\sim}{a}}_{LS} = (X'X)^{-1} X' \underset{\sim}{y}$$

with the model written as

$$\underset{\sim}{Y} = X \underset{\sim}{\beta} + \underset{\sim}{e}$$

$$\underset{\sim}{Y} = (Y_1, \ldots, Y_n)' ; \quad \underset{\sim}{e} = (e_1, \ldots, e_n)'$$

$$\underset{\sim}{\beta} = \underset{\sim}{a} = (a_0, a_1, \ldots, a_k)'$$

$$\underset{n \times \overline{k+1}}{X} = \begin{pmatrix} 1 & 1 & 1^2 & \cdots & 1^k \\ 1 & 2 & 2^2 & \cdots & 2^k \\ \vdots & & & & \\ 1 & n & n^2 & \cdots & n^k \end{pmatrix}$$

$$\underset{\sim}{\hat{a}}_{LS} = \underset{\underset{\sim}{a}}{\arg\min} \, (\underset{\sim}{Y} - X\underset{\sim}{\beta})'(\underset{\sim}{Y} - X\underset{\sim}{\beta})$$

& $\underset{\sim}{\hat{a}}_{LS} = (X'X)^{-1} X' \underset{\sim}{Y}$

Remark: We can arrive at the appropriate order trend in the following manner

(i) estimate linear trend

$$\hat{m}_t^{(1)} = \hat{a}_0 + \hat{a}_1 t$$

(ii) detrend the data

$$\tilde{Y}_{1_t} = Y_t - \hat{m}_t^{(1)}$$

(iii) apply relative ordering test on $\{\tilde{Y}_{1t}\}$ values; If you observe that H_0 of no trend is rejected fit

$$\hat{m}_t^{(2)} = \hat{a}_0 + \hat{a}_1 t + \hat{a}_2 t^2$$

(iv) detrend as $y_t - \hat{m}_t^{(2)}$ and apply relative ordering test to check existence of trend in the detrended series

(v) continue till the detrended series show no significant trend.

Method 2: Trend estimation using moving average

Data: ($y_1, \ldots, y_n$)

Let q be a non-negative integer

Moving average trend estimate at time point t is given by

$$\hat{m}_t = \frac{1}{2q+1} \sum_{j=-q}^{q} y_{t+j} ; \quad q+1 \le t \le n-q$$

$\hat{m}_t$: moving average trend estimate with a window length of $2q+1$

The observations within the moving average window

$$(y_{t-q}, \ldots y_{t-1}, \underset{\uparrow}{y_t}, y_{t+1}, \ldots ; y_{t+q})$$

Note : The above is referred to as equal weighted moving average with weights as $\frac{1}{2q+1}$

Note : We do not get trend values at end points i.e no trend values $\forall t < q+1$ & $\forall t > n-q$

It is so as we do not have y_t for $t<1$ & $t>n$

In such cases, we can use a symmetric padding or end point padding to get rough estimates of trend.

<u>Note</u>: For even order window length moving average, a simple mean of adjacent trend values is computed so as to have trend value correspond to time points. e.g. a 4pt ma

y_1

y_2 $\rightarrow (y_1+y_2+y_3+y_4)/4$

$y_3 \longrightarrow \hat{m}_3 = \frac{1}{2}\left\{\frac{y_1+y_2+y_3+y_4}{4} + \frac{y_2+y_3+y_4+y_5}{4}\right\}$.

y_4 $\rightarrow (y_2+y_3+y_4+y_5)/4$

y_5

y_6

$\vdots$

In general, for even order window length

$$\hat{m}_t = \frac{1}{2q}\left(\frac{1}{2}y_{t-q} + y_{t-q+1} + \cdots + y_t + \cdots + y_{t+q-1} + \frac{1}{2}y_{t+q}\right)$$

$$q+1 \le t \le n-q$$

In such cases, we can use a symmetric padding or end point padding to get rough estimates of trend.

<u>Note</u> : For even order window length moving average, a simple mean of adjacent trend values is computed so as to have trend value correspond to time points. e.g. a 4pt ma

y_1

y_2 $\to (y_1+y_2+y_3+y_4)/4$

y_3 $\to \hat{m}_3 = \frac{1}{2}\left\{ \frac{y_1+y_2+y_3+y_4}{4} + \frac{y_2+y_3+y_4+y_5}{4} \right\}$

y_4 $\to (y_2+y_3+y_4+y_5)/4$

y_5

y_6

$\vdots$

In general, for even order window length

$$\hat{m}_t = \frac{1}{2q}\left(\frac{1}{2} y_{t-q} + y_{t-q+1} + \cdots + y_t + \cdots + y_{t+q-1} + \frac{1}{2} y_{t+q} \right)$$

$$q+1 \le t \le n-q$$

Note: Moving averages are called "low-pass filters" as it filters out the rapidly fluctuating component and passes the low-frequency content (the less volatile smooth part) of the data.

$$y_t \longrightarrow \boxed{} \longrightarrow \hat{m}_t = \sum_{-\infty}^{\infty} a_j\, y_{t+j} \qquad \text{e.g. } a_j = \begin{cases} \frac{1}{2q+1}, & |j| \le q \\ 0, & \text{o/w} \end{cases}$$

filter with coeffs $\{a_j\}$ — a linear filter — equal wts ma filter

Note: $\hat{m}_t$ defined earlier is a 2-sided moving average; one can also define a one-sided moving average and trend estimate at the last pt in the window is considered.

Exponentially Weighted Moving Average (EWMA)

EWMA is an example of one-sided moving average filtering with weights decreasing exponentially inside MA window as one moves further and further away from the time pt , at which the trend is estimated.

For a fixed $\alpha \in (\frac{1}{2}, 1)$, one sided EWMA is defined as

$$\hat{m}_t = \alpha\, Y_t + (1-\alpha)\, \hat{m}_{t-1}\,; \quad t = 2(1)n$$

$$\hat{m}_1 = Y_1$$

Realize that, we have

$$\hat{m}_1 = Y_1$$

$$\hat{m}_2 = \alpha Y_2 + (1-\alpha)\hat{m}_1$$

i.e. $\hat{m}_2 = \alpha Y_2 + (1-\alpha) Y_1$

$$\hat{m}_3 = \alpha Y_3 + (1-\alpha)\hat{m}_2$$

i.e. $\hat{m}_3 = \alpha Y_3 + (1-\alpha)(\alpha Y_2 + (1-\alpha) Y_1)$

i.e. $\hat{m}_3 = \alpha Y_3 + \alpha(1-\alpha) Y_2 + (1-\alpha)(1-\alpha) Y_1$

$$\left(\alpha > \alpha(1-\alpha) > (1-\alpha)^2\right).$$

In general $\forall\, t \geq 2$

$$\hat{m}_t = \alpha Y_t + (1-\alpha)\hat{m}_{t-1}$$

i.e. $\hat{m}_t = \alpha Y_t + (1-\alpha)(\alpha Y_{t-1} + (1-\alpha)\hat{m}_{t-2})$

i.e. $\hat{m}_t = \alpha Y_t + \alpha(1-\alpha) Y_{t-1} + (1-\alpha)^2 \hat{m}_{t-2}$

i.e. $\hat{m}_t = \alpha Y_t + \alpha(1-\alpha) Y_{t-1} + (1-\alpha)^2 (\alpha Y_{t-2} + (1-\alpha)\hat{m}_{t-3})$

i.e. $\hat{m}_t = \alpha Y_t + \alpha(1-\alpha) Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + (1-\alpha)^3 \hat{m}_{t-3}$

$\vdots$

$\alpha(1-\alpha)^{t-2} Y_2 + (1-\alpha)^{t-1} \hat{m}_1$

Y_1

$$\hat{m}_t = \sum_{j=0}^{t-2} \alpha(1-\alpha)^j Y_{t-j} + (1-\alpha)^{t-1} Y_1$$

EWMA estimate of trend at time pt t

Method 3 : Trend removal/elimination by differencing

This is a method of trend elimination without estimating the trend component.

Define,

lag operator : B ; $BY_t = Y_{t-1}$; $B^j Y_t = Y_{t-j}$

First difference operator : ∇

$$\nabla Y_t = Y_t - Y_{t-1} = (1-B) Y_t$$

$$\nabla^2 Y_t = \nabla (\nabla Y_t)$$

i.e. $\nabla^2 Y_t = \nabla (Y_t - Y_{t-1})$

i.e. $\nabla^2 Y_t = \nabla Y_t - \nabla Y_{t-1}$

i.e. $\nabla^2 Y_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2})$

i.e $\nabla^2 Y_t = Y_t - 2Y_{t-1} + Y_{t-2}$

$$= (1 - 2B + B^2) Y_t$$

$$= (1-B)^2 Y_t$$

$$\nabla^j Y_t = \nabla(\nabla^{j-1} Y_t) = (1-B)^j Y_t$$

Suppose, $m_t = a + bt$ - linear time trend

then $\nabla m_t = b$

and suppose that we have the model as

$$Y_t = m_t + e_t \; ; \; E(e_t) = 0$$

$$V(e_t) = \sigma^2 < \infty$$

then $\nabla Y_t = \nabla m_t + \nabla e_t$

i.e $\nabla Y_t = b + (e_t - e_{t-1})$.

∇Y_t series is free from time trend

Sly if $m_t = a + bt + ct^2$

$$\nabla^2 m_t = 2c.$$

& $Y_t = (a + bt + ct^2) + e_t$

$$\nabla^2 Y_t = 2c + \nabla^2 e_t$$

$\nabla^2 Y_t$ would be series free from trend.

In general if m_t is a trend of degree K, then $\nabla^K y_t$ will be a series free from trend as

with $m_t = \sum_{j=0}^{K} a_j t^j$

$$\nabla^K m_t = K!\, a_K$$

and if $y_t = \sum_{j=0}^{K} a_j t^j + e_t$

$$\nabla^K y_t = (K!\, a_K) + \nabla^K e_t$$

$\{\nabla^K y_t\}$ will be a time series with mean $K!\, a_K$ and no time trend component

* Remark: For a given time series, to eliminate trend by differencing we look at the least number of differencing required to reduce it to a series which is free from trend.

(why??)

Estimation / elimination of both trend and seasonality

suppose we have a trend-seasonal model

$$Y_t = m_t + s_t + e_t \; ; \quad E(e_t) = 0$$
$$V(e_t) = \sigma^2 < \infty$$

Note that if d is the period of seasonality, then we assume that

$$\cdots = s_{t-2d} = s_{t-d} = s_t = s_{t+d} = s_{t+2d} = \cdots$$

and also $\sum_{j=1}^{d} s_j = 0$

Suppose, for illustration, that we have monthly data with 12-month period.

Let us write the time index t as

$$t = 12(j-1) + k$$

j : year no. $\quad j = 1(1)J$

k : month no. $\quad k = 1(1)12$

Y_t is written as $Y_{j,k} = Y_{12(j-1)+k}$

<u>Method 1 : Slow trend method</u>

In case the trend is slow, it is assumed that the trend remains constant during a particular year i.e. m_j for a particular year j is constant

<u>Step I</u> : estimate trend as

$$\hat{m}_j = \frac{1}{12} \sum_{k=1}^{12} Y_{j,k} \quad \left(\text{since } \sum_{k=1}^{12} s_k = 0\right)$$

$j = 1(1)J$

Step II : Estimate seasonal factors as

$$\hat{s}_K = \frac{1}{J}\sum_{j=1}^{J}\left(y_{j,K} - \hat{m}_j\right)$$

logic

data

Years

$$
\begin{array}{c|cccc}
 & 1 & 2 & \cdots & J \\
\hline
1 & Y_{1,1} & Y_{2,1} & & Y_{J,1} \\
2 & Y_{1,2} & Y_{2,2} & & Y_{J,2} \\
\text{Months}\ \vdots & \vdots & \vdots & & \vdots \\
12 & Y_{1,12} & Y_{2,12} & & Y_{J,12} \\
 & \downarrow & \downarrow & & \downarrow \\
 & \hat{m}_1 & \hat{m}_2 & & \hat{m}_J
\end{array}
\longrightarrow
\underbrace{
\begin{array}{ccc}
\text{year \#1} & & \text{year \# J} \\
\begin{pmatrix} Y_{1,1} - \hat{m}_1 \\ Y_{1,2} - \hat{m}_1 \\ \vdots \\ Y_{1,12} - \hat{m}_1 \end{pmatrix} & \cdots & \begin{pmatrix} Y_{J,1} - \hat{m}_J \\ \vdots \\ Y_{J,12} - \hat{m}_J \end{pmatrix}
\end{array}
}_{\text{Seasonal factor data.}}
$$

de-trended data

Average values, over years, for the months are (for K^{th} month), $\hat{s}_K$'s

$$\hat{S}_K = \frac{1}{J\,\cancel{12}}\sum_{j=1}^{J}\left(y_{j,K} - \hat{m}_j\right)$$

$$\hat{e}_{j,K} = y_{j,K} - \hat{m}_j - \hat{S}_K \ ; \quad j = 1(1)J, \quad K = 1(1)12$$

Note that
$$\sum_{K=1}^{12}\hat{S}_K = \frac{1}{J\,\cancel{12}}\sum_{K=1}^{12}\sum_{j=1}^{J}\left(y_{j,K} - \hat{m}_j\right)$$
$$= 0 \quad (\text{Similar to the true ones } s_K\text{'s}).$$

Method 2 : Fast trend method

In case there is a significant trend which can not be assumed to be constant for a year, we proceed in the following way :

Step I : Obtain rough estimate of trend

Use an MA filter, filter coefficients are such that seasonal component is eliminated and noise is dampened (i.e. the output process has lower variance than the original time series)

For a monthly data with period of seasonality 12, use a 12 point moving average to achieve the above.

$d = 12 \ (= 2q)$

$\underline{q=6}$

$$\hat{m}_t = \frac{1}{12}\left(\frac{1}{2}y_{t-6} + y_{t-5} + \cdots + y_t + \cdots + y_{t+5} + \frac{1}{2}y_{t+6}\right)$$

In general,

$$\hat{m}_t = \frac{1}{2q}\left(\frac{1}{2}y_{t-q} + y_{t-q+1} + \cdots + y_t + \cdots + y_{t+q-1} + \frac{1}{2}y_{t+q}\right)$$

or $$\hat{m}_t = \frac{1}{2q+1}\left(y_{t-q} + \cdots + y_t + \cdots + y_{t+q}\right)$$

depending on the period of seasonality (even or odd)

Step II : Estimation of seasonal components

For each month K $(K = 1, \ldots, 12)$; compute the average (say w_K) of deviations

$$\left\{ y_{12(j-1)+K} - \hat{m}_{12(j-1)+K} : j = 1, \ldots, J \right\}$$

Over the J years.

Estimate S_K as

$$\hat{S}_K = \omega_K - \frac{1}{12}\sum_{K=1}^{12} \omega_K$$

$$\& \quad \hat{S}_K = \hat{S}_{K-d} \quad \forall \; K > d$$

In general $\hat{S}_K = \omega_K - \frac{1}{d}\sum_{K=1}^{d} \omega_K$

<u>Remark</u>: Note that ω_Ks that we obtained here are similar to the estimates of S_Ks obtained in Method 1. However, the $\{\omega_K\}$ sequence that we have obtained here are not used as estimates of S_K under the current setup as $\sum \omega_K \neq 0$. With the centering, we have $\sum_{K=1}^{d} \hat{S}_K = 0$.

<u>Step iii</u> : Deseasonalize the data

$$d_t = y_t - \hat{S}_t \quad ; \quad t = 1, \ldots, n$$

and get $(d_1, \ldots, d_n)$

<u>Step iv</u> :

Re-estimate trend using $(d_1, \ldots, d_n)$ using any of the trend estimation methods

<u>Remark</u>: Iterate, if required.

Method 3: Elimination of trend & seasonality

using differencing, we can eliminate both trend and seasonality from the data (if they are present)

Define a lag d difference operator

$$\nabla_d y_t = y_t - y_{t-d} = (1-B^d)\, y_t$$

Apply ∇_d to $y_t = m_t + s_t + e_t$; $E(e_t) = 0$, $V(e_t) = \sigma^2 < \infty$

where d is the period of seasonality, hence

$$\cdots = s_{t-d} = s_t = s_{t+d} = \cdots$$

$$Z_t = \nabla^d y_t = \nabla^d m_t + \nabla^d s_t + \nabla^d e_t$$

$$Z_t = \underbrace{(m_t - m_{t-d})}_{\text{deterministic time trend component}} + \underbrace{(s_t - s_{t-d})}_{0} + \underbrace{(e_t - e_{t-d})}_{\text{irregular component}}$$

Trend $(m_t - m_{t-d})$ term in Z_t can be eliminated through differencing of appropriate power of ∇ operator.

Mathematical formulation of a time series

Let $(\Omega, \mathcal{F}, P)$ be a probability space and T be an index set

Def^n: A real valued time series is a real valued function $X(t,\omega)$ defined on $T \times \Omega$, $\ni$ for a fixed t, $X(t,\omega)$ $\left(= X_t(\omega) = X_t, \text{say}\right)$ is a random variable defined on $(\Omega, \mathcal{F}, P)$.

A time series is thus a collection $\{X_t : t \in T\}$ of random variables.

Mathematical formulation of a time series

Let $(\Omega, \mathcal{F}, P)$ be a probability space and T be an index set

<u>Def^n</u>: A real valued time series is a real valued function $X(t,\omega)$ defined on $T \times \Omega$, $\ni$ for a fixed t, $X(t,\omega)\ (= X_t(\omega) = X_t$, say$)$ is a random variable defined on $(\Omega, \mathcal{F}, P)$.

A time series is thus a collection $\{X_t : t \in T\}$ of random variables.

We can define joint distribution function of a finite set of random variables $\{X_{t_1}, X_{t_2}, \ldots, X_{t_n}\}$ from the collection $\{X_t : t \in T\}$ is

$$F_{X_{t_1}, \ldots, X_{t_n}}(x_1, \ldots, x_n) = P(X_{t_1} \le x_1, \ldots, X_{t_n} \le x_n)$$

Concept of stationarity of time series: $\{X_t\}$ process is $\ni$ "statistical properties" of the process do not change over time, i.e. realizations come from a stable physical system which has achieved a "steady-state statistical equilibrium" mode.

<u>Remark</u>: Different forms of stationarity concept is defined under different paradigms of quantifying "statistical equilibrium".

Important definitions of stationarity

(I) Strict stationary: A process $\{X_t\}$ is said to be strict stationary or completely stationary, if for all $n \geq 1$, any admissible $t_1, t_2, \ldots, t_n$ and any k (integer), the joint distn of $(X_{t_1}, \ldots, X_{t_n})$ is identical with the joint distn of $(X_{t_1+k}, \ldots, X_{t_n+k})$.

i.e. $F_{X_{t_1}, \ldots X_{t_n}}(x_1, \ldots, x_n) = F_{X_{t_1+k}, \ldots X_{t_n+k}}(x_1, \ldots, x_n)$ $\forall n \geq 1$ $\forall t_1, \ldots t_n$ & $\forall k \geq 1$

i.e. the jt distn of any finite set of r.v.s is invariant w.r.t. time shift.

(II) Stationarity upto order m: A time series process $\{X_t\}$ is said to be stationary upto order m, if, for all $n \geq 1$, any admissible $t_1, t_2, \ldots, t_n$ and any integer k, all the joint moments upto order m of $\{X_{t_1}, \ldots, X_{t_n}\}$ exist and equal the corresponding joint moments upto order m of $\{X_{t_1+k}, \ldots, X_{t_n+k}\}$

i.e. $E\left(X_{t_1}^{m_1} X_{t_2}^{m_2} \cdots X_{t_n}^{m_n}\right) = E\left(X_{t_1+k}^{m_1} X_{t_2+k}^{m_2} \cdots X_{t_n+k}^{m_n}\right)$

$\forall n \geq 1$

$\forall k$ and $\forall$ integer $m_1, m_2, \ldots, m_n$ $(m_i \geq 0)$

$\ni \sum_{i=1}^{n} m_i \leq m$

Note : In particular, setting $m_2 = m_3 = \cdots = m_n = 0$, we get that for any t and $\forall\, m_1 \le m$

$$E\left(X_{t_1}^{m_1}\right) = E\left(X_{t_1+K}^{m_1}\right) \quad \forall K$$

$$\Rightarrow E\left(X_{t_1}^{m_1}\right) = \underbrace{E\left(X_0^{m_1}\right)}_{\text{const indep of } t} \quad \text{setting } K = -t$$

Also $$E\left(X_t^{m_1} X_s^{m_2}\right) = E\left(X_{t+K}^{m_1} X_{s+K}^{m_2}\right) \quad \forall K$$

$\forall\, m_1, m_2 \ni$
$m_1 + m_2 \le m$

$$= \underbrace{E\left(X_0^{m_1} X_{s-t}^{m_2}\right)}_{\substack{\text{indep of } t \text{ and is a } f^n \text{ of} \\ (s-t) \text{ only}}} \quad \text{for } K = -t$$

Important special cases of "stationarity upto order m"

(i) Order 1 stationary

$$E X_t \text{ exists } \& \ E X_t = \mu \ \forall t$$

This is referred to as "mean stationarity"
Thus a time series $\{X_t\}$ is mean stationary if EX_t exists and is indep of t.

(ii) Order 2 stationary : EX_t^2 and $E X_t X_s$ exist $\forall t, s$ and

(a) $E X_t = \mu \leftarrow$ const, indep of t

(b) $EX_t^2 = \mu_2'$ const, indep of t; hence

$V(X_t) = \sigma^2$ is also indep of t

& (c) $E(X_t X_s) = E(X_{t+h} X_{s+h}) = \underline{E(X_0 X_{s-t})}$ — (*)

f^n of $(s-t)$ only and indep of t

$(s-t)$: time difference

(*) $\Rightarrow$ $Cov(X_t, X_s) = \underline{E(X_t X_s) - \mu^2}$

f^n of time difference $(s-t)$ only and is indep of t

<u>Remark</u>: The order 2 stationarity is also referred to as covariance stationary (or weak stationary or stationary in the wide sense). This form of stationarity is the most widely used form of stationarity.

Remark : If $\{X_t\}$ is strict stationary, then $\{X_t\}$ is also covariance stationary provided moments upto order 2 exist for the joint distribution.

Pf : For $n=1$, def^n of strict stationarity implies that X_t has the same dist^n for each t in the index set,

$$\text{i.e. } F_{X_{t_1}}(x) = F_{X_{t_1+k}}(x) \quad \forall k \qquad -(*)$$

If $EX_t^2 < \infty$, then $(*)$ in particular implies that EX_t & VX_t exists and are indep of t (as the dist^{n}s are identical $\forall t$).

Further, take $n=2$, def^n of strict stationarity implies that the jt dist^n of (X_{t_1}, X_{t_2}) and (X_{t_1+h}, X_{t_2+h}) are identical

$$\text{i.e. } (X_{t_1}, X_{t_2}) \stackrel{d}{=} (X_{t_1+h}, X_{t_2+h})$$

$$\Rightarrow \operatorname{Cov}(X_{t_1}, X_{t_2}) = \operatorname{Cov}(X_{t_1+h}, X_{t_2+h}) \quad \forall t_1, t_2 \;\; \forall h$$

$$= \operatorname{Cov}(X_0, X_{t_2-t_1}); \quad h = -t_1$$

f^n of $(t_2 - t_1)$ only

$\Rightarrow \{X_t\}$ is covariance stationary.

<u>Remark</u>: Converse of the previous result (strict stationarity + existence of moments upto order 2 $\Rightarrow$ weak stationarity) is NOT true (in general)

<u>A counter example</u> for weak stationarity $\not\Rightarrow$ strict stationarity

Let the time series $\{X_t\}$ be a seq of indep r.v.s $\ni$

$$X_t \sim \begin{cases} \exp(1), & \text{if } t \text{ is odd} \\ N(1,1), & \text{if } t \text{ is even} \end{cases}$$

$$EX_t = 1 \;\forall t$$

$$Cov(X_t, X_{t+h}) = \begin{cases} 1, & \text{if } h=0 \\ 0 & \text{o/w} \end{cases} \quad \leftarrow \text{indep of } t, \; f^n \text{ of } h \text{ only}$$

$\Rightarrow \{X_t\}$ is Covariance stationary

Realize that dist^n of X_1 & X_2 are different and hence $\{X_t\}$ can not be strict stationary.

<u>Remark</u>: Can you give an example of a time series which is strict stationary but is not Covariance stationary?

Example is obvious!! Isn't it?

<u>Remark</u>: There is a special (particular) type of time series for which covariance stationarity $\Rightarrow$ strict stationarity.

<u>Def^n</u>: Gaussian process

A time series $\{X_t\}$ is said to be Gaussian if for any n and any admissible $t_1, \ldots, t_n$, the jt dist^n of $(X_{t_1}, \ldots, X_{t_n})$ is multivariate normal; i.e. $(X_{t_1}, \ldots, X_{t_n})' \sim N_n$

[<u>Def^n of multivariate normal</u>

A multivariate random vector $\underset{\sim}{X}_{p\times 1}$ with mean vector $E(\underset{\sim}{X}) = \underset{\sim}{\mu}$ and covariance matrix $E(\underset{\sim}{X} - \underset{\sim}{\mu})(\underset{\sim}{X} - \underset{\sim}{\mu})' = \Sigma = Cov(\underset{\sim}{X})$ is said to follow a multivariate normal iff $\forall \underset{\sim}{\alpha} \in \mathbb{R}^p$ $(\underset{\sim}{\alpha} \neq \underset{\sim}{0})$

$\underset{\sim}{\alpha}' \underset{\sim}{X} \sim N_1$, i.e. $\underset{\sim}{X} \sim N_p$ iff $\underset{\sim}{\alpha}'\underset{\sim}{x} \sim$ univariate normal $\forall \underset{\sim}{\alpha} \in \mathbb{R}^p$ $(\underset{\sim}{\alpha} \neq \underset{\sim}{0})$]

We write $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$

If $\Sigma > 0$, then the jt p.d.f. of $\underset{\sim}{X}$ is

$$f_{\underset{\sim}{X}}(x) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\underset{\sim}{x} - \underset{\sim}{\mu})' \Sigma^{-1} (\underset{\sim}{x} - \underset{\sim}{\mu})\right)$$

Note that N_p dist^n is completely specified by $\underset{\sim}{\mu}$ & Σ

If $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$

Covariance stationary Gaussian process

Suppose $\{X_t\}$ is a Gaussian process and further that the process $\{X_t\}$ is covariance stationary

i.e. $E X_t = \mu \;\; \forall t$

$\mathrm{Cov}(X_t, X_{t+h}) = f^n$ of h only $\forall t$; $V X_t = \sigma^2 \;\forall t$ (finite)

Since $\{X_t\}$ is Gaussian, the jt distr^n of

$$\underset{\sim}{Z} = (X_{t_1}, \ldots, X_{t_n})' \text{ is } N_n(\underset{\sim}{\mu}, \Sigma)$$

where
$$\underset{\sim}{\mu} = E(\underset{\sim}{Z}) = \begin{pmatrix} \mu \\ \mu \\ \vdots \\ \mu \end{pmatrix}$$

$$\cancel{\Sigma = \mathrm{Cov}(\underset{\sim}{Z})} =$$

$$\Sigma = \mathrm{Cov}(\underset{\sim}{Z}) = \begin{pmatrix} V(X_{t_1}) & \mathrm{Cov}(X_{t_1}, X_{t_2}) & \cdots & \mathrm{Cov}(X_{t_1}, X_{t_n}) \\ & V(X_{t_2}) & \cdots & \mathrm{Cov}(X_{t_2}, X_{t_n}) \\ & & \ddots & \vdots \\ & & & V(X_{t_n}) \end{pmatrix}$$

$$\text{i.e. } \Sigma = \begin{pmatrix} \sigma^2 & f^n \text{ of } (t_2 - t_1) & \cdots & f^n \text{ of } (t_n - t_1) \\ & \sigma^2 & \cdots & f^n \text{ of } (t_n - t_2) \\ & & \ddots & \vdots \\ & & & \sigma^2 \end{pmatrix}$$

Consider now shifted set of r.vs

$\underset{\sim}{Y} = (X_{t_1+k}, \ldots, X_{t_n+k})$ for any int k

$\underset{\sim}{Y} \sim N_n$ with $E(\underset{\sim}{Y}) = \begin{pmatrix} \mu \\ \mu \\ \vdots \\ \mu \end{pmatrix}$.

$$Cov(\underset{\sim}{Y}) = \begin{pmatrix} V(X_{t_1+k}) & Cov(X_{t_1+k}, X_{t_2+k}) & \cdots & Cov(X_{t_1+k}, X_{t_n+k}) \\ & V(X_{t_2+k}) & & Cov(X_{t_2+k}, X_{t_n+k}) \\ & & \ddots & \\ & & & V(X_{t_n+k}) \end{pmatrix}$$

$$= \begin{pmatrix} \sigma^2 & f^n \text{ of } (t_2-t_1) \text{ only} & \cdots & f^n \text{ of } (t_n-t_1) \text{ only} \\ & \sigma^2 & & f^n \text{ of } (t_n-t_2) \text{ only} \\ & & \ddots & \\ & & & \sigma^2 \end{pmatrix}$$

$$= \Sigma$$

$$\Rightarrow (X_{t_1}, \ldots, X_{t_n}) \overset{d}{=} (X_{t_1+k}, \ldots, X_{t_n+k})$$

identical in $dist^n$

$\Rightarrow \{X_t\}$ is strict stationary

Thus, if $\{X_t\}$ is Gaussian, then if $\{X_t\}$ is covariance stationary $\Rightarrow \{X_t\}$ is strict stationary

<u>Note</u>: This is a sp case when cov stat $\Rightarrow$ strict stationary

Some examples

Example 1 : $\{X_t\}$ is a seq of i.i.d. random variables

$\{X_t\}$ is strict stationary

Is $\{X_t\}$ covariance stationary

Alternately, suppose $\{X_t\}$ is a seq of i.i.d. random variables with finite variance σ^2. Then $\{X_t\}$ is clearly covariance stationary and strict stationary.

Example 2 :

$$Y_t = \alpha + \beta t + \epsilon_t$$

$\{\epsilon_t\}$ is a seq of i.i.d. random variables

$$\Rightarrow E(\epsilon_t) = 0 \ \forall t \ \& \ Cov(\epsilon_t, \epsilon_s) = \begin{cases} \sigma^2, & t = s \\ 0, & o/w \end{cases}$$

Seq $\{Y_t\}$ is thus an independent random variable

with $EY_t = \alpha + \beta t \leftarrow f^n$ of t

$\& \ VY_t = \sigma^2$

As EY_t is a f^n of t, $\{Y_t\}$ is not even mean stationary and hence is not covariance stationary.

Example 2 : $\{X_t\}$ is $\ni$

$$X_t = A\cos\omega t + B\sin\omega t$$

A & B are uncorrelated r.v.s with

mean zero and variance σ^2

$\omega \in (-\pi, \pi)$ and is fixed

$$E\,X_t = 0 \quad \forall t \quad \text{—(i)}$$

$$\begin{aligned} Cov(X_t, X_{t+h}) &= Cov\big(A\cos\omega t + B\sin\omega t,\ A\cos\omega(t+h) \\ &\qquad + B\sin\omega(t+h)\big) \\ &= \sigma^2\big(\cos\omega t\cos\omega(t+h) + \sin\omega t\sin\omega(t+h)\big) \\ &= \underbrace{\sigma^2\cos\omega h}_{} = \gamma_X(h) \quad \text{—(ii)} \end{aligned}$$

f^n of h only and is indep of t

(i) & (ii) $\Rightarrow$ $\{X_t\}$ is covariance stationary

(*) What happens If I make A & B to have identical but non zero mean ??

Example 3

$\{X_t\}$ is $\ni$

$$X_t = \sum_{j=1}^{K} \left(A_j \cos(j\omega t) + B_j \sin(j\omega t) \right)$$

$\{A_j\}_{j=1}^{K}$ seq of independent $N(0, \sigma^2)$

$\{B_j\}_{j=1}^{K}$ seq of indep $N(0, \sigma^2)$

Further $\{A_j\}$ & $\{B_j\}$ sequences are mutually indep

$$E X_t = 0 \ \forall t \quad \text{(i)}$$

$$\begin{aligned}
\operatorname{Cov}(X_t, X_{t+h}) &= E X_t X_{t+h} \\
&= E \left(\sum_{j=1}^{K} \{ A_j \cos(j\omega t) + B_j \sin(j\omega t) \} \right) \\
&\quad \left(\sum_{j=1}^{K} \{ A_j \cos(j\omega(t+h)) + B_j \sin(j\omega(t+h)) \} \right). \\
&= \sum_{j=1}^{K} \{ E(A_j^2) \cos(j\omega t) \cos(j\omega(t+h)) \\
&\quad + E(B_j^2) \sin(j\omega t) \sin(j\omega(t+h)) \}
\end{aligned}$$

(remaining terms are zero due to serial & mutual independence of $\{A_j\}$ & $\{B_j\}$ seqs)

$$= \sigma^2 \sum_{j=1}^{K} \cos(j\omega h) \leftarrow \text{indep of } t;\ f^n \text{ of } h \text{ only} \quad \text{— (ii)}$$

(i) & (ii) $\Rightarrow \{X_t\}$ is covariance stationary

Example 3 (contd)

Consider the random ~~vec~~ vector $\underset{\sim}{Z} = \begin{pmatrix} X_{t_1} \\ \vdots \\ X_{t_n} \end{pmatrix}$ for any n and any adm $t_1, \dots t_n$

$$\begin{pmatrix} X_{t_1} \\ \vdots \\ X_{t_n} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{K} \left(A_j \cos(j\omega t_1) + B_j \sin(j\omega t_1)\right) \\ \vdots \\ \sum_{j=1}^{K} \left(A_j \cos(j\omega t_n) + B_j \sin(j\omega t_n)\right) \end{pmatrix}$$

$\forall \underset{\sim}{\alpha} \in \mathbb{R}^n$; $\underset{\sim}{\alpha}' \underset{\sim}{Z} = \alpha_1 \left(\sum_{j=1}^{K} \left(A_j \cos(j\omega t_1) + B_j \sin(j\omega t_1)\right)\right)$

$$+ \dots$$

$$\vdots$$

$$+ \alpha_n \left(\sum_{j=1}^{K} \left(A_j \cos(j\omega t_1) + B_j \sin(j\omega t_1)\right)\right)$$

$$= \beta_1 A_1 + \beta_2 A_2 + \dots + \beta_K A_K$$

$$+ \gamma_1 B_1 + \dots + \gamma_K B_K \quad - (*)$$

($\beta_1, \dots, \beta_K$ and $\gamma_1, \dots \gamma_K$ are consts).

(*) is a linear combination of indep. N_1 random variables

$\Rightarrow$ $\underset{\sim}{\alpha}' \underset{\sim}{Z}$ is a linear combination of indep N_1 random variables

$\Rightarrow \underset{\sim}{\alpha}' \underset{\sim}{Z} \sim N_1 \quad \forall \underset{\sim}{\alpha} \in \mathbb{R}^n \; (\underset{\sim}{\alpha} \neq \underset{\sim}{0})$

$$\Rightarrow \begin{pmatrix} X_{t_1} \\ \vdots \\ X_{t_n} \end{pmatrix} \sim N_n(\underset{\sim}{0}, \Sigma) \quad - (*^2)$$

$$\Sigma = \begin{pmatrix} K\sigma^2 & \sigma^2 \sum_{j=1}^{K} \cos(j\omega(t_2-t_1)) & \cdots & \sigma^2 \sum_{j=1}^{K} \cos(j\omega(t_n-t_1)) \\ & K\sigma^2 & & \sigma^2 \sum_{j=1}^{K} \cos(j\omega(t_n-t_2)) \\ & & \ddots & \\ & & & K\sigma^2 \end{pmatrix}$$

$(*^2) \Rightarrow \{X_t\}$ is a Gaussian process.

Further, we have already proved that $\{X_t\}$ is Covariance stationary

$\Rightarrow \{X_t\}$ is strict stationary also.

Note that
$$\begin{pmatrix} X_{t_1+K} \\ \vdots \\ X_{t_n+K} \end{pmatrix} \sim N_n(\underset{\sim}{0}, \Sigma) \quad \forall \text{ int } K$$

Example 4

$X_t = Z_t + \theta Z_{t-1}$; $\{Z_t\}$ is a seq of i.i.d. zero mean finite variance σ^2 process.

$E X_t = 0 \;\; \forall t$ — (i)

$$\begin{aligned}
Cov(X_t, X_{t+h}) &= Cov(Z_t + \theta Z_{t-1}, Z_{t+h} + \theta Z_{t+h-1}) \\
&= Cov(Z_t, Z_{t+h}) + \theta\, Cov(Z_t, Z_{t+h-1}) \\
&\quad + \theta\, Cov(Z_{t-1}, Z_{t+h}) + \theta^2 Cov(Z_{t-1}, Z_{t+h-1}) \\
&= \sigma^2 I_{(0)}(h) + \theta\sigma^2 I_{(1)}(h) \\
&\quad + \theta\sigma^2 I_{(-1)}(h) + \theta^2\sigma^2 I_{(0)}(h) \\
&= \begin{cases} \sigma^2(1+\theta^2), & \text{if } h = 0 \\ \theta\sigma^2, & \text{if } h = \pm 1 \\ 0, & \text{o/w} \end{cases} \quad \text{— (ii)}
\end{aligned}$$

indep of t ; f^n of h only

(i) & (ii) $\Rightarrow \{X_t\}$ is covariance stationary

Example 5

Let $\{X_t\}$ be a seq of i.i.d. random variables with

$$E X_t = 0 \quad \text{and} \quad V X_t = \sigma^2 < \infty$$

Define
$$S_t = \sum_{i=1}^{t} X_i$$

$$E S_t = 0 \quad \forall t$$

$\Rightarrow \{S_t\}$ is a mean stationary process

$$V S_t = t\sigma^2$$

$$\text{Cov}(S_{t+h}, S_t) = \begin{cases} t\sigma^2, & \text{if } h > 0 \\ (t+h)\sigma^2, & \text{if } h \le 0 \end{cases}$$

$\Rightarrow \{S_t\}$ is not Covariance stationary

Note that $\quad S_t = S_{t-1} + X_t \leftarrow$ Random-Walk model

Further, although $\{S_t\}$ is not Covariance stationary

$\nabla S_t = S_t - S_{t-1} = X_t$ is Covariance stationary

Remark: If $\{X_t\}$ is Covariance stationary, then

(i) $Y_t = \alpha X_t \; ; \forall \alpha \in \mathbb{R}$, is covariance stationary

(ii) $Y_t = X_t \pm \alpha \; ; \alpha \in \mathbb{R}$ is covariance stationary

(iii) $$Y_t = \begin{cases} X_t, & \text{If } t \text{ is odd} \\ X_t + \alpha, & \text{If } t \text{ is even} \end{cases}$$

$$E\,Y_t = \begin{cases} E X_t, & t \text{ is odd} \\ \alpha + E X_t, & t \text{ is even} \end{cases}$$

$\Rightarrow$ $\{Y_t\}$ is not covariance stationary (not even mean stationary)

Sly, If $Y_t = \begin{cases} X_t, & t \text{ odd} \\ \alpha X_t, & t \text{ even} \end{cases}$

$\{Y_t\}$ is not covariance stationary (as $V Y_t = \begin{cases} V X_t, & \text{odd} \\ \alpha^2 V X_t, & \text{even} \end{cases}$)

$\{Y_t\}$ is not mean stationary if $E X_t \neq 0$

<u>Remark</u>: If $\{X_t\}$ and $\{Y_t\}$ are covariance stationary processes and $\{X_t\}$ and $\{Y_t\}$ are independent, then

$Z_t = X_t + Y_t$ is $\Rightarrow$

$E Z_t = \mu_x + \mu_y \quad \forall t$ (indep of t)

$$\text{Cov}(Z_{t+h}, Z_t) = \underbrace{\text{Cov}(X_{t+h}, X_t)}_{f^n \text{ of } h \text{ only}} + \underbrace{\text{Cov}(Y_{t+h}, Y_t)}_{f^n \text{ of } h \text{ only}}$$

$\Rightarrow$ $\text{Cov}(Z_{t+h}, Z_t)$ is a f^n of h only

$\Rightarrow$ $\{Z_t\}$ is also covariance stationary

Note that if $\{X_t\}$ & $\{Y_t\}$ are ~~to~~ uncorrelated covariance stationary processes, then also $\{Z_t\}$ is covariance stationary

<u>Remark</u>: If $Z_t = X_t + Y_t$ is covariance stationary, then it is not necessary that $\{X_t\}$ & $\{Y_t\}$ are covariance stationary

Counter example !!

<u>Remark</u> We can also define a complex valued time series process in the following way

Let $\{X_t\}$ & $\{Y_t\}$ be 2 real valued time series processes. Define

$$U_t = X_t + i Y_t \; ; \quad i = \sqrt{-1}$$

$\{U_t\}$ is a complex valued time series with the properties:

(i) $E U_t = E X_t + i E Y_t$

(ii) $Cov(U_{t+h}, U_t) = E\left(U_{t+h} - E(U_{t+h})\right)^* \left(U_t - E(U_t)\right)$

$\{U_t\}$ is said to be covariance stationary if

(a) $E(U_t) = \mu$ indep of t

& (b) $Cov(U_{t+h}, U_t)$ is a f^n of h only

<u>Example</u>: $Y_t = A e^{i\omega t}$

A is random variable $\ni E(A) = 0 \; ; V(A) = \sigma^2 < \infty$

$$Y_t = A \cos \omega t + i A \sin \omega t$$

i.e $Y_t = U_t + i V_t \; ; \; U_t = A \cos \omega t$

$V_t = A \sin \omega t$

$\{U_t\}$ & $\{V_t\}$ are real valued time series.

$$E Y_t = E U_t + i E V_t = 0 \quad \forall t \quad -(i)$$

$$\begin{aligned} Cov(Y_{t+h}, Y_t) &= E(Y_{t+h}^* Y_t) \\ &= E\left(A e^{-i\omega(t+h)} A e^{i\omega t}\right) \\ &= E\left(A^2 e^{-i\omega h}\right) \\ &= \underline{\sigma^2 e^{-i\omega h}} \quad -(ii) \end{aligned}$$

indep of t; f^n of h only

(i) & (ii) $\Rightarrow$ $\{Y_t\}$ is Covariance stationary complex valued time series.

<u>Remark</u>: Suppose

$$X_t = U_t + i V_t$$

For stationarity of $\{X_t\}$, is it necessary that $\{U_t\}$ & $\{V_t\}$ need to be Covariance stationary? Think about it.

Auto Covariance function (ACVF) of a stationary process

Def^n: Let $\{X_t : t \in T\}$ be a covariance stationary time series process. The ACVF of $\{X_t\}$ at lag h is given by

$$\gamma_X(h) = Cov(X_{t+h}, X_t) = Cov(X_t, X_{t+h})$$

i.e. $\gamma_X(h) = E(X_{t+h} - \mu)(X_t - \mu)$

$$h = 0, \pm 1, \pm 2, \ldots$$

$$(\mu = EX_t)$$

Properties of ACVF

Property 1: $\gamma_X(0) \geq 0$ — trivial

$\quad = V(X_t)$

Property 2 $\quad |\gamma_X(h)| \leq \gamma_X(0) \quad \forall h$

Note that by Cauchy-Schwarz inequality, we have

$$|Cov(X_{t+h}, X_t)| = |E(X_{t+h} - \mu)(X_t - \mu)|$$

$$\leq \left(E(X_{t+h} - \mu)^2\right)^{1/2} \left(E(X_t - \mu)^2\right)^{1/2}$$

$$= (V(X_{t+h}))^{1/2} (V(X_t))^{1/2}$$

i.e. $|Cov(X_{t+h}, X_t)| \leq (V(X_{t+h}))^{1/2} (V(X_t))^{1/2} \quad \forall h$

i.e. $|\gamma_X(h)| \leq \gamma_X(0) \quad \forall h$

Property 3 : $\gamma_X(\cdot)$ is even f^n

$$\gamma_X(h) = E(X_{t+h}-\mu)(X_t-\mu)$$
$$= E(X_{t+h-h}-\mu)(X_{t-h}-\mu) \quad [\because \{X_t\} \text{ is cov stat}]$$
$$= E(X_t-\mu)(X_{t-h}-\mu)$$

i.e. $\gamma_X(h) = \gamma_X(-h) \quad \forall h$

Property 4 : $\gamma_X(\cdot)$ is non negative definite

[A real valued f^n on integers ($f : Z \to R$) is said to be non-negative definite iff

$$\sum_{i,j=1}^{n} a_i f(t_i - t_j) a_j \geq 0$$

$\forall$ positive int n

$\forall \underset{\sim}{a} \in R^n \quad \forall \underset{\sim}{t} = (t_1, \ldots t_n)' \in Z^n$]

Proof of property 4 : Let $\{X_t\}$ be a covariance stationary process

$$\underset{\sim}{a} = (a_1, \ldots, a_n)' \in R^n$$
$$\underset{\sim}{t} = (t_1, \ldots, t_n)' \in Z^n \,;\; Z = \{0, \pm 1, \pm 2, \ldots\}$$

Define $U_{t_i} = X_{t_i} - \mu \;;\; i = 1(1)n \;;\; \mu = E(X_{t_i}) \;\forall i$

$$\underset{\sim}{U}_t = (U_{t_1}, \ldots, U_{t_n})'$$

Note that $\quad 0 \leq V(\underset{\sim}{a}' \underset{\sim}{U}_t)$

$$V(\underset{\sim}{a}'\underset{\sim}{U}_t) = E(\underset{\sim}{a}'\underset{\sim}{U}_t - E(\underset{\sim}{a}'\underset{\sim}{U}_t))(\underset{\sim}{a}'\underset{\sim}{U}_t - E(\underset{\sim}{a}'\underset{\sim}{U}_t))'$$
$$= E(\underset{\sim}{a}'\underset{\sim}{U}_t)(\underset{\sim}{a}'\underset{\sim}{U}_t)'$$
$$= \underset{\sim}{a}' E(\underset{\sim}{U}_t \underset{\sim}{U}_t')\underset{\sim}{a}$$

$$= \underset{\sim}{a}' \Gamma_n \underset{\sim}{a}$$

Where, Γ_n = Covariance matrix of $(X_{t_1}, \ldots, X_{t_n})$

$$= \begin{pmatrix} Cov(X_{t_1}, X_{t_1}) & Cov(X_{t_1}, X_{t_2}) & \cdots & Cov(X_{t_1}, X_{t_n}) \\ & Cov(X_{t_2}, X_{t_2}) & \cdots & Cov(X_{t_2}, X_{t_n}) \\ & & \ddots & \vdots \\ & & & Cov(X_{t_n}, X_{t_n}) \end{pmatrix}$$

$$\underset{\text{stationarity}}{\overset{\text{using}}{=}} \begin{pmatrix} \gamma_X(0) & \gamma_X(t_1-t_2) & \cdots & \gamma_X(t_1-t_n) \\ & \gamma_X(0) & \cdots & \gamma_X(t_2-t_n) \\ & & \ddots & \vdots \\ & & & \gamma_X(t_n-t_n) = \gamma_X(0) \end{pmatrix}$$

Thus $\underset{\sim}{a}' \Gamma_n \underset{\sim}{a} = V(\underset{\sim}{a}' \underset{\sim}{U}_t) \geq 0 \quad \forall \underset{\sim}{a} \quad \forall \underset{\sim}{t}$

$$\Rightarrow \sum_{i,j=1}^{n} a_i \gamma_X(t_i - t_j) a_j \geq 0 \quad \forall \underset{\sim}{a} \quad \forall \underset{\sim}{t}$$

$\Rightarrow \gamma_X(.)$ is n.n.d.

<u>Remark</u> : Converse of Property 3 & Property 4 taken together is also true.

i.e. a real valued f^n defined on the set of integers which is even and n.n.d is a Covariance function of a Covariance stationary time series

Remark: In light of the previous remark and the properties (3 & 4) of ACVF, we have the following characterization of ACVF:

"A real valued function defined on integers is the ACVF of a covariance stationary time series Iff it is even and n.n.d."

Remark $\{X_t\}$ & $\{Y_t\}$ u.c. cov stat

Auto Correlation Function (ACF)

ACF of a stationary time series is given by

$$\rho_X(h) = \text{Corr}(X_{t+h}, X_t) = \frac{\gamma_X(h)}{\gamma_X(0)}.$$

$$h = 0, \pm 1, \pm 2, \ldots$$

Using the properties of ACVF, $\gamma_X(\cdot)$, we can easily prove the following properties of ACF

(i) $\rho_X(0) = 1$

(ii) $|\rho_X(h)| \leq 1 \quad \forall h$

$Z_t = X_t + Y_t$

(iii) $\rho_X(h) = \rho_X(-h)$

(iv) $\rho_X(\cdot)$ is n.n.d.

(V) If X_{t+h} & X_t are indep then $\rho_X(h) = 0$.

Remark : ACVF & ACF for complex valued time series

Let $\{X_t\}$ be a complex valued covariance stationary time series

$$X_t = U_t + i V_t$$

ACVF of $\{X_t\}$: $\gamma_X(h) = E\left(X_{t+h} - \mu\right)^* (X_t - \mu)$

$$\mu = E X_t$$

Note that $\gamma_X(h)$ is complex valued $(\forall h \neq 0)$, $\gamma_X(0)$ is real valued.

ACF of $\{X_t\}$: $\rho_X(h) = \dfrac{\gamma_X(h)}{\gamma_X(0)}$.

Estimation of μ and $\gamma_X(h)$ for covariance stationary process

Let $X_1, \ldots, X_n$ be a sample of size n from a covariance stationary $\{X_t\}$ with $E X_t = \mu$ (unknown) and ACVF $\gamma_X(h) = E(X_{t+h} - \mu)(X_t - \mu)$ $\forall h$ (unknown)

Estimation of μ : μ is estimated by sample mean,

$$\bar{X}_n = \frac{1}{n} \sum_{t=1}^{n} X_t$$

Realize that $E(\bar{X}_n) = \frac{1}{n} \sum_{t=1}^{n} E X_t = \mu$,

thus $\bar{X}_n$ is an unbiased estimator of μ

Expression for $V(\bar{X}_n)$ (It is not $\frac{\sigma^2}{n}$!!)

$$V(\bar{X}_n) = V\left(\frac{1}{n}\sum_{t=1}^{n} X_t\right)$$

$$= E\left(\frac{1}{n}\sum_t X_t - \mu\right)^2$$

$$= E\left(\frac{1}{n}\sum_{t=1}^{n}(X_t - \mu)\right)^2$$

$$= \frac{1}{n^2}E\Big((x_1-\mu)+\cdots+(x_n-\mu)\Big)\Big((x_1-\mu)+\cdots+(x_n-\mu)\Big)$$

$$= \frac{1}{n^2}\Big[\left(\gamma_0+\gamma_1+\cdots+\gamma_{n-1}\right)$$
$$+\left(\gamma_1+\gamma_0+\gamma_1+\cdots+\gamma_{n-2}\right)$$
$$\cdots\cdots$$
$$+\left(\gamma_{n-1}+\gamma_{n-2}+\cdots+\gamma_0\right)\Big].$$

$2(n-2)+2 = 2(n-1)$

$$= \frac{1}{n^2}\Big[n\gamma_0 + 2(n-1)\gamma_1 + 2(n-2)\gamma_2 + \cdots$$
$$\cdots + 2(n-(n-1))\gamma_{n-1}\Big]$$

$$= \frac{1}{n^2}\sum_{i-j=-n}^{n}(n-|i-j|)\gamma_{i-j}$$

i.e. $$V(\bar{X}_n) = \frac{1}{n^2}\sum_{h=-n}^{n}(n-|h|)\gamma_h = \frac{1}{n}\sum_{|h|\le n}\left(1-\frac{|h|}{n}\right)\gamma_h$$

Estimation of $\gamma(\cdot)$ and $\rho(\cdot)$

Suppose μ is known, an unbiased estimator of $\gamma(h)$ is

$$\hat{\gamma}_\mu^*(h) = \frac{1}{n-h}\sum_{t=1}^{n-h}(X_t-\mu)(X_{t+h}-\mu)$$

$h = 0(1)n-1$

as
$$E\left(\hat{\gamma}_\mu^*(h)\right) = \frac{1}{n-h}\sum_{t=1}^{n-h}E(X_t-\mu)(X_{t+h}-\mu)$$
$$= \frac{(n-h)\gamma_h}{(n-h)} = \gamma_h$$

An alternate estimator (difference in the divisor only)

$$\hat{\gamma}_\mu(h) = \frac{1}{n}\sum_{t=1}^{n-h}(X_t-\mu)(X_{t+h}-\mu)$$

$$E\left(\hat{\gamma}_\mu(h)\right) = \frac{n-h}{n}\gamma_h \neq \gamma_h$$

Bias: $E\left(\hat{\gamma}_\mu(h)\right) - \gamma_h = -\frac{h}{n}\gamma_h \to 0$ as $n \to \infty$

$\Rightarrow \hat{\gamma}_\mu(h)$ is unbiased in the limit, although

It is a biased estimator

Note that

$$\hat{\Gamma}_n = \frac{1}{n} T T'$$

where $n \times 2n$ matrix T is given by

$$T = \begin{pmatrix} 0 & \cdots & \cdots & 0 & Y_1 & Y_2 & \cdots & \cdots & Y_n \\ 0 & \cdots & \cdots & Y_1 & Y_2 & \cdots & \cdots & Y_n & 0 \\ 0 & \cdots & Y_1 & Y_2 & Y_3 & \cdots & Y_n & 0 & 0 \\ - & - & - & - & - & & & & \\ 0 & Y_1 & \cdots & \cdots & Y_n & 0 & \cdots & \cdots & 0 \end{pmatrix}$$

$Y_i = X_i - \bar{X}_n$; $i = 1(1)n$

Thus, $\forall\, \underset{\sim}{a} \in \mathbb{R}^n$ $\underset{\sim}{a}' \hat{\Gamma}_n \underset{\sim}{a} \geq 0$

$\Rightarrow \hat{\gamma}(h)\ f^n$ is n.n.d.

Note: $\hat{\gamma}^*(h)\ f^n$ is not n.n.d.

Standard models of time series

(I) White noise process

$$X_t = \epsilon_t$$

$\{\epsilon_t\}$: sequence of uncorrelated $(0,\sigma^2)$ random variables

$X_t \sim WN(0,\sigma^2)$

$EX_t = 0 \quad \forall t$

ACVF $\quad \gamma_X(h) = \begin{cases} \sigma^2, & \text{if } h=0 \\ 0, & \text{o/w} \end{cases}$

ACF $\quad \rho_X(h) = \begin{cases} 1, & \text{if } h=0 \\ 0, & \text{o/w} \end{cases}$

(II) Moving Average (MA) process

Suppose $\epsilon_t \sim WN(0,\sigma^2)$

$\{X_t\}$ is $MA(q)$ (q is a positive integer) if

$$X_t = \theta_0 \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}$$

$$\theta_0 \neq 0, \quad \theta_q \neq 0$$

$\theta_0, \theta_1, \ldots, \theta_q$ are unknown consts – MA parameters

Note: w.l.o.g. θ_0 can be taken as 1

o/w define $\epsilon_t' = \theta_0 \epsilon_t \sim WN(0, \theta_0^2 \sigma^2)$

& $X_t = \epsilon_t' + \left(\frac{\theta_1}{\theta_0}\right)\epsilon_{t-1}' + \cdots + \left(\frac{\theta_q}{\theta_0}\right)\epsilon_{t-q}'$

alternate $MA(q)$ representation

Note: 2-sided MA representation

$$X_t = \sum_{j=-M}^{M} \theta_j \epsilon_{t-j}$$

(iii) Auto Regressive (AR) process

Suppose $\epsilon_t \sim WN(0, \sigma^2)$

$\{X_t\}$ is $AR(p)$ if

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t$$

$$\phi_p \neq 0 \ ; \ Cov(\epsilon_t, X_{t-j}) = 0 \ \forall j > 0.$$

$\phi_1, \ldots, \phi_p$ are unknown consts; AR parameters.

Note: w.l.o.g. we can take a model without constant term for a stationary process

Suppose we take

$$X_t = \delta + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t$$

Since $\{X_t\}$ is stationary

$$\mu = EX_t = \delta + \phi_1 \mu + \cdots + \phi_p \mu$$

$$\Rightarrow \mu(1 - \phi_1 - \cdots - \phi_p) = \delta$$

If $(1 - \phi_1 - \cdots - \phi_p) = 0$ then $\delta = 0$

If o/w, i.e. $(1 - \phi_1 - \cdots - \phi_p) \neq 0$, then we can write

$$X_t = \delta + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t$$

as
$$X_t - \mu = \delta + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} - \mu + \epsilon_t$$

i.e. $X_t - \mu = \mu(1-\phi_1 - \cdots - \phi_p) + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} - \mu + \epsilon_t$

i.e. $(X_t - \mu) = \phi_1 (X_{t-1} - \mu) + \cdots + \phi_p (X_{t-p} - \mu) + \epsilon_t$

Define, $Y_t = X_t - \mu$

model is $\quad Y_t = \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \epsilon_t$

$\{X_t\}$ Covariance stationary $\Rightarrow$ $\{Y_t\}$ is also covariance stationary

with identical ACVF & ACF as $\{X_t\}$

as

$$X_t - \mu = \delta + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} - \mu + \epsilon_t$$

i.e. $X_t - \mu = \mu(1-\phi_1 - \cdots - \phi_p) + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} - \mu + \epsilon_t$

i.e. $(X_t - \mu) = \phi_1(X_{t-1} - \mu) + \cdots + \phi_p(X_{t-p} - \mu) + \epsilon_t$

Define, $Y_t = X_t - \mu$

model is $Y_t = \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \epsilon_t$

$\{X_t\}$ Covariance stationary $\Rightarrow$ $\{Y_t\}$ is also covariance stationary

with identical ACVF & ACF as $\{X_t\}$

<u>Note</u>: A special case AR process

<u>Random Walk</u>: $X_t = X_{t-1} + \epsilon_t$; ϵ_t i.i.d (μ, σ^2) (or $\epsilon_t \sim WN(0, \sigma^2)$)

Let $X_0 = 0$ initialization

$$X_1 = \epsilon_1$$

$$X_2 = X_1 + \epsilon_2 = \epsilon_1 + \epsilon_2$$

$$\vdots$$

$X_t = \sum_{i=1}^{t} \epsilon_i$; $E(X_t) = t\mu$ (or 0 if $\epsilon_t \sim WN(0, \sigma^2)$)

$V(X_t) = \underline{\underline{t\sigma^2}} \Rightarrow \{X_t\}$ is non-stationary

Note that for such a process

$Y_t = \nabla X_t = X_t - X_{t-1}$ is always stationary

IV : Auto Regressive Moving Average (ARMA) process

$\epsilon_t \sim WN(0, \sigma^2)$

$\{X_t\}$ is an ARMA(p, q) process if

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p}$$
$$+ \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}$$

$$\phi_p \neq 0, \; \theta_q \neq 0; \; Cov(\epsilon_t, X_{t-j}) = 0 \;\; \forall j > 0$$

$\phi_1, \ldots, \phi_p$: AR parameters of AR part of ARMA(p,q)

$\theta_1, \ldots, \theta_q$; MA parameters of MA part of ARMA(p,q)

Time domain properties of standard models

I : White noise $X_t \sim WN(0, \sigma^2)$

$$\gamma_X(h) = \begin{cases} \sigma^2, & h = 0 \\ 0, & o/w \end{cases} \qquad \rho_X(h) = \begin{cases} 1, & h = 0 \\ 0, & o/w \end{cases}$$

$\{X_t\}$ is always covariance stationary

II : MA models

MA(1) $\quad X_t = \epsilon_t + \theta \epsilon_{t-1}$; $\epsilon_t \sim WN(0, \sigma^2)$

$E X_t = 0 \quad \forall t$

$V X_t = (1 + \theta^2)\sigma^2 \quad \forall t$

$$\gamma_X(h) = Cov(X_{t+h}, X_t) = E(\epsilon_{t+h} + \theta \epsilon_{t+h-1})(\epsilon_t + \theta \epsilon_{t-1})$$

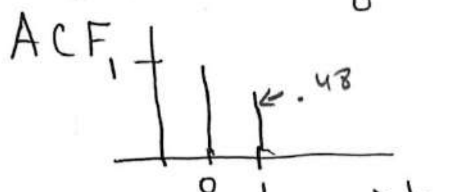
i.e. $$\gamma_X(h) = \begin{cases} (1+\theta^2)\sigma^2, & \text{If } h = 0 \\ \theta\sigma^2, & \text{If } h = \pm 1 \\ 0, & \text{o/w} \end{cases}$$

ACF $$\rho_X(h) = \begin{cases} 1, & h = 0 \\ \theta/1+\theta^2, & h = \pm 1 \\ 0, & \text{o/w} \end{cases}$$

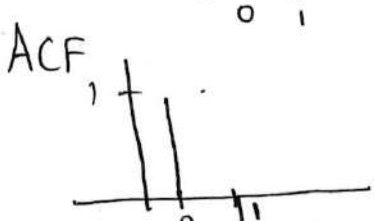
$\{X_t\}$ is covariance stationary $\forall \theta$

Note: Shape of ACF depends on the value of θ

$$X_t = \epsilon_t + 0.8\epsilon_{t-1}$$



$$X_t = \epsilon_t - 0.6\epsilon_{t-1}$$



Note: $\underset{\theta}{\text{Max}}\ \rho_X(1) = \frac{1}{2}$ attained at $\theta = 1$

$\underset{\theta}{\text{Min}}\ \rho_X(1) = -\frac{1}{2}$ attained at $\theta = -1$

Note: No unique representation

Note that

If $X_t = \epsilon_t + \theta\epsilon_{t-1}$ then $\rho_X(1) = \frac{\theta}{1+\theta^2}$

also If $X_t = \epsilon_t + \frac{1}{\theta}\epsilon_{t-1}$ then $\rho_X(1) = \frac{\theta}{1+\theta^2}$

$\Rightarrow \forall \rho \in [-\frac{1}{2}, \frac{1}{2}]$ $\exists$ 2 different MA(1) models that gives the same ACF at lag 1

Note: Lag operator representation

$$X_t = \epsilon_t + \theta \epsilon_{t-1} = \epsilon_t + \theta B \epsilon_t$$

i.e. $X_t = \theta(B)\epsilon_t$

$\theta(B) = 1 + \theta B \leftarrow$ MA polynomial

MA(2) process

$$X_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} \; ; \; \epsilon_t \sim WN(0, \sigma^2)$$

$$X_t = (1 + \theta_1 B + \theta_2 B^2)\epsilon_t$$

$$X_t = \theta(B)\epsilon_t$$

$$E X_t = 0 \;\; \forall t \; ; \; V(X_t) = \sigma^2(1 + \theta_1^2 + \theta_2^2) \;\; \forall t$$

$$\gamma_X(1) = Cov(X_{t+1}, X_t) = Cov(\epsilon_{t+1} + \theta_1 \epsilon_t + \theta_2 \epsilon_{t-1}, \; \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2})$$

$$= \sigma^2(\theta_1 + \theta_1\theta_2) = \gamma_X(-1)$$

$$\gamma_X(2) = Cov(X_{t+2}, X_t) = Cov(\epsilon_{t+2} + \theta_1 \epsilon_{t+1} + \theta_2 \epsilon_t, \; \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2})$$

$$= \sigma^2 \theta_2 = \gamma_X(-2)$$

$$\gamma_X(\pm 3) = \gamma_X(\pm 4) = \dots = 0$$

$$\rho_X(\pm 3) = \rho_X(\pm 4) = \dots = 0$$

$$\rho_X(1) = \frac{\theta_1 + \theta_1\theta_2}{1 + \theta_1^2 + \theta_2^2} = \rho_X(-1); \; \rho_X(\pm 2) = \frac{\theta_2}{1 + \theta_1^2 + \theta_2^2}$$

$\{X_t\}$ is covariance stationary $\forall \theta_1, \theta_2$

MA(q)

$$X_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}\ ; \quad \epsilon_t \sim WN(0,\sigma^2)$$

$$X_t = \theta(B)\,\epsilon_t$$

$$\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$$

$$E\,X_t = 0 \quad \forall\, t$$

$$V\,X_t = \sigma^2\left(1 + \sum_{j=1}^{q} \theta_j^2\right).$$

$$\mathrm{Cov}(X_{t+h}, X_t) = \mathrm{Cov}\big(\epsilon_{t+h} + \theta_1 \epsilon_{t+h-1} + \cdots + \theta_q \epsilon_{t+h-q},\ \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}\big)$$

$= 0 \quad \forall\, |h| > q$ (as there are no common terms)

$$\gamma_X(1) = \sigma^2(\theta_1\theta_0 + \theta_2\theta_1 + \cdots + \theta_q\theta_{q-1}) \qquad (\theta_0 = 1)$$

$$\gamma_X(2) = \sigma^2(\theta_2\theta_0 + \theta_3\theta_1 + \cdots + \theta_q\theta_{q-2})$$

$$\forall\, 0 \le h \le q\ ; \quad \gamma_X(h) = \sigma^2(\theta_h\theta_0 + \theta_{h+1}\theta_1 + \cdots + \theta_q\theta_{q-h})$$

i.e. $\gamma_X(h) = \sigma^2 \sum_{j=h}^{q} \theta_j \theta_{j-h}$ $\qquad j' = j-h$, $j = j'+h$

i.e. $\gamma_X(h) = \sigma^2 \sum_{j=0}^{q-h} \theta_j \theta_{j+h}$

Further $\gamma_X(-h) = \gamma_X(h)$

$$\Rightarrow \gamma_X(h) = \begin{cases} \sigma^2\left(1 + \sum_1^q \theta_j^2\right), & \text{if } h = 0 \\ \left(\sum_{j=0}^{q-|h|} \theta_j \theta_{j+|h|}\right)\sigma^2, & \text{if } |h| \le q \\ 0, & \text{o/w} \end{cases}$$

ACF

$$\rho_X(h) = \begin{cases} 1, & \text{if } h = 0 \\ \dfrac{\sum_{0}^{q-|h|} \theta_j \theta_{j+|h|}}{\left(1 + \sum_{1}^{q} \theta_j^2\right)}, & \text{if } |h| \le q \\ 0, & \text{o/w} \end{cases}$$

MA(q) is stationary $\forall\ \theta_1, \dots, \theta_q$

Remark: All finite order MA processes are always covariance stationary, irrespective of the values of MA parameters.

Remark: ACF value is always equal to zero beyond the lag which equals the order of the finite order MA model. This gives a way to identify MA processes from sample ACF plot.

MA(∞) process $\quad X_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$; $\epsilon_t \sim WN(0, \sigma^2)$

$EX_t = 0\ \forall t$; $VX_t = \sigma^2\left(\sum_j \psi_j^2\right)$

$< \infty$ if $\sum \psi_j^2 < \infty$

or if $\sum |\psi_j| < \infty$

MA(∞) is not always covariance stationary

For covariance stationary process

$$\gamma_h = \left(\sum_{j=0}^{\infty} \psi_j \psi_{j+h} \right) \sigma^2$$

Note: Although square summability, $\sum \psi_j^2 < \infty$ is enough to ensure that $MA(\infty)$ is covariance stationary, absolute summability $\sum |\psi_j| < \infty$ is assumed for $MA(\infty)$.

$\sum_j |\psi_j| < \infty$ ensures that $\sum_h |\gamma_h| < \infty$ which is required for further theoretical results for $MA(\infty)$ process.

iii) Auto Regressive (AR) processes

AR(1) or Markov process

$$X_t = \phi X_{t-1} + \epsilon_t \; ; \; \epsilon_t \sim WN(0, \sigma^2)$$

$$\mathrm{cov}(\epsilon_t, X_{t-j}) = 0 \quad \forall j > 0$$

unlike MA, finite order, processes, AR processes are not always stationary

Note that for AR(1) process with $\phi = 1$ is a random walk and as we have already discussed it is non-stationary

Further, for $|\phi|>1$ the series explodes and hence can't be stationary

[Note that
$$\begin{aligned} X_t &= \phi X_{t-1} + \epsilon_t \\ &= \phi(\phi X_{t-2} + \epsilon_{t-1}) + \epsilon_t \\ &= \phi^2 X_{t-2} + \phi\epsilon_{t-1} + \epsilon_t \\ &= \phi^2(\phi X_{t-3} + \epsilon_{t-2}) + \phi\epsilon_{t-1} + \epsilon_t \\ &= \phi^3 X_{t-3} + \phi^2\epsilon_{t-2} + \phi\epsilon_{t-1} + \epsilon_t \\ &\ \vdots \\ &= \phi^t X_0 + \phi^{t-1}\epsilon_1 + \cdots + \epsilon_t \end{aligned}$$

Starting from any arbitrary X_0, series would explode for $|\phi|>1$]

If $\{X_t\}$ is stationary, then

$$V(X_t) = V(\phi X_{t-1} + \epsilon_t) = \phi^2 V(X_{t-1}) + V(\epsilon_t)$$

i.e $\sigma_X^2 = \phi^2\sigma_X^2 + \sigma^2$

$$\Rightarrow \sigma_X^2 = \gamma_X(0) = \frac{\sigma^2}{1-\phi^2} \leftarrow |\phi|<1 \text{ is a valid region for this}$$

For an AR(1) process $|\phi|<1$ is the region for stationarity.

Alternate formulation for stationarity:

$$\phi(B)X_t = \epsilon_t$$
$$\phi(B) = 1-\phi B$$

Consider root of $\phi(z)=0$; i.e. $1-\phi z=0 \Rightarrow z=\frac{1}{\phi}$

$|\phi|<1 \Leftrightarrow$ roots of $\phi(z)=0$ lie outside unit circle

Note: Condition for stationarity of AR is usually in terms of the above

Consider a covariance stationary AR(1)

$$X_t = \phi X_{t-1} + \epsilon_t$$

$$E(X_t) = E(\phi X_{t-1} + \epsilon_t)$$

$$\mu_X = \phi \mu_X$$

i.e $\mu_X(1-\phi) = 0$

$1-\phi \neq 0$ (condition of stationarity)

$\Rightarrow \mu_X = 0$

$$\gamma_X(1) = Cov(X_{t+1}, X_t) = E(X_{t+1} X_t)$$
$$= E(\phi X_t + \epsilon_{t+1}) X_t$$
$$= \phi \sigma_X^2 + 0 \quad (Cov(\epsilon_t, X_{t-j}) = 0 \ \forall j > 0)$$

i.e. $\gamma_X(1) = \phi \dfrac{\sigma^2}{1-\phi^2} = \gamma_X(-1)$

$$\gamma_X(2) = Cov(X_{t+2}, X_t) = E(X_{t+2} X_t)$$
$$= E(\phi X_{t+1} + \epsilon_{t+2}) X_t$$
$$= \phi \gamma_X(1) = \phi^2 \frac{\sigma^2}{1-\phi^2}.$$

$\forall h > 0$

$$\gamma_X(h) = E(X_{t+h} X_t)$$

Now $X_{t+h} = \phi X_{t+h-1} + \epsilon_{t+h}$

$$= \phi(\phi X_{t+h-2} + \epsilon_{t+h-1}) + \epsilon_{t+h}$$
$$= \phi^2 X_{t+h-2} + \phi \epsilon_{t+h-1} + \epsilon_{t+h}$$
$$\vdots$$
$$= \phi^h X_{t+h-h} + \phi^{h-1} \epsilon_{t+1} + \cdots + \epsilon_{t+h}$$

$$\Rightarrow \gamma_X(h) = E(X_{t+h} X_t)$$
$$= \phi^h \gamma_X(0) = \phi^h \frac{\sigma^2}{1-\phi^2} \quad \left(\begin{array}{c} Cov(\epsilon_t, X_{t-j}) = 0 \\ \forall j > 0 \end{array} \right).$$

$$\gamma_X(h) = \gamma_X(-h) = \frac{\sigma^2}{1-\phi^2} \phi^{|h|} ; \quad h = 0, \pm 1, \pm 2, \ldots$$

ACF $\rho_X(h) = \begin{cases} 1, & h = 0 \\ \phi^{|h|} & \text{If } h \neq 0 \end{cases}$

Note: unlike MA(q), AR process's ACVF/ACF does not cut off (to zero) beyond the lag order of the model

Note: $\forall h > 0$

$$\gamma_X(h) = E(X_{t+h} X_t)$$
$$= E(\phi X_{t+h-1} + \epsilon_{t+h}) X_t$$
$$= \phi \gamma_X(h-1)$$

$$\text{i.e. } \gamma_X(h) = \phi \gamma_X(h-1) \quad - (*)$$

i.e. the ACVF satisfies relationship similar to the data eq^n

The above eq^n (*) is called the Yule-Walker eq^n.

AR(2) process or Yule process

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t ; \quad \epsilon_t \sim WN(0, \sigma^2)$$

$$Cov(\epsilon_t, X_{t-j}) = 0 \quad \forall j > 0$$

$$\phi(B) X_t = \epsilon_t$$

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2$$

AR(2) process is stationary if roots of $\phi(z) = 0$ all lie outside the unit circle

i.e. roots of $1 - \phi_1 z - \phi_2 z^2 = 0$ lie outside unit circle

i.e. roots of $y^2 - \phi_1 y - \phi_2 = 0$ lie inside unit circle

Let π_1 & π_2 be the roots of $y^2 - \phi_1 y - \phi_2 = 0$

$|\pi_i| < 1 ; i = 1, 2$, if

$$\left| \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{2} \right| < 1$$

The 2 roots are

$$\pi_1 = \frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2} \quad \& \quad \pi_2 = \frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2}$$

Case 1 : Roots are real

$$|\pi_i| < 1$$

$$\Leftrightarrow \quad \underbrace{-1 < \frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2}}_{(ii)} < \underbrace{\frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2} < 1}_{(i)}$$

$$(i) \Rightarrow \quad \sqrt{\phi_1^2 + 4\phi_2} < 2 - \phi_1$$

$$\phi_1^2 + 4\phi_2 < 4 + \phi_1^2 - 4\phi_1$$

$$\text{i.e. } 1 - \phi_1 - \phi_2 > 0 \quad \text{—— (1)}$$

(ii) $\Rightarrow \quad -2 < \phi_1 - \sqrt{\phi_1^2 + 4\phi_2}$

$$(2+\phi_1)^2 > \phi_1^2 + 4\phi_2$$

i.e. $1 + \phi_1 - \phi_2 > 0 \quad (2)$

Also $\pi_1 + \pi_2 = \phi_1$ & $-\pi_1 \pi_2 = \phi_2$

$|\pi_i| < 1 \quad \Rightarrow \quad |\phi_2| < 1$

<u>Case 2</u> : Roots are complex

$$\pi_1 = a + ib = \frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2}$$

$$\pi_2 = a - ib = \frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2}$$

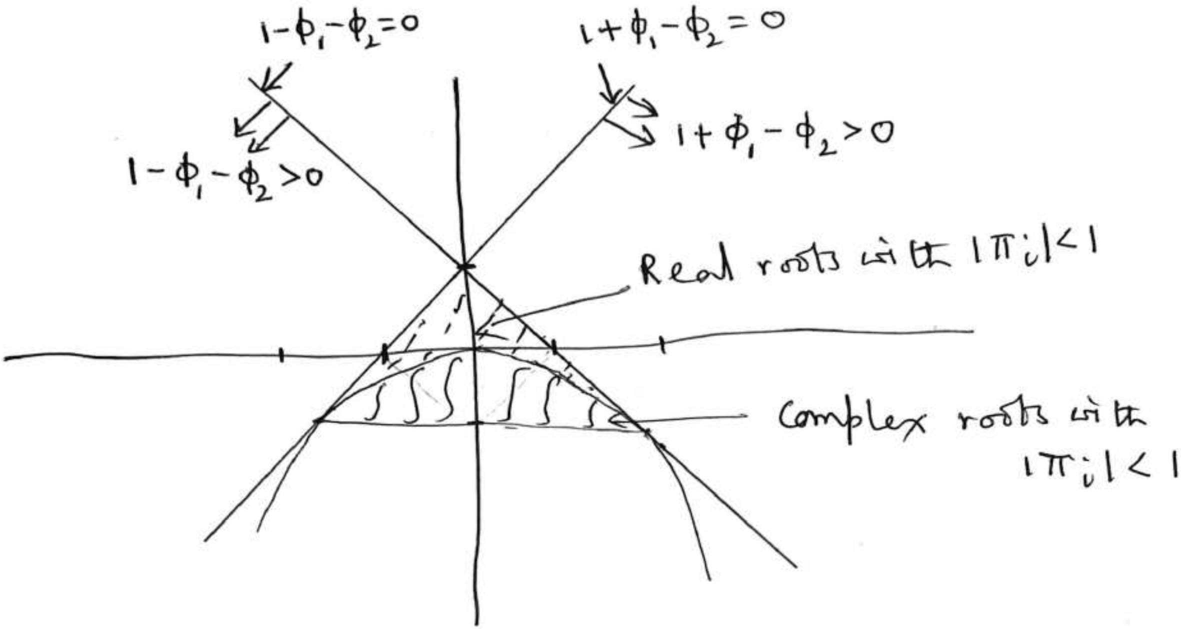
$$a = \phi_1/2 \qquad ib = \frac{\sqrt{\phi_1^2 + 4\phi_2}}{2}$$

$$\Rightarrow b = \frac{-\sqrt{-\phi_1^2 - 4\phi_2}}{2}$$

$$a^2 + b^2 = -\phi_2$$

$|\pi_i| < 1 \quad \Rightarrow \quad 1 + \phi_2 > 0.$

Region of stationarity in $\phi_1 - \phi_2$ plane



Region of stationarity is the triangular region bounded by

$$\left.\begin{array}{l} 1-\phi_1-\phi_2>0 \\ 1+\phi_1-\phi_2>0 \\ |\phi_2|<1 \end{array}\right\}.$$

Suppose $\{X_t\}$ is covariance stationary

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t$$

$$\mu_X = E(X_t) = E(\phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t) = \mu_X \phi_1 + \mu_X \phi_2$$

i.e. $\mu_X (1 - \phi_1 - \phi_2) = 0$

$1 - \phi_1 - \phi_2 \neq 0$ as $\{X_t\}$ is covariance stationary and hence root of $\phi(z) = 0$ can't be on unit circle

$\Rightarrow \mu_X = 0$.

$$\sigma_X^2 = V(X_t) = V(\phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t)$$

$$= \phi_1^2 \sigma_X^2 + \phi_2^2 \sigma_X^2 + 2\phi_1\phi_2 \gamma_1 + \sigma^2$$

$$\left(\mathrm{cov}(\epsilon_t, X_{t-1}) = 0 = \mathrm{Cov}(\epsilon_t, X_{t-2}) \right)$$

i.e. $\sigma_X^2 = \phi_1^2 \sigma_X^2 + \phi_2^2 \sigma_X^2 + 2\phi_1\phi_2 \rho_1 \sigma_X^2 + \sigma^2$

$$\Rightarrow \gamma_0 = \sigma_X^2 = \frac{\sigma^2}{1 - \phi_1^2 - \phi_2^2 - 2\phi_1\phi_2\rho_1}$$

Now $\gamma_1 = E(X_{t+1} X_t)$

$$= E(\phi_1 X_t + \phi_2 X_{t-1} + \epsilon_{t+1}) X_t$$

i.e. $\gamma_1 = \phi_1 \gamma_0 + \phi_2 \gamma_1$

i.e. $\rho_1 = \phi_1 + \phi_2 \rho_1$

i.e. $\rho_1 = \phi_1 / 1 - \phi_2$

$$\Rightarrow \gamma_0 = \sigma_X^2 = \frac{\sigma^2}{1 - \phi_1^2 - \phi_2^2 - 2\phi_1\phi_2\left(\frac{\phi_1}{1-\phi_2}\right)}$$

$$\gamma_0 = \frac{\sigma^2(1-\phi_2)}{(1-\phi_1-\phi_2)(1-\phi_2+\phi_1)(1+\phi_2)}$$

Note: factors in γ_0 are corresponding to the region of stationarity conditions

AR Yule-Walker eqⁿ and ACF for AR(2)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t$$

$$\gamma_X(h) = Cov(X_{t+h}, X_t)$$

$\forall h > 0$; $\gamma_X(h) = Cov(\phi_1 X_{t+h-1} + \phi_2 X_{t+h-2} + \epsilon_{t+h}, X_t)$

i.e. $\gamma_X(h) = \phi_1 \gamma_X(h-1) + \phi_2 \gamma_X(h-2) \quad \forall h > 0$

The ACVF/ACF satisfies the same 2^{nd} order diff eqⁿ as the data process.

Yule-Walker eqⁿ for ACF

$$\rho_X(h) = \phi_1 \rho_X(h-1) + \phi_2 \rho_X(h-2)$$

Yule-Walker eqⁿ can be used to express the ACF/ACVF seq in terms of ϕ_1 & ϕ_2

Take $h=1$ in Y-W; $\rho_1 = \phi_1 + \phi_2 \rho_1$

$h=2$ in Y-W; $\rho_2 = \phi_1 \rho_1 + \phi_2$

$$\Rightarrow \rho_1 = \frac{\phi_1}{1-\phi_2}$$

$$\rho_2 = \phi_1 \frac{\phi_1}{1-\phi_2} + \phi_2 = \frac{\phi_1^2 + \phi_2(1-\phi_2)}{1-\phi_2}$$

$$\rho_3 = \phi_1 \left(\frac{\phi_1^2 + \phi_2(1-\phi_2)}{1-\phi_2} \right) + \phi_2 \left(\frac{\phi_1}{1-\phi_2} \right)$$

Converversely, Y-W eqⁿ also enables us to express ϕs in terms of ρs.

$$\phi_1 = \frac{\rho_1(1-\rho_2)}{1-\rho_1^2} \quad \& \quad \phi_2 = \frac{\rho_2 - \rho_1^2}{1-\rho_1^2}$$

The above suggest a way for parameter estimation from sample ACF sequence.

AR(p) process

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \epsilon_t$$

$$\phi(B) X_t = \epsilon_t$$

$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$$

AR(p) is stationary if roots of $\phi(z) = 0$, i.e, roots of $1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p = 0$ all lie outside the unit circle.

i.e. roots of $y^p - \phi_1 y^{p-1} - \cdots - \phi_p = 0$ all lie inside the unit circle.

Suppose $\{X_t\}$ is covariance stationary AR(p)

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t$$

$$\mu_X = E(X_t) = E(\phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t)$$

$$= \phi_1 \mu_X + \cdots + \phi_p \mu_X$$

i.e. $\mu_X(1 - \phi_1 - \cdots - \phi_p) = 0$

$1 - \phi_1 - \cdots - \phi_p \neq 0$ (as $\{X_t\}$ is covariance stat)

$\Rightarrow \mu_X = 0$

ACVF structure of AR(p)

$$\gamma_X(h) = Cov(X_{t+h}, X_t)$$
$$= E(X_{t+h} X_t)$$
$$= E(\phi_1 X_{t+h-1} + \phi_2 X_{t+h-2} + \cdots + \phi_p X_{t+h-p} + \epsilon_{t+h}) X_t$$
$$= \begin{cases} \phi_1 \gamma_X(h-1) + \phi_2 \gamma_X(h-2) + \cdots + \phi_p \gamma_X(h-p) + \sigma^2, & h=0 \\ \phi_1 \gamma_X(h-1) + \phi_2 \gamma_X(h-2) + \cdots + \phi_p \gamma_X(h-p), & h=1,2,\ldots \end{cases}$$

& $\gamma_X(-h) = \gamma_X(h)$

$$\rho_h = \phi_1 \rho_{h-1} + \cdots + \phi_p \rho_{h-p} \quad ; \quad h=1,2,\ldots$$

$h=1$; $\quad \rho_1 = \phi_1 \rho_0 + \phi_2 \rho_1 + \cdots + \phi_p \rho_{p-1}$

$h=2$; $\quad \rho_2 = \phi_1 \rho_1 + \phi_2 \rho_0 + \cdots + \phi_p \rho_{p-2}$

$\vdots$

$h=p$; $\quad \rho_p = \phi_1 \rho_{p-1} + \cdots + \phi_p \rho_0$

i.e.
$$\underset{\sim}{\rho}_p = \begin{pmatrix} \rho_1 \\ \vdots \\ \rho_p \end{pmatrix} = \begin{pmatrix} 1 & \rho_1 & \cdots & \rho_{p-1} \\ & 1 & \cdots & \rho_{p-2} \\ & & \ddots & \\ & & & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \vdots \\ \phi_p \end{pmatrix}$$

i.e. $\underset{\sim}{\rho}_p = A_\rho \underset{\sim}{\phi} \Rightarrow \underset{\sim}{\phi} = A_\rho^{-1} \underset{\sim}{\rho}_p$

ARMA(p,q) process

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p}$$
$$+ \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}$$

$\phi_p \neq 0$, $\theta_q \neq 0$; $\epsilon_t \sim WN(0, \sigma^2)$; $Cov(\epsilon_t, X_{t-j}) = 0 \ \forall j > 0$

$$\phi(B) X_t = \theta(B) \epsilon_t$$
$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$$
$$\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$$

ARMA(p,q) process is stationary if AR part is stationary

i.e. ARMA(p,q) is stationary if the roots of $\phi(z) = 0$ all lie outside the unit circle

i.e. If the roots of $1 - \phi_1 z - \cdots - \phi_p z^p = 0$ all lie outside unit circle

i.e. If the roots of $y^p - \phi_1 y^{p-1} - \cdots - \phi_p = 0$ all lie inside the unit circle.

It is easy to see that

$E X_t = 0$ if $\{X_t\}$ is covariance stationary

$$\gamma_X(h) = Cov(X_{t+h}, X_t) = Cov(X_t, X_{t-h})$$
$$= E(X_t X_{t-h})$$
$$= E(\phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}) X_{t-h}$$

$$\gamma_X(h) = \phi_1 \gamma_X(h-1) + \cdots + \phi_p \gamma_X(h-p)$$
$$+ E(\epsilon_t X_{t-h}) + \theta_1 E(\epsilon_{t-1} X_{t-h}) + \cdots + E(\theta_q \epsilon_{t-q} X_{t-h}) \longleftarrow (*)$$

Note that

$$\gamma_X(h) = \phi_1 \gamma_X(h-1) + \cdots + \phi_p \gamma_X(h-p)$$
$$\forall h > q$$

otherwise, r.h.s. of (*) will have additional terms involving θ_js and ϕ_js & σ^2

e.g. suppose $q > 1$ and take $h = 1$

$$\gamma_X(1) = \phi_1 \gamma_X(0) + \phi_2 \gamma_X(1) + \cdots + \phi_p \gamma_X(p-1)$$
$$+ \sum_{i=1}^{q} \theta_i E(\epsilon_{t-i} X_{t-1})$$

$$E(\epsilon_{t-1} X_{t-1}) = E(\epsilon_{t-1}(\phi_1 X_{t-2} + \cdots + \phi_p X_{t-p-1} + \epsilon_{t-1}$$
$$+ \theta_1 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q-1}))$$
$$= \sigma^2 \quad \text{(always the case when time indices of } \epsilon \text{ \& } X \text{ are same)}$$

$$E(\epsilon_{t-2} X_{t-1}) = E\,\epsilon_{t-2}(\phi_1 X_{t-2} + \cdots + \phi_p X_{t-p-1} + \epsilon_{t-1}$$
$$+ \theta_1 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q-1})$$
$$= \sigma^2 \phi_1 + \sigma^2 \theta_1$$

Invertibility of stationary processes

Invertibility of AR(1) : AR(1) to MA(∞) representation

Let $\{X_t\}$ be a covariance stationary AR(1)

$$X_t = \phi X_{t-1} + \epsilon_t \; ; \quad \epsilon_t \sim WN(0, \sigma^2) ; |\phi| < 1$$

$$(1-\phi B) X_t = \epsilon_t \quad \text{i.e.} \quad \phi(B) X_t = \epsilon_t$$

Define $\phi_j^*(B) = 1 + \phi B + \phi^2 B^2 + \cdots + \phi^j B^j$

We have by multiplying both sides of AR(1) model equation with $\phi_j^*(B)$

$$\phi_j^*(B)\, \phi(B)\, X_t = \phi_j^*(B)\, \epsilon_t$$

Note that $\phi_j^*(B)\, \phi(B)$

$$= (1 + \phi B + \cdots + \phi^j B^j)(1 - \phi B)$$

$$= (1 + \phi B + \cdots + \phi^j B^j) - (\phi B + \phi^2 B^2 + \cdots + \phi^{j+1} B^{j+1})$$

$$= 1 - \phi^{j+1} B^{j+1}$$

$$\Rightarrow \phi_j^*(B)\, \phi(B)\, X_t = X_t - \phi^{j+1} B^{j+1} X_t$$

$$= X_t - \phi^{j+1} X_{t-j-1}$$

$$= \phi_j^*(B)\, \epsilon_t$$

$$\Rightarrow X_t - \phi^{j+1} X_{t-j-1} = \epsilon_t + \phi\epsilon_{t-1} + \cdots + \phi^j \epsilon_{t-j}$$

Realize that

$$\phi_j^*(B)\phi(B)X_t - X_t = -\phi^{j+1} X_{t-j-1}$$

hence

$$E\left(\phi_j^*(B)\phi(B)X_t - X_t\right)^2 = \phi^{2(j+1)} E\left(X_{t-j-1}^2\right)$$

$$\lim_{j\to\infty} E\left(\phi_j^*(B)\phi(B)X_t - X_t\right)^2 = \lim_{j\to\infty} \phi^{2(j+1)} E\left(X_{t-j-1}^2\right).$$

Since $V(X_t) = E(X_t^2) < \infty \ \forall t$ (as the process is covariance stationary)

& $|\phi| < 1$

$$\lim_{j\to\infty} \phi^{2(j+1)} E\left(X_{t-j-1}^2\right) = 0$$

$$\Rightarrow \lim_{j\to\infty} E\left(\phi_j^*(B)\phi(B)X_t - X_t\right)^2 = 0$$

i.e. $\lim_{j\to\infty} E\left(\phi_j^*(B)\epsilon_t - X_t\right)^2 = 0$

$\left(\phi_j^*(B)\phi(B)X_t = \phi_j^*(B)\epsilon_t\right)$

i.e. $\lim_{j\to\infty} E\left(\sum_{i=0}^{j} \phi^i \epsilon_{t-i} - X_t\right)^2 = 0$

i.e. $\sum_{i=0}^{j} \phi^i \epsilon_{t-i} \xrightarrow{m.s.} X_t$ as $j \to \infty$

($\xrightarrow{m.s.}$: means convergence in mean square sense)

$$\lim_{j \to \infty} \sum_{i=0}^{j} \phi^i \epsilon_{t-i} = \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}$$

Thus

$$X_t \overset{m.s.}{=\!=} \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i} \leftarrow MA(\infty)$$

i.e. $X_t = \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}$ (usually we just say this in m.s. sense).

with $\lim_{j \to \infty} (1+\phi B + \cdots + \phi^j B^j)$ acting as $(1-\phi B)^{-1}$

so that $(1-\phi B) X_t = \epsilon_t$

$$\Rightarrow (1-\phi B)^{-1}(1-\phi B) X_t = (1-\phi B)^{-1} \epsilon_t$$

i.e. $X_t = \sum_{i=0}^{\infty} \phi^i B^i \epsilon_t$

i.e. $X_t = \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}$

MA(∞) representation of stationary AR(1).

Note: $\forall |\phi| < 1$; we will take

$\sum_{i=0}^{\infty} \phi^i B^i = (1-\phi B)^{-1}$ the inverse operator of $(1-\phi B)$ operator

AR(2) to MA(∞) representation

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t$$

$$\phi(B) X_t = \epsilon_t$$

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 = (1-\lambda_1 B)(1-\lambda_2 B) \text{ say}$$

$\Rightarrow$ Roots of $\phi(z) = 0$ are $\frac{1}{\lambda_1}$ & $\frac{1}{\lambda_2}$

For $\{X_t\}$, a covariance stationary process,

$|\lambda_i| < 1 \quad i = 1, 2$

Further if $|\lambda_i| < 1 \ \forall i$ then $(1-\lambda_i B)^{-1}$

exists $\forall i$ and is given by

$$(1-\lambda_i B)^{-1} = \sum_{k=0}^{\infty} \lambda_i^k B^k$$

Let us consider a partial fraction approach to obtain MA(∞) representation of AR(2).

$$X_t = \frac{1}{(1-\lambda_1 B)(1-\lambda_2 B)} \epsilon_t$$

Let
$$\frac{1}{(1-\lambda_1 B)(1-\lambda_2 B)} = \frac{a}{1-\lambda_1 B} + \frac{b}{1-\lambda_2 B}$$

$$= \frac{a(1-\lambda_2 B) + b(1-\lambda_1 B)}{(1-\lambda_1 B)(1-\lambda_2 B)}$$

$$= \frac{(a+b) - B(a\lambda_2 + b\lambda_1)}{(1-\lambda_1 B)(1-\lambda_2 B)}$$

$\Rightarrow a + b = 1 \ \& \ a\lambda_2 + b\lambda_1 = 0$

$(1-b)\lambda_2 + b\lambda_1 = 0 \Rightarrow b = \frac{\lambda_2}{\lambda_2 - \lambda_1}$

& $a = \dfrac{\lambda_1}{\lambda_1 - \lambda_2}$

$$\Rightarrow \frac{1}{(1-\lambda_1 B)(1-\lambda_2 B)} = (1-\lambda_1 B)^{-1}(1-\lambda_2 B)^{-1}$$

$$= \frac{\lambda_1}{\lambda_1 - \lambda_2}(1-\lambda_1 B)^{-1} + \frac{\lambda_2}{\lambda_2 - \lambda_1}(1-\lambda_2 B)^{-1}$$

$$= \frac{\lambda_1}{\lambda_1 - \lambda_2}\sum_{j=0}^{\infty} \lambda_1^j B^j + \frac{\lambda_2}{\lambda_2 - \lambda_1}\sum_{j=0}^{\infty} \lambda_2^j B^j$$

$\Rightarrow (1-\lambda_1 B)(1-\lambda_2 B) X_t = \epsilon_t$ can be expressed as

$$X_t = (1-\lambda_1 B)^{-1}(1-\lambda_2 B)^{-1}\epsilon_t$$

i.e. $$X_t = \left(\frac{\lambda_1}{\lambda_1 - \lambda_2}\sum_{j=0}^{\infty} \lambda_1^j B^j + \frac{\lambda_2}{\lambda_2 - \lambda_1}\sum_{j=0}^{\infty} \lambda_2^j B^j\right)\epsilon_t$$

i.e. $$X_t = \frac{\lambda_1}{\lambda_1 - \lambda_2}\sum_{j=0}^{\infty} \lambda_1^j \epsilon_{t-j} + \frac{\lambda_2}{\lambda_2 - \lambda_1}\sum_{j=0}^{\infty} \lambda_2^j \epsilon_{t-j}$$

i.e. $$X_t = \sum_{j=0}^{\infty}\left(\frac{\lambda_1}{\lambda_1 - \lambda_2}\lambda_1^j + \frac{\lambda_2}{\lambda_2 - \lambda_1}\lambda_2^j\right)\epsilon_{t-j}$$

MA(∞) representation of AR(2)

<u>Note</u>: The MA(∞) representation of AR(2) can also be obtained using the "method of comparing coefficients".

Illustration of "method of comparing coefficients".

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t$$

$$\phi(B) X_t = \epsilon_t$$

$$X_t = \phi(B)^{-1} \epsilon_t \quad \text{(as } \{X_t\} \text{ is covariance stationary)}$$

$$X_t = \Psi(B) \epsilon_t \text{, say, with}$$

$$\Psi(B) = \Psi_0 + \Psi_1 B + \Psi_2 B^2 + \cdots.$$

i.e. $\phi(B)^{-1} = \Psi(B)$

i.e. $1 = \phi(B)\Psi(B)$

i.e. $1 = (1 - \phi_1 B - \phi_2 B^2)(\Psi_0 + \Psi_1 B + \Psi_2 B^2 + \cdots)$

i.e. $1 = \Psi_0 + B(\Psi_1 - \Psi_0 \phi_1) + B^2(\Psi_2 - \phi_1 \Psi_1 - \phi_2 \Psi_0) + \cdots$

Comparing coeffs of B^j from both sides we can solve for $\Psi_0, \Psi_1, \Psi_2, \ldots$ and hence the MA(∞) rep $X_t = \sum_{j=0}^{\infty} \Psi_j \epsilon_{t-j}$

e.g.

Comparing coeff

of B^0 : $\Psi_0 = 1$

of B^1 : $\Psi_1 = \Psi_0 \phi_1 = \phi_1$

B^2 : $\Psi_2 = \phi_1 \Psi_1 + \phi_2 \Psi_0 = \phi_1^2 + \phi_2$

$\vdots$

AR(p) to MA(∞)

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t$$

$$\phi(B) X_t = \epsilon_t$$

$$X_t = \phi(B)^{-1} \epsilon_t$$

Let $\theta_1, \ldots, \theta_p$ be the roots of $\phi(z) = 0$

For $\{X_t\}$ covariance stationary $|\theta_i| > 1 \quad \forall i = 1(1)p$

$$\phi(B) = (1 - \theta_1^{-1} B)(1 - \theta_2^{-1} B) \cdots (1 - \theta_p^{-1} B)$$

$$\text{or} = (1 - \lambda_1 B)(1 - \lambda_2 B) \cdots (1 - \lambda_p B)$$

$$|\lambda_i| < 1 \quad \forall i = 1(1)p.$$

$\Rightarrow (1 - \lambda_i B)^{-1}$ exists and

$$(1 - \lambda_i B)^{-1} = \sum_{j=0}^{\infty} \lambda_i^j B^j$$

Now, using partial fraction approach

$$\phi(B)^{-1} = \frac{1}{\phi(B)} = \frac{c_1}{1 - \lambda_1 B} + \cdots + \frac{c_p}{1 - \lambda_p B}$$ for some suitable consts $c_1, \ldots, c_p$

$$= c_1 \sum_{j=0}^{\infty} \lambda_1^j B^j + \cdots + c_p \sum_{j=0}^{\infty} \lambda_p^j B^j$$

$$= \sum_{i=1}^{p} c_i \sum_{j=0}^{\infty} \lambda_i^j B^j$$

$$\Rightarrow X_t = \phi(B)^{-1} \epsilon_t = \left(\sum_{i=1}^{p} c_i \sum_{j=0}^{\infty} \lambda_i^j B^j \right) \epsilon_t$$

$$= \sum_{j=0}^{\infty} \underbrace{\left(\sum_{i=1}^{p} c_i \lambda_i^j \right)}_{\psi_j} \epsilon_{t-j} = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} \leftarrow \text{MA}(\infty)$$

with $\psi_j = \sum_{i=1}^{p} c_i \lambda_i^j = \sum_{i=1}^{p} c_i \theta_i^{-j}$

Note : Method of comparing coefficients can be used to find the seq $\psi_0, \psi_1, \psi_2, \ldots$ and hence the $MA(\infty)$ representation

$$X_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$$

Remark

Causal AR process : An $AR(p)$ process $X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t$ is said to be causal if it can be expressed in terms of a white noise sequence in $MA(\infty)$ form $X_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$ for appropriate constants.

Thus an $AR(p)$ is causal if roots of $\phi(z) = 0$ all lie outside unit circle (i.e. if is covariance stationary).

Invertible representations are thus causal representations of stationary AR processes.

Invertibility of MA processes

MA(1) $\quad X_t = \epsilon_t + \theta \epsilon_{t-1}$; $\epsilon_t \sim WN(0, \sigma^2)$

always covariance stationary $\forall \theta$

$$X_t = \theta(B)\,\epsilon_t$$

$$\theta(B) = 1 + \theta B$$

Suppose now that $|\theta| < 1$, then $(1+\theta B)^{-1}$ exists and is given by

$$(1+\theta B)^{-1} = (1-(-\theta)B)^{-1}$$
$$= 1 - \theta B + \theta^2 B^2 - \theta^3 B^3 + \cdots$$

$\Rightarrow \quad X_t = \theta(B)\,\epsilon_t$ can be expressed as

$$\epsilon_t = \theta(B)^{-1} X_t$$

i.e. $\quad \epsilon_t = X_t - \theta X_{t-1} + \theta^2 X_{t-2} - \cdots$

i.e. $\quad \epsilon_t = \sum_{i=0}^{\infty} (-\theta)^i X_{t-i} \quad \leftarrow$ AR(∞) form

i.e. $\quad X_t = -\sum_{i=1}^{\infty} (-\theta)^i X_{t-i} + \epsilon_t$

Note that the mean square sense convergence interpretation is also valid for this MA setup.

$$X_t = \epsilon_t + \theta \epsilon_{t-1}$$
$$= \epsilon_t + \theta(X_{t-1} - \theta \epsilon_{t-2})$$
$$= \epsilon_t + \theta X_{t-1} - \theta^2 \epsilon_{t-2}$$

i.e. $X_t = \theta X_{t-1} - \theta^2 \epsilon_{t-2} + \epsilon_t$

$$= \theta X_{t-1} - \theta^2 (X_{t-2} - \theta \epsilon_{t-3}) + \epsilon_t$$

i.e. $X_t = \theta X_{t-1} - \theta^2 X_{t-2} + \theta^3 \epsilon_{t-3} + \epsilon_t$

$\vdots$ continuing K times substitution

$$X_t = -\sum_{i=1}^{K} (-\theta)^i X_{t-i} - (-\theta)^{K+1} \epsilon_{t-(K+1)} + \epsilon_t$$

$$\Rightarrow E\left(X_t - \epsilon_t + \sum_{i=1}^{K} (-\theta)^i X_{t-i}\right)^2 = E\left((-\theta)^{K+1} \epsilon_{t-K-1}\right)^2$$

Since $E(\epsilon_t^2) < \infty \ \forall t$ & $|\theta| < 1$

$$\lim_{K \to \infty} E\left(X_t - \epsilon_t + \sum_{i=1}^{K} (-\theta)^i X_{t-i}\right)^2 = 0$$

$$\Rightarrow \quad X_t \overset{m.s.}{=\!=} \epsilon_t - \sum_{i=1}^{\infty} (-\theta)^i X_{t-i}$$

<u>Note</u>: (unlike AR(1)) Covariance stationary MA(1) is not necessarily invertible.

MA(q)

$$X_t = (1+\theta_1 B + \cdots + \theta_q B^q)\,\epsilon_t$$

$$X_t = \theta(B)\,\epsilon_t$$

Let $\theta(B) = (1-\lambda_1 B)\cdots(1-\lambda_q B)$

If $|\lambda_i| < 1 \ \forall i$ then roots of $\theta(z)=0$ all lie outside the unit circle and each of $(1-\lambda_i B)$ is invertible and $\theta(B)^{-1}$ exists and

$$\theta(B)^{-1} = (1-\lambda_1 B)^{-1}\cdots(1-\lambda_q B)^{-1}$$

$$\text{with } (1-\lambda_i B)^{-1} = \sum_{j=0}^{\infty} \lambda_i^j B^j \quad \forall\, i = 1(1)q$$

We can either use partial fraction approach or method of comparing coefficients to find the AR(∞) representation of the invertible MA(q).

e.g.

$$X_t = \theta(B)\,\epsilon_t$$

$$\Rightarrow \theta(B)^{-1} X_t = \epsilon_t$$

i.e. $\epsilon_t = \Psi(B) X_t$ say

$$\Psi(B) = \Psi_0 + \Psi_1 B + \cdots$$

$$\Rightarrow \Psi(B) = \theta(B)^{-1}$$

$$\text{or } \theta(B)\Psi(B) = 1$$

$$\text{i.e. } (1+\theta_1 B + \cdots + \theta_q B^q)(\Psi_0 + \Psi_1 B + \cdots) = 1$$

Comparing coeffs of B^j from both the sides we can express Ψ_j s in terms of θ_j s.

ARMA(p, q) $\{X_t\}$ is $\ni$

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}$$

$$\phi(B) X_t = \theta(B) \epsilon_t \qquad \epsilon_t \sim WN(0, \sigma^2)$$

$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$$

$$\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$$

(i) $\{X_t\}$ is said to be Causal if roots of $\phi(z) = 0$ all lie outside the unit circle. In such a case the MA(∞) causal representation is

$$X_t = \phi(B)^{-1} \theta(B) \epsilon_t = \psi(B) \epsilon_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$$

$$\Rightarrow \phi(B)^{-1} \theta(B) = \psi(B)$$

i.e. $\theta(B) = \phi(B) \psi(B)$

i.e. $(1 + \theta_1 B + \cdots + \theta_q B^q)$

$$= (1 - \phi_1 B - \cdots - \phi_p B^p)(\psi_0 + \psi_1 B + \cdots + \cdots)$$

Comparing coeffs of B^j, we get

$$\psi_s = \phi_1 \psi_{s-1} + \cdots + \phi_p \psi_{s-p} + \theta_s \quad \forall\, s \leq q$$

$$\& = \phi_1 \psi_{s-1} + \cdots + \phi_p \psi_{s-p} \quad \forall\, s > q$$

with $\psi_r = 0 \ \forall\, r < 0$ & $\psi_0 = 1$

(ii) $\{X_t\}$ is said to be invertible if it can be expressed as AR(∞), i.e. if roots of $\theta(z)=0$ all lie outside the unit circle. In such a case

$$\phi(B)X_t = \theta(B)\epsilon_t$$

$$\epsilon_t = \theta(B)^{-1}\phi(B)X_t$$

$$\underbrace{\epsilon_t = \psi(B)X_t = \sum_{j=0}^{\infty}\psi_j X_{t-j}}_{AR(\infty)}$$

We can use method of comparing coefficients to express ψs interms of θ_js & ϕ_js as done for the causal representation.

$$\theta(B)^{-1}\phi(B) = \psi(B)$$

$$\text{i.e. } \phi(B) = \theta(B)\psi(B)$$

Auto Covariance Generating Function (AGGF)

ACGF is a simple concept and usually easy to calculate. The auto covariances at different lags be determined through ACGF

If $\{X_t\}$ is a covariance stationary time series with ACVF $\gamma(\cdot)$, then it's ACGF is defined by

$$g_X(z) = \sum_{j=-\infty}^{\infty} \gamma(j) z^j \quad -(*)$$

provided the series converges for all z in some annulus $r^{-1} < |z| < r$ with $r > 1$.

Note: Coeff of z^j in $(*)$ is $\gamma(j)$, auto covariance at lag j.

MA(1) ACGF

$$X_t = \epsilon_t + \theta\epsilon_{t-1}; \quad \epsilon_t \sim WN(0,\sigma^2)$$

$$X_t = \theta(B)\epsilon_t$$

$$\gamma(j) = \begin{cases} \sigma^2(1+\theta^2), & j=0 \\ \theta\sigma^2, & j=\pm 1 \\ 0, & o/w \end{cases}$$

ACGF: $g_X(z) = (\theta\sigma^2)z^{-1} + \sigma^2(1+\theta^2)z^0 + (\theta\sigma^2)z^1$

i.e. $g_X(z) = \sigma^2\left(\theta z^{-1} + (1+\theta^2) + \theta z\right)$

i.e $g_X(z) = \sigma^2 (1+\theta z)(1+\theta z^{-1})$

i.e. $g_X(z) = \sigma^2 \theta(z)\, \theta(z^{-1})$

MA(q) ACGF

$$X_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}$$

$$X(t) = \theta(B)\epsilon_t \qquad \epsilon_t \sim WN(0,\sigma^2)$$

$$\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q.$$

ACGF:

$$\begin{aligned} g_X(z) = \sigma^2 \Big[& (\theta_0^2 + \theta_1^2 + \cdots + \theta_q^2) z^0 \\ & + (\theta_0\theta_1 + \cdots + \theta_{q-1}\theta_q) z^1 \\ & + (\theta_0\theta_1 + \cdots + \theta_{q-1}\theta_q) z^{-1} \\ & + (\theta_0\theta_2 + \cdots + \theta_{q-2}\theta_q) z^2 \\ & + (\theta_0\theta_2 + \cdots + \theta_{q-2}\theta_q) z^{-2} + \\ & \vdots \\ & + \theta_0\theta_q z^q + \theta_0\theta_q z^{-q} \Big] \end{aligned}$$

$\theta_0 = 1$

(using the already known auto covariance sequence)

i.e $g_X(z) = \sigma^2 (\theta_0 + \theta_1 z + \cdots + \theta_q z^q)$
$(\theta_0 + \theta_1 z^{-1} + \cdots + \theta_q z^{-q})$

$$g_X(z) = \sigma^2\, \theta(z)\, \theta(z^{-1})$$

MA(∞) : $X_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$ with $\epsilon_t \sim WN(0,\sigma^2)$ & $\sum_j |\psi_j| < \infty$

$X_t = \psi(B)\,\epsilon_t$

ACGF $g_X(z) = \sigma^2 \psi(z)\psi(z^{-1})$.

AR(1) $\{X_t\}$ is stationary AR(1)

$(1-\phi B)X_t = \epsilon_t$

$\theta_1+\theta_1\theta_2$

$\phi(B)X_t = \epsilon_t$

$X_t = \phi(B)^{-1}\epsilon_t$ ($\{X_t\}$ is causal)

$= \psi(B)\,\epsilon_t$ $\quad |\phi z|<1 \quad |z| < \frac{1}{|\phi|} \quad \left|\frac{\phi}{z}\right| < 1 \quad |\phi| < |z|$

ACGF $g_X(z) = \sigma^2 \psi(z)\psi(z^{-1})$

i.e. $g_X(z) = \dfrac{\sigma^2}{\phi(z)\,\phi(z^{-1})} = \dfrac{\sigma^2}{(1-\phi z)(1-\phi z^{-1})}$

i.e. $g_X(z) = \sigma^2(1+\phi z+\phi^2 z^2+\cdots)$
$(1+\phi z^{-1}+\phi^2 z^{-2}+\cdots)$ — (*)

Note that coeff of z^j in the r.h.s. of (*) is

$= \sigma^2(\phi^j + \phi^{j+1}\phi + \phi^{j+2}\phi^2 + \cdots)$

$= \dfrac{\phi^j\sigma^2}{1-\phi^2} = \gamma(j)$ as we expected from ACGF.

Note: for AR(p) $g_X(z) = \dfrac{\sigma^2}{\phi(z)\,\phi(z^{-1})}$; $\phi(B)X_t = \epsilon_t$

ARMA (p, q)

Suppose $\{X_t\}$ is covariance stationary ARMA(p, q)

$$\phi(B) X_t = \theta(B) \epsilon_t$$

Using the same causal representation (as used for AR, ACGF of ARMA(p, q) is

$$g_X(z) = \sigma^2 \frac{\theta(z)\,\theta(z^{-1})}{\phi(z)\,\phi(z^{-1})}.$$

Note : ACGF of WN is a constant.

ACGF of a filtered process

Let $\{X_t\}$ be a covariance stationary process with ACVF $\gamma(\cdot)$ and ACGF

$$g_X(z) = \sum_{j=-\infty}^{\infty} z^j \gamma(j)$$

Consider a linear filtered process $\{Y_t\}$ ∋

$$Y_t = \sum_{i=0}^{q} \theta_i X_{t-i} = \theta(B) X_t$$

$$\theta(B) = \theta_0 + \theta_1 B + \cdots + \theta_q B^q$$

$$\gamma_Y(h) = Cov(Y_{t+h}, Y_t)$$

$$= Cov\left(\sum_{i=0}^{q} \theta_i X_{t+h-i}, \sum_{j=0}^{q} \theta_j X_{t-j}\right)$$

ACGF of a filtered process

Let $\{X_t\}$ be a covariance stationary process with ACVF $\gamma(\cdot)$ and ACGF

$$g_X(z) = \sum_{j=-\infty}^{\infty} z^j \gamma(j)$$

Consider a linear filtered process $\{Y_t\}$ $\ni$

$$Y_t = \sum_{i=0}^{q} \theta_i X_{t-i} = \theta(B) X_t$$

$$\theta(B) = \theta_0 + \theta_1 B + \cdots + \theta_q B^q$$

$$\gamma_Y(h) = Cov(Y_{t+h}, Y_t)$$

$$= Cov\left(\sum_{i=0}^{q} \theta_i X_{t+h-i}, \sum_{j=0}^{q} \theta_j X_{t-j}\right)$$

$$\gamma_y(h) = \sum_{i=0}^{q} \sum_{j=0}^{q} \theta_i \theta_j \gamma_X(h-i+j)$$

$$g_y(z) = \sum_{h=-\infty}^{\infty} z^h \gamma_y(h)$$

$$= \sum_{h=-\infty}^{\infty} z^h \left(\sum_{i=0}^{q} \sum_{j=0}^{q} \theta_i \theta_j \gamma_X(h-i+j) \right)$$

$$= \sum_{i=0}^{q} \theta_i \sum_{j=0}^{q} \theta_j \sum_{h=-\infty}^{\infty} z^h \gamma_X(h-i+j)$$

$$= \sum_{i=0}^{q} \theta_i z^i \sum_{j=0}^{q} \theta_j z^{-j} \sum_{h=-\infty}^{\infty} z^{h-i+j} \gamma_X(h-i+j)$$

$$h' = h-i+j$$

$$= \sum_{i=0}^{q} \theta_i z^i \sum_{j=0}^{q} \theta_j z^{-j} \sum_{h'=-\infty}^{\infty} z^{h'} \gamma_X(h')$$

i.e. $\underline{g_y(z)} = \theta(z)\, \theta(z^{-1})\, g_X(z)$

ACGF of the filtered process

Multivariate time series processes

In many situations, it is desirable to study behaviour of related time series processes under a multivariate setup rather than studying these processes in isolation as univariate processes. under the multivariate setup it is possible to exploit the interrelationship between these processes for identifying the dynamics of these processes.

Let $\{X_{1_t}\}, \{X_{2_t}\}, \ldots \{X_{m_t}\}$ be m time series processes $\ni E X_{i_t}^2 < \infty \ \forall i, \forall t$ and define m-dimensional time series

$$\underset{\sim}{X}_t = (X_{1_t}, \ldots, X_{m_t})'$$

$$E(\underset{\sim}{X}_t) = \begin{pmatrix} E X_{1_t} \\ \vdots \\ E X_{m_t} \end{pmatrix} = \underset{\sim}{\mu}_t$$

$$E(\underset{\sim}{X}_t \underset{\sim}{X}'_{t+h}) = \begin{pmatrix} E(X_{1_t} X_{1_{t+h}}) & \cdots & E(X_{1_t} X_{m_{t+h}}) \\ & E(X_{2_t} X_{2_{t+h}}) \cdots & E(X_{2_t} X_{m_{t+h}}) \\ & \ddots & \vdots \\ & & E(X_{m_t} X_{m_{t+h}}) \end{pmatrix}$$

$$\text{Cov}(\underset{\sim}{X}_t, \underset{\sim}{X}_{t+h}) = E(\underset{\sim}{X}_t - \underset{\sim}{\mu}_t)(\underset{\sim}{X}_{t+h} - \underset{\sim}{\mu}_{t+h})'$$

$$= E(\underset{\sim}{X}_t \underset{\sim}{X}'_{t+h}) - \underset{\sim}{\mu}_t \underset{\sim}{\mu}'_{t+h}$$

$$= \begin{pmatrix} \gamma_{11}(t,t+h) & \cdots & \gamma_{1m}(t,t+h) \\ \cdots & \cdots & \cdots \\ \gamma_{m1}(t,t+h) & \cdots & \gamma_{mm}(t,t+h) \end{pmatrix}$$

$$= \Gamma_X(t, t+h) = \Gamma_t(h)$$

where $\gamma_{ij}(t,t+h) = \text{Cov}(X_{i_t}, X_{j_{t+h}})$

<u>Def^n</u> : An m-variate process $\{\underset{\sim}{X}_t\}$ is covariance stationary If

(i) $\underset{\sim}{\mu}_t$ is independent of t

(ii) $\Gamma_X(t, t+h)$ is independent of t and is a f^n of h only

In such a situation

$$\underset{\sim}{\mu} = E\,\underset{\sim}{X}_t \quad \forall t$$

$$E(\underset{\sim}{X}_t - \underset{\sim}{\mu})(\underset{\sim}{X}_{t+h} - \underset{\sim}{\mu})' = \Gamma_X(h)$$

$$= \begin{pmatrix} \gamma_{11}(h) & \cdots & \gamma_{1m}(h) \\ & \ddots & \\ & & \gamma_{mm}(h) \end{pmatrix}$$

$\Gamma_X(h)$ is called the Auto covariance matrix function

Remark: If a vector process is stationary then all the components are stationary. Converse is not true, i.e. If the component processes are stationary, the vector process can be non stationary. Can you give a counter example to show that converse is not true?

Basic properties of $\Gamma_X(\cdot)$ for a stationary process

Let $\{\underset{\sim}{X}_t\}$ be a covariance stationary m-variate process. The autocovariance matrix f^n of $\underset{\sim}{X}_t$ is defined as

$$\Gamma_X(h) = Cov(\underset{\sim}{X}_t, \underset{\sim}{X}_{t+h}) = E(\underset{\sim}{X}_t - \underset{\sim}{\mu})(\underset{\sim}{X}_{t+h} - \underset{\sim}{\mu})'$$

$$h = 0, \pm 1, \pm 2, \ldots$$

$$E(\underset{\sim}{X}_t) = \underset{\sim}{\mu} \quad \forall t$$

Properties

(i) $\Gamma_X(h) = \Gamma_X'(-h)$

$$\Gamma_X(h) = E(\underset{\sim}{X}_t - \underset{\sim}{\mu})(\underset{\sim}{X}_{t+h} - \underset{\sim}{\mu})'$$

$$= E(\underset{\sim}{X}_{t-h} - \underset{\sim}{\mu})(\underset{\sim}{X}_t - \underset{\sim}{\mu})' \quad \left(\because \underset{\sim}{X}_t \text{ is covariance stationary}\right)$$

$$= \left(E(\underset{\sim}{X}_t - \underset{\sim}{\mu})(\underset{\sim}{X}_{t-h} - \underset{\sim}{\mu})'\right)' = \left(\Gamma_X(-h)\right)'$$

(ii) $|\gamma_{ij}(h)| \leq [\gamma_{ii}(0)\, \gamma_{jj}(0)]^{1/2}$; $i,j = 1(1)m$

This follows from Cauchy-Schwarz inequality

$\gamma_{ij} = Cov(X_{i_t}, X_{j_{t+h}})$

(iii) $\Gamma_X(h) = ((\gamma_{ij}(h)))$

$\{\gamma_{ii}(h)\}$ is the ACVF seq of $\{X_{i_t}\}$
$i = 1(1)m$

(iv) $\forall\, \underset{\sim}{a}_j \in \mathbb{R}^m$; $j = 1(1)m$, $\sum_{k,j=1}^{n} \underset{\sim}{a}_j' \Gamma_X(k-j)\, \underset{\sim}{a}_k \geq 0$

Note that $\forall\, \underset{\sim}{a}_j \in \mathbb{R}^m$ $j = 1(1)m$, the random variable

$$Y = (\underset{\sim}{a}_1', \ldots, \underset{\sim}{a}_n') \begin{pmatrix} (\underset{\sim}{X}_1 - \underset{\sim}{\mu}) \\ \vdots \\ (\underset{\sim}{X}_n - \underset{\sim}{\mu}) \end{pmatrix}$$

$$E\, Y^2 = E\left(\sum_{j=1}^{n} \underset{\sim}{a}_j' (\underset{\sim}{X}_j - \underset{\sim}{\mu})\right)^2 \geq 0$$

i.e. $E\left(\sum_{j=1}^{n} \underset{\sim}{a}_j' (\underset{\sim}{X}_j - \underset{\sim}{\mu})\right)\left(\sum_{j=1}^{n} \underset{\sim}{a}_j' (\underset{\sim}{X}_j - \underset{\sim}{\mu})\right)' \geq 0$

i.e. $E\left(\sum_j \sum_k \underset{\sim}{a}_j' (\underset{\sim}{X}_j - \underset{\sim}{\mu})(\underset{\sim}{X}_k - \underset{\sim}{\mu})' \underset{\sim}{a}_k\right) \geq 0$

i.e. $\sum_{j,k} \underset{\sim}{a}_j' \left(E(\underset{\sim}{X}_j - \underset{\sim}{\mu})(\underset{\sim}{X}_k - \underset{\sim}{\mu})'\right) \underset{\sim}{a}_k \geq 0$

i.e. $\sum_{j,k} \underset{\sim}{a}_j' \Gamma_X(k-j)\, \underset{\sim}{a}_k \geq 0$

Note: We can also define a autocorrelation matrix function for a covariance stationary process as

$$R_X(h) = \begin{pmatrix} \rho_{11}(h) & \rho_{12}(h) & \cdots & \rho_{1m}(h) \\ & \ddots & & \\ & & \ddots & \\ & & & \rho_{mm}(h) \end{pmatrix}$$

$$\rho_{ij}(h) = \frac{\gamma_{ij}(h)}{\left(\gamma_{ii}(0)\,\gamma_{jj}(0)\right)^{1/2}}$$

$$R_X(h) = D_0^{-1/2}\,\Gamma_X(h)\,D_0^{-1/2}$$

$$D_0 = \mathrm{diag}\left(\gamma_{11}(0), \ldots, \gamma_{mm}(0)\right)$$

Standard multivariate processes

(I) Vector White noise (VWN)

$$\underset{\sim}{Z}_t \sim VWN(\underset{\sim}{0}, \Sigma)$$

($\underset{\sim}{Z}_t$ is a VWN with mean vector $\underset{\sim}{0}$ and covariance matrix f^n Σ)

$E\,\underset{\sim}{Z}_t = \underset{\sim}{0}$ and covariance matrix f^n as

$$\Gamma(h) = \begin{cases} \Sigma, & h=0 \\ 0, & o/w \end{cases} \quad \text{i.e. } E(\underset{\sim}{Z}_t \underset{\sim}{Z}_s') = \begin{cases} \Sigma, & t=s \\ 0, & t\neq s \end{cases}$$

Note : Note that for VWN process the vector process is uncorrelated; however, It is not necessary that the components are uncorrelated.

II : Vector Moving Average (VMA)

VMA(1): $\underset{K\times 1}{\underset{\sim}{X}_t} = \underset{K\times K}{\Theta}\, \underset{\sim}{\epsilon}_{t-1} + \underset{\sim}{\epsilon}_t$

$\underset{\sim}{\epsilon}_t \sim VWN(\underset{\sim}{0}, \Sigma)$

Θ : $K\times K$ matrix of constants

$$\underset{\sim}{X}_t = (I_K + \Theta B)\, \underset{\sim}{\epsilon}_t$$

i.e. $\underset{\sim}{X}_t = \Theta(B)\, \underset{\sim}{\epsilon}_t$

$\Theta(B) = I_K + \Theta B$: MA matrix polynomial

$E\, \underset{\sim}{X}_t = \underset{\sim}{0} \quad \forall\, t$

$$\begin{aligned}
Cov(\underset{\sim}{X}_t, \underset{\sim}{X}_{t+h}) &= E(\underset{\sim}{X}_t \underset{\sim}{X}_{t+h}') \\
&= E\left(\underset{\sim}{\epsilon}_t + \Theta\, \underset{\sim}{\epsilon}_{t-1}\right)\left(\underset{\sim}{\epsilon}_{t+h} + \Theta\, \underset{\sim}{\epsilon}_{t+h-1}\right)' \\
&= E\left(\underset{\sim}{\epsilon}_t \underset{\sim}{\epsilon}_{t+h}'\right) + E\left(\underset{\sim}{\epsilon}_t \underset{\sim}{\epsilon}_{t+h-1}' \Theta'\right) \\
&\quad + E\left(\Theta\, \underset{\sim}{\epsilon}_{t-1} \underset{\sim}{\epsilon}_{t+h}'\right) \\
&\quad + E\left(\Theta\, \underset{\sim}{\epsilon}_{t-1} \underset{\sim}{\epsilon}_{t+h-1}' \Theta'\right)
\end{aligned}$$

i.e. $$\Gamma_X(h) = \Sigma\, I_0(h) + \Sigma\, \Theta' I_1(h) + \Theta \Sigma\, I_{-1}(h) + \Theta \Sigma \Theta' I_0(h)$$

i.e. $\Gamma_X(h) = \begin{cases} \Sigma + \Theta\Sigma\Theta', & h = 0 \\ \Sigma\Theta', & h = 1 \\ \Theta\Sigma, & h = -1 \\ 0, & |h| > 1 \end{cases}$

$\underset{\sim}{X}_t$ is always covariance stationary $\forall$ $K \times K$ Θ.

VMA(q)

$$\underset{\sim}{X}_t = \underset{\sim}{\epsilon}_t + \Theta_1 \underset{\sim}{\epsilon}_{t-1} + \cdots + \Theta_q \underset{\sim}{\epsilon}_{t-q}$$

$$\Theta_q \neq 0 \; ; \; \underset{\sim}{\epsilon}_t \sim VWN(\underset{\sim}{0}, \Sigma)$$

$$\underset{\sim}{X}_t = (I_K + \Theta_1 B + \cdots + \Theta_q B^q) \underset{\sim}{\epsilon}_t$$

i.e. $\underset{\sim}{X}_t = \Theta(B) \underset{\sim}{\epsilon}_t$

$\Theta(B) = I_K + \Theta_1 B + \cdots + \Theta_q B^q$ is the MA matrix polynomial

$$E \underset{\sim}{X}_t = \underset{\sim}{0}$$

$Cov(\underset{\sim}{X}_t, \underset{\sim}{X}_{t+h}) = E(\underset{\sim}{X}_t \underset{\sim}{X}'_{t+h})$

$(\Theta_0 = I_K) \longrightarrow \; = E(\Theta_0 \underset{\sim}{\epsilon}_t + \Theta_1 \underset{\sim}{\epsilon}_{t-1} + \cdots + \Theta_q \underset{\sim}{\epsilon}_{t-q})(\Theta_0 \underset{\sim}{\epsilon}_{t+h} + \Theta_1 \epsilon_{t+h-1} + \cdots + \Theta_q \epsilon_{t+h-q})'$

$$\Gamma_0 = Cov(\underset{\sim}{X}_t, \underset{\sim}{X}_t) = \sum_{j=0}^{q} \Theta_j \Sigma \Theta_j'$$

$$\Gamma(1) = Cov(\underset{\sim}{X}_t, \underset{\sim}{X}_{t+1}) = \sum_{j=0}^{q-1} \Theta_j \Sigma \Theta_{j+1}'$$

$$\Gamma(-1) = \Gamma(1)'$$

$$\Gamma(2) = Cov(\underset{\sim}{X}_t, \underset{\sim}{X}_{t+2}) = \sum_{j=0}^{q-2} \Theta_j \Sigma \Theta_{j+2}'$$

$$\Gamma(-2) = \Gamma(2)'$$

$\forall\ h \le q$

$$\Gamma(h) = \sum_{j=0}^{q-h} \Theta_j \Sigma \Theta_{j+h}'$$

$$\Gamma(-h) = \Gamma(h)' = \sum_{j=0}^{q-h} \Theta_{j+h} \Sigma \Theta_j'$$

e.g. $\Gamma(-1) = E\left(\Theta_0 \underset{\sim}{\epsilon}_t + \Theta_1 \underset{\sim}{\epsilon}_{t-1} + \cdots + \Theta_q \underset{\sim}{\epsilon}_{t-q}\right)$

$$\left(\underset{\sim}{\epsilon}_{t-1}' \Theta_0' + \underset{\sim}{\epsilon}_{t-2}' \Theta_1' + \cdots + \underset{\sim}{\epsilon}_{t-q-1}' \Theta_q'\right)$$

$$= \Theta_1 \Sigma \Theta_0' + \cdots + \Theta_q \Sigma \Theta_{q-1}'$$

i.e. $\Gamma(-1) = \sum_{j=0}^{q-1} \Theta_{j+1} \Sigma \Theta_j' = \Gamma(1)'$

Also $\Gamma(h) = Cov(\underset{\sim}{X}_t, \underset{\sim}{X}_{t+h})$

$$= 0 \quad \forall\ |h| > q$$

VMA(q) is always Covariance stationary $\forall\ \Theta_1, \ldots, \Theta_q$, matrices of consts (a property like the univariate MA processes)

VMA(∞)

$$\underset{\sim}{X}_t = \sum_{j=0}^{\infty} \Theta_j \, \underset{\sim}{\epsilon}_{t-j} \; ; \; \underset{\sim}{\epsilon}_t \sim VWN(\underset{\sim}{0}, \Sigma)$$

Def^n: A seq of $n \times m$ matrices $\{H_s\}_{s=0}^{\infty}$ is said to be absolutely summable if each of its elements forms an absolutely summable scalar sequence

i.e. $\sum_{s=0}^{\infty} |H_{ij(s)}| < \infty \quad \forall i,j; \; i = 1(1)n, \; j = 1(1)m$

Suppose that for the VMA(∞) process the sequence of matrices $\{\Theta_j\}_{j=0}^{\infty}$ is absolutely summable,

i.e $\sum_{s=0}^{\infty} |\Theta_{ij(s)}| < \infty \quad \forall i,j = 1(1)K$

$(\Theta_{ij(s)} : (i,j)^{th}$ element of $\Theta_s)$

then VMA(∞) is covariance stationary

Note: If $\{\Theta_j\}$ is square summable, then also VMA(∞) would be covariance stationary.

If VMA(∞) is covariance stationary, then it's auto covariance matrix function is

$$\Gamma(h) = \sum_{j=0}^{\infty} \Theta_j \Sigma \Theta_{j+h}' \; ; \; h = 0, 1, 2, \ldots$$

$$\Gamma(-h) = \Gamma(h)'$$

VAR models

Consider a VAR(p) process (say K-variate)

$$\underset{K\times 1}{\underset{\sim}{X}_t} = \Phi_1 \underset{\sim}{X}_{t-1} + \Phi_2 \underset{\sim}{X}_{t-2} + \cdots + \Phi_p \underset{\sim}{X}_{t-p} + \underset{\sim}{\epsilon}_t ;$$

$$\Phi_p \neq 0, \quad \underset{\sim}{\epsilon}_t \sim VWN(\underset{\sim}{0}, \Sigma)$$

$$Cov(\underset{\sim}{\epsilon}_t, \underset{\sim}{X}_{t-j}) = 0 \quad \forall j > 0$$

i.e. $E(\underset{\sim}{\epsilon}_t \underset{\sim}{X}'_{t-j}) = 0 \quad \forall j > 0$

[$\underset{\sim}{X}_{p\times 1}$, $\underset{\sim}{Y}_{q\times 1}$ random vectors

$Cov(\underset{\sim}{X}, \underset{\sim}{Y}) = E(\underset{\sim}{X} - E(\underset{\sim}{X}))(\underset{\sim}{Y} - E(\underset{\sim}{Y}))'$]

$\Phi_1, \Phi_2, \ldots, \Phi_p$ are $K\times K$ matrices of constants – VAR matrices of parameters

Model can be written in terms of AR operator matrix polynomial

$$\Phi(B) \underset{\sim}{X}_t = \underset{\sim}{\epsilon}_t$$

where $\Phi(B) = I_K - \Phi_1 B - \cdots - \Phi_p B^p$.

VAR matrix polynomial

Let us see what the VAR(p) gives us. The l^{th} row of the VAR(p) system is (at time point t)

$$\begin{aligned} X_{l,t} = & \left(\Phi_{l1}^{(1)} X_{1,t-1} + \Phi_{l2}^{(1)} X_{2,t-1} + \cdots \cdots + \Phi_{lK}^{(1)} X_{K,t-1}\right) \\ & + \left(\Phi_{l2}^{(2)} X_{1,t-2} + \Phi_{l2}^{(2)} X_{2,t-2} + \cdots \cdots + \Phi_{lK}^{(2)} X_{K,t-2}\right) \\ & \quad \vdots \\ & + \left(\Phi_{l1}^{(p)} X_{1,t-p} + \Phi_{l2}^{(p)} X_{2,t-p} + \cdots \cdots + \Phi_{lK}^{(p)} X_{K,t-p}\right) \\ & + \epsilon_{lt} \end{aligned}$$

$l = 1(1)K \; ; \; t = 1(1)n$

i.e.

$$\begin{aligned} X_{l,t} = & \left(\Phi_{l1}^{(1)} X_{1,t-1} + \Phi_{l1}^{(2)} X_{1,t-2} + \cdots \cdots + \Phi_{l1}^{(p)} X_{1,t-p}\right) \\ & + \left(\Phi_{l2}^{(1)} X_{2,t-1} + \Phi_{l2}^{(2)} X_{2,t-2} + \cdots \cdots + \Phi_{l2}^{(p)} X_{2,t-p}\right) \\ & \quad \vdots \\ & + \left(\Phi_{lK}^{(1)} X_{K,t-1} + \Phi_{lK}^{(2)} X_{K,t-2} + \cdots \cdots + \Phi_{lK}^{(p)} X_{K,t-p}\right) \\ & + \epsilon_{lt} \end{aligned}$$

Thus the model eq^n for the l^{th} variable is expressed in terms of p lags of the l^{th} variable (as it would have been for AR(p))
+ p lags of all the remaining $K-1$ variables

present inside the VAR(p) system

+ instantaneous noise variable for the l^{th} variable system

Thus the VAR(p) system takes care of using inputs from related variables for a predictive model.

Example : VAR(1) with 2 variables

$$\underset{2\times 1}{\underset{\sim}{X}_t} = \Phi \underset{\sim}{X}_{t-1} + \underset{\sim}{\epsilon}_t \quad ; \quad \underset{\sim}{\epsilon}_t \sim VWN(\underset{\sim}{0}, \Sigma)$$

i.e $$\begin{pmatrix} X_{1,t} \\ X_{2,t} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} X_{1,t-1} \\ X_{2,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix}.$$

i.e. $$\begin{pmatrix} X_{1,t} \\ X_{2,t} \end{pmatrix} = \begin{pmatrix} \phi_{11} X_{1,t-1} + \phi_{12} X_{2,t-1} + \epsilon_{1,t} \\ \phi_{21} X_{1,t-1} + \phi_{22} X_{2,t-1} + \epsilon_{2,t} \end{pmatrix}$$

Also

$$(I_2 - \Phi B) \underset{\sim}{X}_t = \underset{\sim}{\epsilon}_t$$

i.e $$\Phi(B) \underset{\sim}{X}_t = \underset{\sim}{\epsilon}_t$$

$$\Phi(B) = \begin{pmatrix} 1-\phi_{11}B & -\phi_{12}B \\ -\phi_{21}B & 1-\phi_{22}B \end{pmatrix}$$ VAR matrix polynomial

Condition for stationarity of VAR(p)

A K-variate VAR(p) process is covariance stationary if all values of z satisfying $|\Phi(z)| = 0$ ($|A|$ is determinant of A) all lie outside the unit circle

i.e. all z satisfying

$$\left| I_K - \Phi_1 z - \Phi_2 z^2 - \cdots - \Phi_p z^p \right| = 0$$

lie outside the unit circle.

i.e. all y satisfying

$$\left| I_K y^p - \Phi_1 y^{p-1} - \cdots - \Phi_p \right| = 0$$

lie inside the unit circle

Example VAR(1) with $K=2$

$$\begin{pmatrix} X_{1,t} \\ X_{2,t} \end{pmatrix} = \begin{pmatrix} 1 & -.6 \\ .5 & -.7 \end{pmatrix} \begin{pmatrix} X_{1,t-1} \\ X_{2,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix}$$

$$\Phi(B) = I_2 - \begin{pmatrix} 1 & -.6 \\ .5 & -.7 \end{pmatrix} B$$

$$|\underset{\sim}{\Phi}(z)| = \begin{vmatrix} 1-z & .6z \\ -.5z & 1+.7z \end{vmatrix} = (1-z)(1+.7z)+.3z^2$$

$$= (1+.7z - z - .7z^2) + .3z^2$$

$$= 1 - .3z - .4z^2$$

$$= (1-.8z)(1+.5z)$$

Roots of $|\underset{\sim}{\Phi}(z)| = 0$ are $\frac{1}{.8}$, $-\frac{1}{.5}$

$\Rightarrow$ all z satisfying $|\underset{\sim}{\Phi}(z)| = 0$ lie outside the unit circle

$\Rightarrow$ the VAR(1) process is covariance stationary.

<u>Remark</u>: w.l.o.g. we can take mean vector of a covariance stationary VAR(1) as null vector, i.e. we take w.l.o.g. a VAR(1) (covariance stationary) without a const vector in the model,

o/w $\underset{\sim}{Y}_{t\,(K\times 1)} = \underset{\sim}{\delta} + \underset{\sim}{\Phi}_1 \underset{\sim}{Y}_{t-1} + \cdots + \underset{\sim}{\Phi}_p \underset{\sim}{Y}_{t-p} + \underset{\sim}{\epsilon}_t$ $\quad (\underset{\sim}{\epsilon}_t \sim VWN(\underset{\sim}{0}, I))$

$$E(\underset{\sim}{Y}_t) = E(\underset{\sim}{\delta} + \underset{\sim}{\Phi}_1 \underset{\sim}{Y}_{t-1} + \cdots + \underset{\sim}{\Phi}_p \underset{\sim}{Y}_{t-p} + \underset{\sim}{\epsilon}_t)$$

i.e. $\underset{\sim}{\mu} = \underset{\sim}{\delta} + \underset{\sim}{\Phi}_1 \underset{\sim}{\mu} + \cdots + \underset{\sim}{\Phi}_p \underset{\sim}{\mu}$

$$\Rightarrow (I_K - \underline{\Phi}_1 - \cdots - \underline{\Phi}_p)\,\underset{\sim}{\mu} = \underset{\sim}{\delta}$$

$$\Rightarrow \underset{\sim}{\mu} = (I_K - \underline{\Phi}_1 - \cdots - \underline{\Phi}_p)^{-1}\,\underset{\sim}{\delta}$$

i.e. $\underset{\sim}{\mu} = \left(\underline{\Phi}(1)\right)^{-1}\,\underset{\sim}{\delta}$

$$\left(\begin{array}{l} \text{Note that } \underline{\Phi}(z) = I_K - \underline{\Phi}_1 z - \cdots - \underline{\Phi}_p z^p \\ \qquad\qquad\quad \underline{\Phi}(1) = I_K - \underline{\Phi}_1 - \cdots - \underline{\Phi}_p \\ \qquad\quad \text{\& } \underline{\Phi}(1) \text{ is non singular} \\ \qquad\quad \text{as } |\underline{\Phi}(1)| \neq 0 \text{ for stationary process.} \end{array}\right)$$

$$\underset{\sim}{\delta} = \underline{\Phi}(1)\,\underset{\sim}{\mu}$$

$$\Rightarrow \underset{\sim}{Y}_t - \underset{\sim}{\mu} = \underline{\Phi}_1(\underset{\sim}{Y}_{t-1} - \underset{\sim}{\mu}) + \cdots + \underline{\Phi}_p(\underset{\sim}{Y}_{t-p} - \underset{\sim}{\mu}) + \underset{\sim}{\epsilon}_t$$

let $\underset{\sim}{X}_t = \underline{Y}_t - \underset{\sim}{\mu}$

$$\Rightarrow \underset{\sim}{X}_t = \underline{\Phi}_1(\underset{\sim}{X}_{t-1}) + \cdots + \underline{\Phi}_p \underset{\sim}{X}_{t-p} + \underline{\epsilon}_t$$

equivalent VAR(p) with same VWN and $\underline{\Phi}_1, \ldots, \underline{\Phi}_p$ and without const vector

A. Mean vector of covariance stationary VAR(p)

$$\underset{\sim}{X}_t = \underline{\Phi}_1 \underset{\sim}{X}_{t-1} + \cdots + \underline{\Phi}_p \underset{\sim}{X}_{t-1} + \underset{\sim}{\epsilon}_t$$

$$\underline{\Phi}_p \neq 0, \quad \epsilon_t \sim VWN(\underline{0}, \Sigma)$$

$$Cov(\underset{\sim}{\epsilon}_t, \underset{\sim}{X}_{t-j}) = 0 \quad \forall j > 0$$

$$\underset{\sim}{X}_t = \Phi_1 \underset{\sim}{X}_{t-1} + \cdots + \Phi_p \underset{\sim}{X}_{t-p} + \underset{\sim}{\epsilon}_t$$

$$\Rightarrow \underset{\sim}{\mu} = E(\underset{\sim}{X}_t) = \Phi_1 E(\underset{\sim}{X}_{t-1}) + \cdots + \Phi_p E(\underset{\sim}{X}_{t-p}) + \underset{\sim}{0}$$

i.e. $\underset{\sim}{\mu} = \Phi_1 \underset{\sim}{\mu} + \cdots + \Phi_p \underset{\sim}{\mu}$

i.e. $(I_K - \Phi_1 - \cdots - \Phi_p)\underset{\sim}{\mu} = \underset{\sim}{0}$

i.e. $\Phi(1)\underset{\sim}{\mu} = \underset{\sim}{0}$

$$\Rightarrow \underset{\sim}{\mu} = \underset{\sim}{0} \quad (|\Phi(1)| \neq 0).$$

Auto covariance matrix function

$$\Gamma_0 = \mathrm{Cov}(\underset{\sim}{X}_t, \underset{\sim}{X}_t) = E(\underset{\sim}{X}_t \underset{\sim}{X}_t')$$

$$= E\left(\underset{\sim}{X}_t \left(\Phi_1 \underset{\sim}{X}_{t-1} + \cdots + \Phi_p \underset{\sim}{X}_{t-p} + \underset{\sim}{\epsilon}_t\right)'\right)$$

$$= E(\underset{\sim}{X}_t \underset{\sim}{X}_{t-1}')\Phi_1' + E(\underset{\sim}{X}_t \underset{\sim}{X}_{t-2}')\Phi_2' + \cdots$$

$$+ E(\underset{\sim}{X}_t \underset{\sim}{X}_{t-p}')\Phi_p' + E(\underset{\sim}{X}_t \underset{\sim}{\epsilon}_t')$$

Note that

$$E(\underset{\sim}{X}_t \underset{\sim}{\epsilon}_t') = E\left(\left(\Phi_1 \underset{\sim}{X}_{t-1} + \cdots + \Phi_p \underset{\sim}{X}_{t-p}\right) + \epsilon_t\right)\underset{\sim}{\epsilon}_t'$$

$$= \Phi_1 \mathrm{Cov}(\underset{\sim}{X}_{t-1}, \underset{\sim}{\epsilon}_t) + \cdots + \Phi_p \mathrm{Cov}(\underset{\sim}{X}_{t-p}, \underset{\sim}{\epsilon}_t)$$

$$+ \mathrm{Cov}(\underset{\sim}{\epsilon}_t, \underset{\sim}{\epsilon}_t)$$

$$= \Sigma \quad \left(\mathrm{Cov}(\underset{\sim}{X}_{t-j}, \underset{\sim}{\epsilon}_t) = 0 \ \forall j > 0\right)$$

$$\Rightarrow \Gamma_0 = \Gamma(-1)\,\Phi_1' + \Gamma(-2)\,\Phi_2' + \cdots + \Gamma(-p)\,\Phi_p' + \Sigma \quad \text{—} (*')$$

Note further that

$$\Gamma_0 = E(\underline{X}_t \underline{X}_t') \quad \text{can be derived through}$$

$$= E(\Phi_1 \underset{\sim}{X}_{t-1} + \cdots + \Phi_p \underset{\sim}{X}_{t-p} + \underset{\sim}{\epsilon}_t)\, \underline{X}_t'$$

$$= \Phi_1 E(\underset{\sim}{X}_{t-1} \underset{\sim}{X}_t') + \cdots + \Phi_p E(\underset{\sim}{X}_{t-p} \underline{X}_t') + E(\underline{\epsilon}_t \underline{X}_t')$$

$$= \Phi_1 \left(E(\underline{X}_t \underset{\sim}{X}_{t-1}') \right)' + \cdots + \Phi_p \left(E(\underline{X}_t \underset{\sim}{X}_{t-p}') \right)' + \Sigma$$

i.e $\Gamma_0 = \Phi_1 (\Gamma(-1))' + \cdots + \Phi_p (\Gamma(-p))' + \Sigma$

i.e $\Gamma_0 = \Phi_1 \Gamma(1) + \cdots + \Phi_p \Gamma(p) + \Sigma$ — $(*^2)$

$\left((\Gamma(-h))' = \Gamma(h) \right)$

$(*') = (*^2)$ — Γ_0 is anyway symmetric matrix

$$Cov(\underset{\sim}{X}_t, X_{t+h}) = E(\underset{\sim}{X}_t \underset{\sim}{X}'_{t+h})$$

$$= E(\underset{\sim}{X}_t (\underline{\Phi}_1 \underset{\sim}{X}_{t+h-1} + \cdots + \underline{\Phi}_p \underset{\sim}{X}_{t+h-p} + \underset{\sim}{\epsilon}_{t+h})')$$

$\forall h > 0$

$$\Gamma_h = Cov(\underset{\sim}{X}_t, X_{t+h})$$

$$= E(\underset{\sim}{X}_t \underset{\sim}{X}'_{t+h-1})\underline{\Phi}'_1 + \cdots + E(\underset{\sim}{X}_t X'_{t+h-p})\underline{\Phi}'_p$$

$$\left(E(\underset{\sim}{X}_t \underset{\sim}{\epsilon}'_{t+h}) = Cov(\underset{\sim}{X}_t, \epsilon_{t+h}) = 0 \text{ as } h>0\right)$$

$$\Rightarrow \Gamma(h) = \Gamma(h-1)\underline{\Phi}'_1 + \cdots + \Gamma(h-p)\underline{\Phi}'_p$$

matrix Yule-Walker eq^n.

Further $\Gamma(-h) = (\Gamma(h))'$

Causal VAR process

Def^n: A VAR(p) process is said to be Causal if it can be expressed in terms of a VWN sequence as a VMA(∞) form

VMA(∞) representation of VAR

VAR(1) $\quad \underset{\sim}{X}_t = \Phi\, \underset{\sim}{X}_{t-1} + \underset{\sim}{\epsilon}_t \,;\quad \underset{\sim}{\epsilon}_t \sim VWN(\underset{\sim}{0}, \Sigma)$

$$\Phi(B)\, \underset{\sim}{X}_t = \underset{\sim}{\epsilon}_t$$

$$\Phi(B) = I_K - \Phi B$$

Suppose $\Phi(B)^{-1}$ is the inverse of the operator $\Phi(B)$

i.e. $\Phi(B)^{-1}\, \Phi(B) = I_K$

then $\quad \underset{\sim}{X}_t = \Phi(B)^{-1} \underset{\sim}{\epsilon}_t = \Psi(B)\, \underset{\sim}{\epsilon}_t$, say

where $\Psi(B) = \Psi_0 + \Psi_1 B + \Psi_2 B^2 + \cdots\cdots$

$$\Phi(B)^{-1} = \Psi(B)$$

i.e $\quad I_K = \Phi(B)\, \Psi(B)$

i.e. $\quad I_K = (I_K - \Phi B)(\Psi_0 + \Psi_1 B + \Psi_2 B^2 + \cdots)$

$$= (\Psi_0 + \Psi_1 B + \Psi_2 B^2 + \cdots\ )$$
$$- (\Phi \Psi_0 B + \Phi \Psi_1 B^2 + \cdots\ )$$

i.e. $I_K = \Psi_0 + (\Psi_1 - \Phi \Psi_0) B + (\Psi_2 - \Phi \Psi_1) B^2 + \cdots$

Comparing coefficient of B^j

B^0 : $\Psi_0 = I_K$

B^1 : $\Psi_1 = \Phi$

B^2 : $\Psi_2 = \Phi \Psi_1 = \Phi^2$

$\vdots$

B^j : $\Psi_j = \Phi^j$

$$\Rightarrow \underset{\sim}{X}_t = \sum_{j=0}^{\infty} \Phi^j \underset{\sim}{\epsilon}_{t-j}$$

This is the causal representation a covariance stationary VAR(1)

<u>VAR(p)</u> Suppose $\{X_t\}$ is covariance stationary VAR(p)

$$\underset{\sim}{X}_t = \Phi_1 \underset{\sim}{X}_{t-1} + \cdots + \Phi_p \underset{\sim}{X}_{t-p} + \underset{\sim}{\epsilon}_t$$

$$\underset{\sim}{\epsilon}_t \sim VWN(\underset{\sim}{0}, \Sigma)$$

$$\Phi(B) \underset{\sim}{X}_t = \underset{\sim}{\epsilon}_t \;;\; \Phi(B) = I_K - \sum_{i=1}^{p} \Phi_i B^i$$

Suppose $\underset{\sim}{X}_t = \Phi(B)^{-1} \underset{\sim}{\epsilon}_t = \Psi(B) \underset{\sim}{\epsilon}_t = \sum_{j=0}^{\infty} \Psi_j \underset{\sim}{\epsilon}_{t-j}$

$\Psi(B)$ is $\ni$ $\Phi(B)\,\Psi(B) = I_K$

i.e. $\left(I_K - \sum_{i=1}^{p} \Phi_i B^i \right) \left(\sum_{j=0}^{\infty} \Psi_j B^j \right) = I_K$

i.e. $\left(I_K - \Phi_1 B - \Phi_2 B^2 - \cdots - \Phi_p B^p\right)$

$$\left(\Psi_0 + \Psi_1 B + \Psi_2 B^2 + \cdots\right) = I_K$$

i.e. $\Psi_0 B^0 + (\Psi_1 - \Phi_1 \Psi_0) B + (\Psi_2 - \Phi_1 \Psi_1 - \Phi_2 \Psi_0) B^2$

$$+ (\Psi_3 - \Phi_1 \Psi_2 - \Phi_2 \Psi_1 - \Phi_3 \Psi_0) B^3 + \cdots$$

$$= I_K$$

Comparing coefficients, we have

$$\Psi_0 = I_K$$

$$\Psi_1 = \Phi_1$$

$$\Psi_2 = \Phi_1 \Psi_1 + \Phi_2 \Psi_0$$

$$\Psi_3 = \Phi_1 \Psi_2 + \Phi_2 \Psi_1 + \Phi_3 \Psi_0$$

In general, $\forall s \geq 2$

$$\Psi_s = \Phi_1 \Psi_{s-1} + \Phi_2 \Psi_{s-2} + \cdots + \Phi_p \Psi_{s-p}$$

with $\Psi_0 = I_K$ & $\Psi_l = 0 \;\forall\, l < 0$

$\Psi_0, \Psi_1, \Psi_2, \ldots$ gives the Causal VMA(∞) representation of VAR(p).

<u>Remark</u> : It is always possible to write a stationary VAR(p) process as VMA(∞) process with the VWN process having elements that are mutually uncorrelated.

Consider the VMA(∞) representation of VAR(p).

$$\underset{K\times 1}{\underset{\sim}{X}_t} = \underset{\sim}{\epsilon}_t + \Psi_1 \underset{\sim}{\epsilon}_{t-1} + \Psi_2 \underset{\sim}{\epsilon}_{t-2} + \cdots$$

$$\underset{\sim}{\epsilon}_t \sim VWN(\underline{0}, \Sigma), \ \Sigma > 0$$

Note that $\exists$ a non-singular matrix $H \ni$

$$H \Sigma H' = D_\lambda \text{ (diagonal matrix)}$$

Using the above non-singular matrix H, we can write

$$\underset{\sim}{X}_t = H^{-1} H \underset{\sim}{\epsilon}_t + \Psi_1 H^{-1} H \underset{\sim}{\epsilon}_{t-1} + \Psi_2 H^{-1} H \underset{\sim}{\epsilon}_{t-2} + \cdots$$

i.e. $\underset{\sim}{X}_t = \Psi_0^* \underset{\sim}{\eta}_t + \Psi_1^* \underset{\sim}{\eta}_{t-1} + \Psi_2^* \underset{\sim}{\eta}_{t-2} + \cdots$

$$\Psi_i^* = \Psi_i H^{-1} \ ; \ i = 0, 1, \cdots$$

$$\Psi_0 = I_K$$

$\underset{\sim}{\eta}_t = H \underset{\sim}{\epsilon}_t$ is $\ni$ $E(\underset{\sim}{\eta}_t) = H\, E(\underset{\sim}{\epsilon}_t) = \underset{\sim}{0}$

$$Cov(\underset{\sim}{\eta}_t, \underset{\sim}{\eta}_s) = E(\underset{\sim}{\eta}_t \underset{\sim}{\eta}_s')$$

i.e. $Cov(\underset{\sim}{\eta}_t, \underset{\sim}{\eta}_s) = E(H\underset{\sim}{\epsilon}_t)(H\underset{\sim}{\epsilon}_s)'$

$$= H\, E(\underset{\sim}{\epsilon}_t \underset{\sim}{\epsilon}_s')\, H'$$

$$= \begin{cases} H\Sigma H' = D_\lambda, & \text{If } t = s \\ 0, & \text{If } t \neq s \end{cases}$$

$\Rightarrow \underset{\sim}{\eta}_t \sim VWN(\underset{\sim}{0}, D_\lambda)$

$\Rightarrow$ elements of $\underset{\sim}{\eta}_t$ are uncorrelated

$\Rightarrow \{\underset{\sim}{X}_t\}$ is expressed as VMA(∞) with VWN process with uncorrelated components.

Vector ARMA(p,q)

$\underset{\sim}{X}_t \sim \text{VARMA}(p,q)$ if

$$\underset{k\times 1}{\underset{\sim}{X}_t} = \Phi_1 \underset{\sim}{X}_{t-1} + \cdots + \Phi_p \underset{\sim}{X}_{t-p} + \underset{\sim}{\epsilon}_t + \Theta_1 \underset{\sim}{\epsilon}_{t-1} + \cdots + \Theta_q \underset{\sim}{\epsilon}_{t-q}$$

$$\Phi_p \neq 0, \quad \Theta_q \neq 0, \quad \underset{\sim}{\epsilon}_t \sim \text{VWN}(\underset{\sim}{0}, \Sigma)$$

$$\text{Cov}(\underset{\sim}{\epsilon}_t, \underset{\sim}{X}_{t-j}) = 0 \quad \forall j > 0$$

Conditions for stationarity

$\underset{\sim}{X}_t$ is covariance stationary if all values of z satisfying

$$|I_k - \Phi_1 z - \cdots - \Phi_p z^p| = 0$$ lie outside the unit circle

i.e. all z satisfying $|\Phi(z)| = 0$ lie outside the unit circle (with model as $\Phi(B)\underset{\sim}{X}_t = \Theta(B)\underset{\sim}{\epsilon}_t$)

Every covariance stationary vector ARMA(p,q) has a causal representation through,

$$\underset{\sim}{X}_t = \Phi(B)^{-1}\Theta(B)\underset{\sim}{\epsilon}_t = \Psi(B)\underset{\sim}{\epsilon}_t, \text{ say}$$

i.e. $\Phi(B)^{-1}\Theta(B) = \Psi(B)$

$$\Rightarrow \Theta(B) = \Phi(B)\Psi(B)$$

i.e. $(I_k + \Theta_1 B + \cdots + \Theta_q B^q) = (I_k - \Phi_1 B - \cdots - \Phi_p B^p)(\Psi_0 + \Psi_1 B + \cdots)$

i.e.

$$(I_K + \Theta_1 B + \cdots + \Theta_q B^q)$$

$$= \Psi_0 + (\Psi_1 - \Phi_1 \Psi_0) B + (\Psi_2 - \Phi_1 \Psi_1 - \Phi_2 \Psi_0) B^2 +$$

$$(\Psi_3 - \Phi_1 \Psi_2 - \Phi_2 \Psi_1 - \Phi_3 \Psi_0) + \cdots .$$

comparing coefficients,

B^0 : $\Psi_0 = I_K$

B^1 : $\Psi_1 = \Theta_1 + \Phi_1$

B^2 : $\Psi_2 = \Phi_1 \Psi_1 + \Phi_2 \Psi_0 + \Theta_2$

B^3 : $\Psi_3 = \Phi_1 \Psi_2 + \Phi_2 \Psi_1 + \Phi_3 \Psi_0 + \Theta_3$

⋮

In general,

$$\Psi_s = \Phi_1 \Psi_{s-1} + \Phi_2 \Psi_{s-2} + \cdots + \Phi_p \Psi_{s-p} + \Theta_s$$

$$\forall \ s \le q$$

$$\& \quad \Psi_s = \Phi_1 \Psi_{s-1} + \Phi_2 \Psi_{s-2} + \cdots + \Phi_p \Psi_{s-p}$$

$$\forall \ s > q$$

with $\Psi_r = 0 \ \forall \ r < 0$ & $\Psi_r = I_K$ if $r = 0$

Def: $\underset{\sim}{X}_t$ is said to be invertible if all values of z satisfying $|\Theta(z)| = 0$ lie outside the unit circle. In this case,

$$\underset{\sim}{\epsilon}_t = \Theta(B)^{-1} \Phi(B) \underset{\sim}{X}_t = \Psi(B) \underset{\sim}{X}_t$$

Vector ARMA$(p,q) \longrightarrow$ VAR(∞) use comparing coeffs.

VMA(q) to VAR(∞)

VMA(q) model

VMA(q) is invertible if all z satisfying $|\Theta(z)| = 0$ lie outside the unit circle.

$$\underset{\sim}{X}_t = \underset{\sim}{\epsilon}_t + \Theta_1 \underset{\sim}{\epsilon}_{t-1} + \cdots + \Theta_q \underset{\sim}{\epsilon}_{t-q}$$

$$\Theta_q \neq 0, \quad \underset{\sim}{\epsilon}_t \sim VWN(\underset{\sim}{0}, \Sigma)$$

$$\underset{\sim}{X}_t = \Theta(B) \underset{\sim}{\epsilon}_t$$

$$\Theta(B) = I_K + \Theta_1 B + \cdots + \Theta_q B^q$$

Let $\Theta(B)^{-1}$ denote the inverse operator of $\Theta(B)$, then

$$\underset{\sim}{\epsilon}_t = \Theta(B)^{-1} \underset{\sim}{X}_t = \Psi(B) \underset{\sim}{X}_t = \sum_{j=0}^{\infty} \Psi_j \underset{\sim}{X}_{t-j}$$

$$\Psi(B) = \Theta(B)^{-1}$$

i.e. $I_K = \Theta(B)\, \Psi(B)$

$$I_K = (I_K + \Theta_1 B + \cdots + \Theta_q B^q)(\Psi_0 + \Psi_1 B + \Psi_2 B^2 + \cdots)$$

i.e. $I_K = \Psi_0 + (\Theta_1 \Psi_0 + \Psi_1) B + (\Psi_2 + \Theta_1 \Psi_1 + \Theta_2 \Psi_0) B^2 + \cdots$

Comparing coeff of B^j:

B^0: $\Psi_0 = I_K$

B^1: $\Psi_1 = -\Theta_1 \Psi_0 = -\Theta_1$

B^2: $\Psi_2 = -\Theta_1 \Psi_1 - \Theta_2 \Psi_0$

$\Psi_2 = \Theta_1^2 - \Theta_2$ $\quad \rightarrow \Psi_s = -(\Theta_1 \Psi_{s-1} + \Theta_2 \Psi_{s-2} + \cdots + \Theta_q \Psi_{s-q})$

$\Psi_0 = I \quad \Psi_l = 0 \ \forall\, l < 0$

B^3: $\Psi_3 = -\Theta_1 \Psi_2 - \Theta_2 \Psi_1 - \Theta_3 \Psi_0$

i.e. $\Psi_3 = -\Theta_1(\Theta_1^2 - \Theta_2) - \Theta_2(-\Theta_1) - \Theta_3$

$$-\Theta_1^2 + \Theta_1 \Theta_2 + \Theta_2 \Theta_1 - \Theta_3$$

Remark : Impulse response function (IRF)

Covariance stationary VAR(p) → VMA(∞)

Covariance stationary VARMA(p,q) → VMA(∞)

$$X_t = \underset{\sim}{\epsilon}_t + \Psi_1 \underset{\sim}{\epsilon}_{t-1} + \Psi_2 \underset{\sim}{\epsilon}_{t-2} + \cdots$$

$$\begin{pmatrix} X_{1t} \\ \vdots \\ X_{Kt} \end{pmatrix} = \begin{pmatrix} \epsilon_{1t} \\ \vdots \\ \epsilon_{Kt} \end{pmatrix} + \begin{pmatrix} \psi_{11}^{(1)} & \cdots & \psi_{1K}^{(1)} \\ \cdots & \cdots & \cdots \\ \psi_{K1}^{(1)} & \cdots & \psi_{KK}^{(1)} \end{pmatrix} \begin{pmatrix} \epsilon_{1,t-1} \\ \vdots \\ \epsilon_{K,t-1} \end{pmatrix} + \cdots$$

$$\cdots + \begin{pmatrix} \psi_{11}^{(s)} & \cdots & \psi_{1K}^{(s)} \\ \cdots & \cdots & \cdots \\ \psi_{K1}^{(s)} & \cdots & \psi_{KK}^{(s)} \end{pmatrix} \begin{pmatrix} \epsilon_{1,t-s} \\ \vdots \\ \epsilon_{K,t-s} \end{pmatrix} + \cdots$$

$$X_{it} = \epsilon_{it} + \left(\psi_{i1}^{(1)} \epsilon_{1,t-1} + \cdots + \psi_{iK}^{(1)} \epsilon_{K,t-1}\right) + \cdots$$

$$\cdots + \left(\psi_{i1}^{(s)} \epsilon_{1,t-s} + \cdots + \psi_{iK}^{(s)} \epsilon_{K,t-s}\right) + \cdots$$

$i = 1(1)K$

$$\frac{\partial X_{i,t+s}}{\partial \epsilon_{j,t}} = \psi_{ij}^{(s)}$$

$$\frac{\partial \underset{\sim}{X}_{t+s}}{\partial \underset{\sim}{\epsilon}_t} = \Psi_s$$

Remark : Impulse response function (IRF)

Covariance stationary $VAR(p) \rightarrow VMA(\infty)$

Covariance stationary $VARMA(p,q) \rightarrow VMA(\infty)$

$$\underset{\sim}{X}_t = \underset{\sim}{\epsilon}_t + \Psi_1 \underset{\sim}{\epsilon}_{t-1} + \Psi_2 \underset{\sim}{\epsilon}_{t-2} + \cdots$$

$$\begin{pmatrix} X_{1t} \\ \vdots \\ X_{Kt} \end{pmatrix} = \begin{pmatrix} \epsilon_{1t} \\ \vdots \\ \epsilon_{Kt} \end{pmatrix} + \begin{pmatrix} \psi_{11}^{(1)} & \cdots & \psi_{1K}^{(1)} \\ - & - & - \\ \psi_{K1}^{(1)} & \cdots & \psi_{KK}^{(1)} \end{pmatrix} \begin{pmatrix} \epsilon_{1,t-1} \\ \vdots \\ \epsilon_{K,t-1} \end{pmatrix} + \cdots$$

$$\cdots + \begin{pmatrix} \psi_{11}^{(s)} & \cdots & \psi_{1K}^{(s)} \\ - & - & - \\ \psi_{K1}^{(s)} & \cdots & \psi_{KK}^{(s)} \end{pmatrix} \begin{pmatrix} \epsilon_{1,t-s} \\ \vdots \\ \epsilon_{K,t-s} \end{pmatrix} + \cdots$$

$$X_{it} = \epsilon_{it} + \left(\psi_{i1}^{(1)} \epsilon_{1,t-1} + \cdots + \psi_{iK}^{(1)} \epsilon_{K,t-1} \right) + \cdots$$

$$\cdots + \left(\psi_{i1}^{(s)} \epsilon_{1,t-s} + \cdots + \psi_{iK}^{(s)} \epsilon_{K,t-s} \right) + \cdots$$

$i = 1(1)K$

$$\frac{\partial X_{i,t+s}}{\partial \epsilon_{j,t}} = \psi_{ij}^{(s)}$$

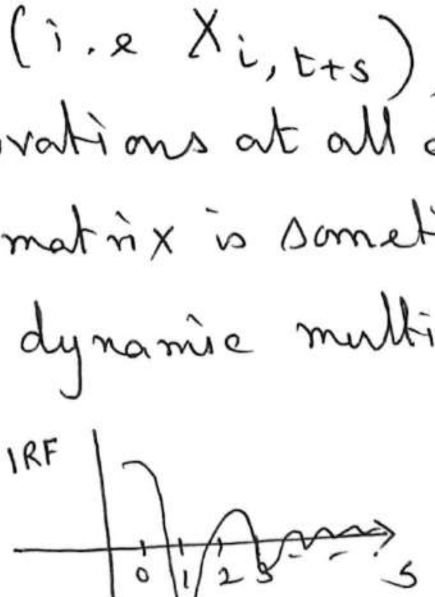
$$\frac{\partial \underset{\sim}{X}_{t+s}}{\partial \underset{\sim}{\epsilon}_t} = \Psi_s$$

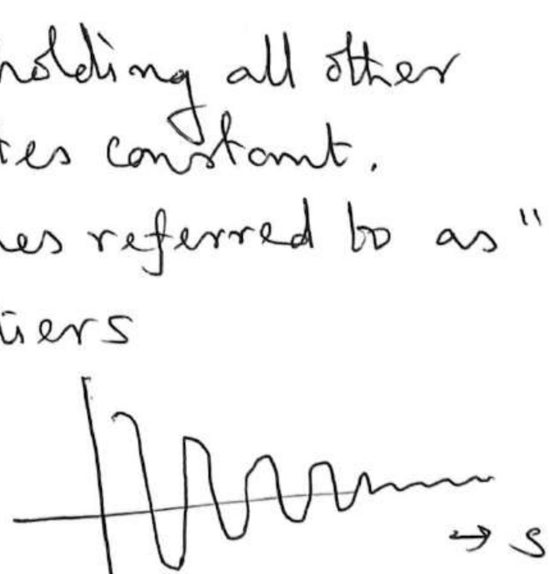
Def^n: Plot of $\dfrac{\partial X_{i,t+s}}{\partial \epsilon_{j,t}}$ as a f^n of s is called the impulse response function (IRF) plot, as it describes the response of X_i at time point $(t+s)$ to a one time impulse in X_j at time point t

Note: IRF can be interpreted as response of X_i at future time point at lead periods $1, 2, \ldots$; w.r.t. a shock in the jth variable X_j at a time point t.

Note: $\psi_{ij}^{(s)}$ indicates the consequence of one unit increase in the jth variable's innovation at time point t $(\epsilon_{j,t})$ on the value of the ith variable at time $t+s$ (i.e $X_{i,t+s}$), holding all other innovations at all dates constant.

Note: Ψ_s matrix is sometimes referred to as "matrix of dynamic multipliers

Example: 



Auto Regressive Integrated Moving Average (ARIMA) model

Realise that many time series processes which are non-stationary but an appropriate order differenced process, derived from the non-stationary process, can be a stationary ARMA.

e.g. (i) $X_t = X_{t-1} + \epsilon_t$; $\epsilon_t \sim WN(0, \sigma^2)$ is a non-stationary process, but

$$Y_t = \nabla X_t = X_t - X_{t-1} = \epsilon_t \text{ is stationary}$$

$Y_t \sim ARMA(0,0) \equiv WN$

(ii) $X_t = m_t + Y_t$

Y_t : Covariance stationary

$m_t = \sum_{j=0}^{K} \beta_j t^j$ is the time trend

$$\nabla^K X_t = \nabla^K m_t + \nabla^K Y_t$$

$Z_t = \nabla^K X_t = K!\, \beta_K + \nabla^K Y_t$ is stationary and can be modeled using ARMA

Def^n: Integrated process

A time series $\{X_t\}$ is said to be integrated of order d ($X_t \sim I_d$) if d is the smallest integer $\ni$ $\nabla^d X_t$ is a stationary process.

Note: In (i) $X_t \sim I_1$ (or $I(1)$)

(ii) $X_t \sim I_K$ (or $I(K)$)

Def^n: ARIMA model

$\{X_t\}$ is said to follow an Auto Regressive Integrated Moving Average (ARIMA) model of order (p, d, q) if

$$Z_t = \nabla^d X_t = (1-B)^d X_t \sim ARMA(p, q)$$

i.e. $Z_t = \phi_1 Z_{t-1} + \cdots + \phi_p Z_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}$

i.e. $\phi(B) Z_t = \theta(B) \epsilon_t$ $\left(\begin{array}{l} \phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p \\ \theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q \end{array} \right)$

i.e. $\phi(B) \nabla^d X_t = \theta(B) \epsilon_t$

i.e. $\underbrace{\phi(B)(1-B)^d X_t = \theta(B) \epsilon_t}_{\text{model in terms of } \{X_t\}}$ — (*)

Note that (i) $ARIMA(p, 0, q) \equiv ARMA(p, q)$

(ii) $ARIMA(p, d, q) \equiv ARMA(p+d, q)$

from (*), the model for $\{X_t\}$ is

$$\phi^*(B) X_t = \theta(B) \epsilon_t.$$

$$\phi^*(B) = \phi(B)(1-B)^d \Rightarrow X_t \sim ARMA(p+d, q)$$

(iii) $X_t \sim ARIMA(p,d,q)$

i.e. $X_t \sim ARMA(p+d, q)$

X_t is always a non-stationary ARMA, ~~irrespective~~ irrespective of covariance stationarity of $Z_t \sim ARMA(p,q)$ $(Z_t = \nabla^d X_t)$

$\phi^*(z) = \phi(z)(1-z)^d$ has d roots on unit circle $\Rightarrow X_t$ is non-stationary.

Example: $X_t \sim ARIMA(1,1,1)$

i.e. $\nabla X_t = Z_t$ is $\ni$

$$Z_t = \phi Z_{t-1} + \epsilon_t + \theta \epsilon_{t-1}$$

$$|\phi|<1, \quad \epsilon_t \sim WN(0, \sigma^2)$$

i.e $X_t - X_{t-1} = \phi(X_{t-1} - X_{t-2}) + \epsilon_t + \theta\epsilon_{t-1}$

i.e. $X_t = (1+\phi)X_{t-1} - \phi X_{t-2} + \epsilon_t + \theta \epsilon_{t-1}$

i.e. $X_t \sim ARMA(2,1)$

Also $(1-\phi B)(1-B) X_t = (1+\theta B)\epsilon_t$

i.e. $\phi^*(B) X_t = \theta(B)\epsilon_t$

$\phi^*(z)$ has one unit root

$\Rightarrow X_t$ is non-stationary ARMA(2,1)

Seasonal ARMA model

Seasonal MA(Q): $MA(Q)_s$

$$X_t = \epsilon_t + \theta_1^{(s)} \epsilon_{t-s} + \theta_2^{(s)} \epsilon_{t-2s} + \cdots + \theta_Q^{(s)} \epsilon_{t-Qs}$$

$$\epsilon_t \sim WN(0, \sigma^2)$$

s: period of seasonality

Seasonal AR(P): $AR(P)_s$

$$X_t = \phi_1^{(s)} X_{t-s} + \phi_2^{(s)} X_{t-2s} + \cdots + \phi_P^{(s)} X_{t-Ps} + \epsilon_t$$

$$\epsilon_t \sim WN(0, \sigma^2)$$

s: period of seasonality

Seasonal ARMA(P, Q): $ARMA(P,Q)_s$

$$X_t = \phi_1^{(s)} X_{t-s} + \phi_2^{(s)} X_{t-2s} + \cdots + \phi_P^{(s)} X_{t-Ps}$$
$$+ \epsilon_t + \theta_1^{(s)} \epsilon_{t-s} + \theta_2^{(s)} \epsilon_{t-2s} + \cdots + \theta_Q^{(s)} \epsilon_{t-Qs}$$

i.e. $\left(1 - \phi_1^{(s)} B^s - \phi_2^{(s)} B^{2s} - \cdots - \phi_P^{(s)} B^{Ps}\right) X_t$

$$= \left(1 + \theta_1^{(s)} B^s + \theta_2^{(s)} B^{2s} + \cdots + \theta_Q^{(s)} B^{Qs}\right) \epsilon_t$$

$$\underbrace{\Phi^{(s)}(B^s)}_{\text{Seasonal AR polynomial}} X_t = \underbrace{\Theta^{(s)}(B^s)}_{\text{Seasonal MA polynomial}} \epsilon_t$$

Note: $ARMA(P,Q)_s$ is stationary and causal if roots of $0 = \Phi^{(s)}(z)$ all lie outside the unit circle; it is invertible if roots of $\Theta^{(s)}(z) = 0$ all lie outside unit circle

Mixed seasonal model

$$\text{ARMA}(p,q)\times\text{ARMA}(P,Q)_s \equiv \text{ARMA}(p,q)(P,Q)_s$$

$$\Phi^{(s)}(B^s)\,\phi(B)\,X_t = \Theta^{(s)}(B^s)\,\theta(B)\,\epsilon_t$$

$\epsilon_t \sim WN(0,\sigma^2)$

$\Phi^{(s)}(B^s)$: seasonal AR polynomial (order P)

$\phi(B)$: non-seasonal AR polynomial (order p)

$\Theta^{(s)}(B^s)$: seasonal MA polynomial (order Q)

$\theta(B)$: non-seasonal MA polynomial (order q)

Note that $\text{ARMA}(p,q)(P,Q)_s \equiv \text{ARMA}(Ps+p,\, Qs+q)$

Mixed Seasonal ARIMA model

Let $\{X_t\}$ be a non-stationary process and suppose

$Y_t = \nabla^d \nabla_s^D X_t$ is stationary

s: period of seasonality

d: order of non-seasonal differencing

D: order of seasonal differencing

Suppose $Y_t = (1-B)^d(1-B^s)^D X_t$ follows an

ARMA(p,q) for non-seasonal part and ARMA$(P,Q)_s$ for the seasonal part

i.e. $Y_t \sim$ ARMA $(p,q)(P,Q)_s$

i.e. $\Phi^{(s)}(B^s)\,\phi(B)\,Y_t = \Theta^{(s)}(B^s)\,\theta(B)\,\epsilon_t$

$\epsilon_t \sim WN(0,\sigma^2)$

then X_t is said to have a mixed seasonal ARIMA $(p,d,q)(P,D,Q)_s$ (or just S-ARIMA $(p,d,q)(P,D,Q)_s$ model).

Note: $Y_t \sim$ ARMA$(p,q)(P,Q)_s$

$\Rightarrow Y_t \sim$ ARMA$(Ps+p,\ Qs+q)$ with some coeffs 0, i.e. a restricted ARMA$(Ps+p, Qs+q)$

Note: $Y_t = \nabla^d \nabla_s^D X_t$

$Y_t \sim$ ARMA$(p,q)(P,Q)_s$

$$\Phi^{(s)}(B^s)\,\phi(B)\,Y_t = \Theta^{(s)}(B^s)\,\theta(B)\,\epsilon_t$$

i.e. $\Phi^{(s)}(B^s)\,\phi(B)\,(1-B)^d(1-B^s)^D X_t$

$= \Theta^{(s)}(B^s)\,\theta(B)\,\epsilon_t$

i.e. $\phi^*(B) X_t = \theta^*(B) \epsilon_t$

$$\phi^*(B) = \Phi^{(s)}(B^s)\,\phi(B)\,(1-B)^d (1-B^s)^D$$

$$\theta^*(B) = \Theta^{(s)}(B^s)\,\theta(B)$$

i.e. $X_t \sim \text{ARMA}(Ps + p + d + Ds,\ Qs + q)$

a restricted, non-stationary ARMA

Parameter estimation in time series models

Parameter estimation for AR models:

AR(p) $\quad X_t = c + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t$

Given an observed sample $(x_1, \ldots, x_n)$ from the above model, the problem is to estimate the model parameter $(c, \phi_1, \ldots, \phi_p)$

Approach I: Least squares estimation

Let $\Psi(c, \phi_1, \ldots, \phi_p) = \sum_{t=p+1}^{n} (x_t - c - \phi_1 x_{t-1} - \cdots - \phi_p x_{t-p})^2$

$\hat{c}, \hat{\phi}_1, \ldots, \hat{\phi}_p$ which minimizes $\Psi(c, \phi_1, \ldots, \phi_p)$ is the least squares estimates.

$$\frac{\partial \Psi}{\partial c} = 0 \Rightarrow \sum_{t=p+1}^{n} (x_t - c - \phi_1 x_{t-1} - \cdots - \phi_p x_{t-p}) = 0$$

$$\frac{\partial \Psi}{\partial \phi_i} = 0 \Rightarrow \sum_{t=p+1}^{n} (x_t - c - \phi_1 x_{t-1} - \cdots - \phi_p x_{t-p}) x_{t-i} = 0.$$

$i = 1(1)p$

This lead to the following linear system of equations

$$\sum_{p+1}^{n} x_t = (n-p)c + \phi_1 \sum_{p+1}^{n} x_{t-1} + \cdots + \phi_p \sum_{p+1}^{n} x_{t-p}$$

$$\sum x_t x_{t-1} = c \sum x_{t-1} + \phi_1 \sum x_{t-1}^2 + \cdots + \phi_p \sum x_{t-1} x_{t-p}$$

$$\sum x_t x_{t-2} = c \sum x_{t-2} + \phi_1 \sum x_{t-1} x_{t-2} + \cdots + \phi_p \sum x_{t-2} x_{t-p}.$$

$$\vdots$$

$$\sum x_t x_{t-p} = c \sum x_{t-p} + \phi_1 \sum x_{t-1} x_{t-p} + \cdots + \phi_p \sum x_{t-p}^2$$

Solving the above we get the least squares estimates of $(c, \phi_1, \ldots, \phi_p)$.

<u>Approach II</u> : Exact Maximum Likelihood estimation

Suppose $\{X_t\}$ is Gaussian AR(1)

$$X_t = c + \phi X_{t-1} + \epsilon_t \,; \quad \epsilon_t \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$$

Model parameters : c, ϕ

Noise parameter : σ^2

Observation set : $(x_1, \ldots, x_n)$

Note that,

$$E(X_1) = \frac{c}{1-\phi} = \mu \text{ (say)}$$

$$V(X_1) = \gamma_0 \text{ of AR(1)} = \frac{\sigma^2}{1-\phi^2}$$

$$X_1 \sim N\left(\frac{c}{1-\phi}, \frac{\sigma^2}{1-\phi^2}\right) \,; \quad X_2 \sim N\left(\frac{c}{1-\phi}, \frac{\sigma^2}{1-\phi^2}\right)$$

But these are not independent!

$$f_{X_1}(x_1;\theta) = \left(\sqrt{2\pi}\sqrt{\sigma^2/1-\phi^2}\right)^{-1} \exp\left(-\frac{1}{2}\frac{(x_1 - c/1-\phi)^2}{\sigma^2/1-\phi^2}\right)$$

$$X_2 = c + \phi X_1 + \epsilon_2$$

$$\Rightarrow \quad X_2 | X_1 \sim N(c + \phi x_1, \sigma^2)$$

$$f_{X_2|X_1} = \left(\sqrt{2\pi\sigma^2}\right)^{-1} \exp\left(-\frac{1}{2}\frac{(x_2 - c - \phi x_1)^2}{\sigma^2}\right)$$

$$X_3 = c + \phi X_2 + \epsilon_3$$

dist^n of $X_3 | X_2, X_1 \equiv$ dist^n of $X_3 | X_2 \sim N(c + \phi x_2, \sigma^2)$

In general, $\forall\, t \geq 2$

$$X_t | X_{t-1}, \ldots X_1 \equiv X_t | X_{t-1}$$
$$\sim N(c + \phi x_{t-1}, \sigma^2)$$

jt p.d.f. of $X_1, \ldots, X_n$

$$f_{X_1, \ldots X_n} = f_{X_n | X_{n-1}, \ldots, X_1} \; f_{X_{n-1}, \ldots X_1}$$

i.e. $f_{X_1, \ldots X_n} = f_{X_n | X_{n-1}, \ldots X_1} \left(f_{X_{n-1} | X_{n-2} \ldots X_1} f_{X_{n-2}, \ldots X_1} \right)$

split

$\vdots$

$$f_{X_1, \ldots X_n} = f_{X_n | X_{n-1}} f_{X_{n-1} | X_{n-2}} \cdots f_{X_2 | X_1} f_{X_1}$$
$$= f_{X_1} \prod_{t=2}^{n} f_{X_t | X_{t-1}}$$

Likelihood f^n

$$L(\underset{\sim}{\theta}) = f_{X_1}(x_1;\underset{\sim}{\theta}) \prod_{t=2}^{n} f_{X_t|X_{t-1}}(x_t;\underset{\sim}{\theta}|x_{t-1}).$$

Explicit form of log likelihood f^n

$$l(\underset{\sim}{\theta}) = \log f_{X_1}(x_1;\underset{\sim}{\theta}) + \sum_{t=2}^{n} \log f_{X_t|X_{t-1}}(x_t;\underset{\sim}{\theta}|x_{t-1})$$

$$l(\underset{\sim}{\theta}) = \left(-\frac{1}{2}\log 2\pi - \frac{1}{2}\log\frac{\sigma^2}{1-\phi^2} - \frac{1}{2}\frac{(x_1 - c/1-\phi)^2}{\sigma^2/1-\phi^2}\right)$$

$$\left(-\frac{n-1}{2}\log 2\pi - \frac{n-1}{2}\log\sigma^2 - \sum_{t=2}^{n}\frac{(x_t - c - \phi x_{t-1})^2}{2\sigma^2}\right). \quad (*')$$

$$\hat{\underset{\sim}{\theta}}_{EMLE} = \underset{\underset{\sim}{\theta}}{\arg\max}\ l(\underset{\sim}{\theta})$$

Note that the above EMLE does not have a closed form solution. Iterative search procedures are used for obtaining the values of EMLE (e.g. Newton-Raphson, Levenburg-Marquardt, Downhill simplex)

<u>Remark</u>: Alternate multivariate approach to derive the likelihood f^n.

Consider $(X_1, \ldots, X_n)$ as a random vector from an n-dimensional Gaussian ($\{X_t\}$ is a Gaussian process).

$$E(\underset{\sim}{X}) = \mu \underset{\sim}{1}_n \; ; \qquad \mu = \frac{c}{1-\phi}$$

$$\mathrm{Cov}(\underset{\sim}{X}) = \Sigma = \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{n-1} \\ & \ddots & & \vdots \\ & & \ddots & \gamma_1 \\ & & & \gamma_0 \end{pmatrix}$$

i.e. $$\Sigma = \frac{\sigma^2}{1-\phi^2} \begin{pmatrix} 1 & \phi & \phi^2 & \cdots & \phi^{n-1} \\ & 1 & \phi & \cdots & \phi^{n-2} \\ & & \ddots & & \vdots \\ & & & \ddots & \phi \\ & & & & 1 \end{pmatrix}$$

$$\underset{\sim}{X} \sim N_n(\mu \underset{\sim}{1}_n, \Sigma)$$

Likelihood f^n

$$L(\underset{\sim}{\theta}) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\underset{\sim}{x} - \mu \underset{\sim}{1}_n)' \Sigma^{-1} (\underset{\sim}{x} - \mu \underset{\sim}{1}_n)\right)$$

log likelihood f^n

$$l(\underset{\sim}{\theta}) = -\frac{n}{2}\log(2\pi) + \frac{1}{2}\log(|\Sigma^{-1}|)$$
$$-\frac{1}{2}(\underset{\sim}{x} - \mu \underset{\sim}{1}_n)' \Sigma^{-1} (\underset{\sim}{x} - \mu \underset{\sim}{1}_n)$$

Approach iii : Conditional MLE approach

Regard the value of X_1 (observed as x_1) as deterministic and maximize the likelihood conditional on the first observation

i.e. $L_c(\underset{\sim}{\theta}) = f_{X_n, \ldots X_2 | X_1}(x_n, \ldots x_2 | x_1)$

Realize that

$$f_{X_n, \ldots X_2 | X_1} = f_{X_n | X_{n-1} \ldots X_1} \, f_{X_{n-1}, \ldots X_2 | X_1}$$

i.e. $f_{X_n, \ldots X_2 | X_1} = \left(f_{X_n | X_{n-1}}\right) \left(f_{X_{n-1} | X_{n-2} \ldots X_1} \, f_{X_{n-2} \ldots X_2 | X_1}\right)$

$$f_{X_n \ldots X_2 | X_1} = f_{X_n | X_{n-1}} \, f_{X_{n-1} | X_{n-2}} \cdots f_{X_2 | X_1}$$

$$f_{X_n \ldots X_2 | X_1} = \prod_{t=2}^{n} f_{X_t | X_{t-1}}(x_t ; \underset{\sim}{\theta} | x_{t-1})$$

i.e $L_c(\underset{\sim}{\theta}) = \prod_{t=2}^{n} f_{X_t | X_{t-1}}(x_t ; \underset{\sim}{\theta} | x_{t-1})$

Now $X_t | X_{t-1} \sim N_1(c + \phi x_{t-1}, \sigma^2)$

$\forall \, t \geq 2$

Conditional log likelihood

$$l_c(\underset{\sim}{\theta}) = -\frac{n-1}{2}\log 2\pi - \frac{n-1}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_2^n\left(x_t - c - \phi x_{t-1}\right)^2$$

$$\hat{\underset{\sim}{\theta}}_{CMLE} = \underset{\underset{\sim}{\theta}}{\arg\max}\ l_c(\underset{\sim}{\theta})$$

Realize that maximization of $l_c(\underset{\sim}{\theta})$ w.r.t. c & ϕ is equivalent to minimization of

$$\sum_{t=2}^n\left(x_t - c - \phi x_{t-1}\right)^2$$

i.e. $\Rightarrow$ CMLEs of c & ϕ are the ordinary LSE (that was described in Approach I).

i.e. CMLEs of c & ϕ are obtained as solutions of

$$\left.\begin{aligned}\sum_2^n x_t &= c(n-1) + \phi\sum_2^n x_{t-1}\\ \sum_2^n x_t x_{t-1} &= c\sum_2^n x_{t-1} + \phi\sum_2^n x_{t-1}^2\end{aligned}\right\}.$$

$$\begin{pmatrix}\hat{c}\\ \hat{\phi}\end{pmatrix} = \begin{pmatrix} n-1 & \sum_2^n x_t \\ \sum_2^n x_t & \sum x_{t-1}^2\end{pmatrix}^{-1}\begin{pmatrix}\sum_2 x_t \\ \sum_2^n x_t x_{t-1}\end{pmatrix}$$

Further,

$$\hat{\sigma}^2_{CMLE} = \frac{1}{n-1} \sum_{2}^{n} \left(x_t - \hat{c} - \hat{\phi} x_{t-1}\right)^2$$

<u>Note</u> : Unlike EMLE, we get closed form solution of CMLE

<u>Remark</u> : CMLE & EMLE has the same asymptotic distribution (provided $|\phi|<1$) i.e. stationary

MLE for Gaussian AR(p)

Exact MLE formulation:

$$X_t = c + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t$$

$$\epsilon_t \overset{i.i.d.}{\sim} N(0, \sigma^2)$$

$$\underset{\sim}{\theta} = (c, \phi_1, \ldots, \phi_p, \sigma^2)'$$

Note first that $\underset{\sim}{X}_p = (X_1, \ldots, X_p)' \sim N_p$

$$E \underset{\sim}{X}_p = \mu \underset{\sim}{1}_p \quad ; \quad \mu = c/(1-\phi_1-\cdots-\phi_p)$$

Let $\sigma^2 V_p$ denote the covariance matrix of $\underset{\sim}{X}_p$

i.e. $\operatorname{cov}(\underset{\sim}{X}_p) = \sigma^2 V_p = \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{p-1} \\ & \ddots & & \vdots \\ & & \ddots & \gamma_1 \\ & & & \gamma_0 \end{pmatrix}$

$$\left(\text{for } p=1, \; V_p = (1-\phi^2)^{-1}\right)$$

$$\underset{\sim}{X}_p \sim N_p(\mu \underset{\sim}{1}_p, \sigma^2 V_p)$$

$$f_{X_p, X_{p-1}, \ldots X_1} = (2\pi)^{-p/2} (\sigma^2)^{-p/2} |V_p|^{-1/2} \exp\left(-\frac{1}{2\sigma^2} (\underset{\sim}{x}_p - \mu \underset{\sim}{1}_p)' V_p^{-1} (\underset{\sim}{x}_p - \mu \underset{\sim}{1}_p)\right)$$

For $\forall\, t > p$; $X_t | X_{t-1}, \ldots X_1 \sim N(c + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p}, \sigma^2)$

$$\& \; X_t | X_{t-1}, \ldots, X_1 \equiv X_t | X_{t-1}, \ldots X_{t-p}$$

$\forall\, t > p$

$$f_{X_t | X_{t-1}, \ldots, X_{t-p}} = (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2} \Big(x_t - c - \sum_1^p \phi_i x_{t-i}\Big)^2\right)$$

jt distⁿ of $X_n, X_{n-1}, \ldots X_1$

$$f_{X_n,\ldots X_1} = f_{X_n|X_{n-1},\ldots,X_1} \, f_{X_{n-1},\ldots X_1}$$

$$= f_{X_n|X_{n-1}\ldots X_{n-p}} , f_{X_{n-1}|X_{n-2},\ldots X_1} \, f_{X_{n-2},\ldots X_1}$$

$$\vdots \qquad\qquad \vdots$$

$$f_{X_n,\ldots X_1} = f_{X_n|X_{n-1},\ldots X_{n-p}} \, f_{X_{n-1}|X_{n-2}\ldots n-1-p} \cdots$$

$$\cdots f_{X_{p+1}|X_p,\ldots X_1} \, f_{X_p, X_{p-1},\ldots X_1}$$

$$= f_{X_p,\ldots X_1} \prod_{t=p+1}^{n} f_{X_t|X_{t-1},\ldots X_{t-p}}.$$

Likelihood f^n

$$L(\underset{\sim}{\theta}) = f_{\underset{\sim}{X}_p}(\underset{\sim}{x}_p ; \underset{\sim}{\theta}) \prod_{t=p+1}^{n} f_{X_t|X_{t-1},\ldots,X_{t-p}}(x_t ; \underset{\sim}{\theta} | x_{t-1},\ldots x_{t-p})$$

log likelihood f^n

$$l(\underset{\sim}{\theta}) = \Big(-\frac{p}{2}\log 2\pi - \frac{p}{2}\log\sigma^2 + \frac{1}{2}\log|V_p^{-1}|$$

$$- \frac{1}{2\sigma^2}(\underset{\sim}{x}_p - \mu\underset{\sim}{1}_p)' V_p^{-1}(\underset{\sim}{x}_p - \mu\underset{\sim}{1}_p)\Big)$$

$$+ \Big(-\frac{n-p}{2}\log 2\pi - \frac{n-p}{2}\log\sigma^2$$

$$- \frac{1}{2\sigma^2}\sum_{t=p+1}^{n}\big(x_t - c - \sum_{i=1}^{p}\phi_i x_{t-i}\big)^2\Big)$$

$$\hat{\underset{\sim}{\theta}}_{EMLE} = \underset{\underset{\sim}{\theta}}{\arg\max}\; l(\underset{\sim}{\theta})$$

<u>Note</u>: No closed form solution of $\hat{\underset{\sim}{\theta}}_{EMLE}$

<u>Note</u>: Iterative methods applied to obtain exact maximum likelihood estimates.

Note: For a general p, $(i,j)^{th}$ element of V_p^{-1}, $v^{ij}(p)$ is given by (result is due to Galbraith, 1974, Jr of Applied Probability paper):

$$v^{ij}(p) = \left(\sum_{k=0}^{i-1} \phi_k \phi_{k+j-i} - \sum_{k=p+1-j}^{p+i-j} \phi_k \phi_{k+j-i} \right)$$

$1 \leq i \leq j \leq p$ with $\phi_0 = -1$ above

The above is called the Galbraith's formula and can be used to write $L(\underset{\sim}{\theta})$ or $l(\underset{\sim}{\theta})$ explicitly in terms of ϕ_i s.

Conditional MLE formulation : AR(p)

Regard the values of the first p observations as deterministic and maximize the likelihood conditional on the first p observations.

Now, realize that the joint conditional p.d.f of $X_n, \ldots, X_{p+1}$ given $X_p, \ldots X_1$ is given by:

$$f_{X_n, \ldots X_{p+1} | X_p, \ldots X_1} = f_{X_n | X_{n-1} \ldots X_1} f_{X_{n-1} | X_{n-2}, \ldots X_1} \cdots f_{X_{p+1} | X_p, \ldots, X_1}$$

$$= f_{X_n | X_{n-1}, \ldots X_{n-p}} f_{X_{n-1} | X_{n-2}, \ldots X_{n-1-p}} \cdots f_{X_{p+1} | X_p, \ldots, X_1}$$

i.e. $f_{X_n,\dots X_{p+1}|X_p,\dots,X_1} = \prod_{t=p+1}^{n} f_{X_t|X_{t-1},\dots,X_{t-p}}$

$\forall\, t > p$, we have

$$X_t|X_{t-1},\dots,X_{t-p} \equiv X_t|X_{t-1},\dots X_1$$

$$\sim N\left(c+\phi_1 x_{t-1}+\dots+\phi_p x_{t-p}, \sigma^2\right)$$

Hence conditional log likelihood

$$l_c(\underset{\sim}{\theta}) = -\frac{n-p}{2}\log 2\pi - \frac{n-p}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=p+1}^{n}\left(x_t - c - \sum_{i=1}^{p}\phi_i x_{t-i}\right)^2 \quad — *$$

Note that maximization of $(*)$ w.r.t. $(c, \phi_1,\dots,\phi_p)$ is equivalent to ~~maximization~~ ~~of~~ minimization of

$$\sum_{t=p+1}^{n}\left(x_t - c - \sum_{i=1}^{p}\phi_i x_{t-i}\right)^2 \quad \text{w.r.t } (c,\phi_1,\dots\phi_p)$$

$\Rightarrow$ CMLEs of $(c,\phi_1,\dots,\phi_p)$ are same as the OLS estimates.

<u>Note</u>: CMLE & EMLE have the same asymptotic distⁿ.

(21)

<u>Remark</u>: Estimation of AR parameters using Yule-Walker eq^n

$$X_t = c + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t$$

Yule-Walker eq^n

$$\gamma_h = \phi_1 \gamma_{h-1} + \phi_2 \gamma_{h-2} + \cdots + \phi_p \gamma_{h-p} \quad ; h > 0$$

Using p Yule-Walker eq^{n}s and estimated $\hat{\gamma}_i$s one can obtain Y-W eq^n based estimates of AR model parameters.

e.g. consider an AR(3) setup

$$\hat{\gamma}_1 = \phi_1 \hat{\gamma}_0 + \phi_2 \hat{\gamma}_1 + \phi_3 \hat{\gamma}_2$$

$$\hat{\gamma}_2 = \phi_1 \hat{\gamma}_1 + \phi_2 \hat{\gamma}_0 + \phi_3 \hat{\gamma}_1$$

$$\hat{\gamma}_3 = \phi_1 \hat{\gamma}_2 + \phi_2 \hat{\gamma}_1 + \phi_3 \hat{\gamma}_0$$

i.e. $$\begin{pmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \\ \hat{\phi}_3 \end{pmatrix} = \begin{pmatrix} \hat{\gamma}_0 & \hat{\gamma}_1 & \hat{\gamma}_2 \\ & \hat{\gamma}_0 & \hat{\gamma}_1 \\ & & \hat{\gamma}_0 \end{pmatrix}^{-1} \begin{pmatrix} \hat{\gamma}_1 \\ \hat{\gamma}_2 \\ \hat{\gamma}_3 \end{pmatrix}.$$

Maximum likelihood estimation for MA models

MA(1): Conditional MLE formulation

$$X_t = \mu + \epsilon_t + \theta \epsilon_{t-1}; \quad \epsilon_t \overset{i.i.d.}{\sim} N(0, \sigma^2)$$

$$\underset{\sim}{\eta} = (\mu, \theta, \sigma^2)' \leftarrow \text{parameter vector}$$

Note that $X_t | \epsilon_{t-1} \sim N(\mu + \theta \epsilon_{t-1}, \sigma^2)$

Suppose, we assume that $\epsilon_0 = 0$ (it's expected value) is given, then

$$X_1 | \epsilon_0 \sim N(\mu, \sigma^2)$$

$\epsilon_1 = X_1 - \mu - \theta \epsilon_0$; hence ϵ_1 given $X_1 = x_1$ & $\epsilon_0 = 0$ is $\epsilon_1 = x_1 - \mu$

$$X_2 | X_1, \epsilon_0 = 0 \sim N(\mu + \theta \epsilon_1, \sigma^2)$$
$$\text{i.e. } N(\mu + \theta(x - \mu), \sigma^2)$$

$$X_3 | X_2, X_1, \epsilon_0 \sim N(\mu + \theta \epsilon_2, \sigma^2)$$

ϵ_2 given $X_2, X_1, \epsilon_0 = 0$ is $\epsilon_2 = x_2 - \mu - \theta \epsilon_1$

$$\text{i.e. } \epsilon_2 = x_2 - \mu - \theta(x_1 - \mu)$$

i.e. $X_3 | X_2, X_1, \epsilon_0 \sim N\left(\mu + \theta\left((x_2 - \mu) - \theta(x_1 - \mu)\right), \sigma^2\right)$

Thus given $\epsilon_0 = 0$, the full sequence $\epsilon_1, \ldots \epsilon_n$ can be expressed in terms of $(x_1, \ldots, x_n)$, μ & θ through the relationship

$$\epsilon_t = x_t - \mu - \theta \epsilon_{t-1}$$
$$\epsilon_{t-1} = x_{t-1} - \mu - \theta \epsilon_{t-2} \quad - - - .$$

(221)

$\forall\, t \geq 2;\ X_t \mid X_{t-1}, \ldots X_1, \epsilon_0 = 0 \equiv X_t \mid \epsilon_{t-1} \sim N(\mu + \theta \epsilon_{t-1}, \sigma^2)$

Conditional likelihood f^n, conditioned on $\epsilon_0 = 0$, is

$$L(\underset{\sim}{\eta}) = f_{X_n, \ldots, X_1 \mid \epsilon_0 = 0}(x_n, \ldots, x_1; \underset{\sim}{\eta} \mid \epsilon_0 = 0)$$

$$= f_{X_n \mid X_{n-1}, \ldots X_1, \epsilon_0 = 0}\, f_{X_{n-1}, \ldots, X_1 \mid \epsilon_0 = 0}$$

$$= f_{X_n \mid X_{n-1} \ldots X_1, \epsilon_0 = 0}\, f_{X_{n-1} \mid X_{n-2}, \ldots X_1, \epsilon_0 = 0}\, f_{X_{n-2} \ldots X_1 \mid \epsilon_0 = 0}$$

$$= f_{X_1 \mid \epsilon_0 = 0} \prod_{t=2}^{n} f_{X_t \mid X_{t-1}, \ldots, X_1, \epsilon_0 = 0}$$

$$= f_{X_1 \mid \epsilon_0 = 0} \prod_{t=2}^{n} f_{X_t \mid \epsilon_{t-1}}$$

Conditional likelihood f^n, conditioned on $\epsilon_0 = 0$, is

$$L(\underset{\sim}{\eta}) = f_{X_n,\ldots,X_1 | \epsilon_0=0}(x_n,\ldots,x_1; \underset{\sim}{\eta} | \epsilon_0 = 0)$$

$$= f_{X_n | X_{n-1},\ldots,X_1, \epsilon_0=0} \, f_{X_{n-1},\ldots,X_1 | \epsilon_0=0}$$

$$= f_{X_n | X_{n-1},\ldots,X_1, \epsilon_0=0} \, f_{X_{n-1} | X_{n-2},\ldots,X_1, \epsilon_0=0} \, f_{X_{n-2},\ldots,X_1 | \epsilon_0=0}$$

$$= f_{X_1 | \epsilon_0=0} \prod_{t=2}^{n} f_{X_t | X_{t-1},\ldots,X_1, \epsilon_0=0}$$

$$= f_{X_1 | \epsilon_0=0} \prod_{t=2}^{n} f_{X_t | \epsilon_{t-1}}$$

Thus, the conditional log likelihood f^n is

$$l(\underset{\sim}{\eta}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_1^n (x_t - \mu - \theta\epsilon_{t-1})^2$$

$$l(\underset{\sim}{\eta}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_1^n \epsilon_t^2$$

Note that
$$\epsilon_t = (x_t - \mu) - \theta\epsilon_{t-1}$$

$$\epsilon_t = (x_t - \mu) - \theta\left((x_{t-1} - \mu) - \theta\epsilon_{t-2}\right)$$

$$= (x_t - \mu) - \theta(x_{t-1} - \mu) + \theta^2\epsilon_{t-2}$$

$$= (x_t - \mu) - \theta(x_{t-1} - \mu) + \theta^2(x_{t-2} - \mu - \theta\epsilon_{t-3})$$

$$= (x_t - \mu) - \theta(x_{t-1} - \mu) + \theta^2(x_{t-2} - \mu)$$
$$- \theta^3\epsilon_{t-3}$$

$$\vdots$$

$$\epsilon_t = (x_t - \mu) - \theta(x_{t-1} - \mu) + \theta^2 (x_{t-2} - \mu) - \cdots$$
$$\cdots + (-\theta)^{t-1}(x_1 - \mu) + (-\theta)^t \epsilon_0$$

If $\epsilon_0 = 0$ at the initialization,

$$\epsilon_t = \sum_{i=1}^{t} (x_i - \mu)(-\theta)^{t-i} \quad \text{and } l(\underset{\sim}{\eta}) \text{ is}$$

$$l(\underset{\sim}{\eta}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log \sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{n}\left(\sum_{i=1}^{t}(x_i - \mu)(-\theta)^{t-i}\right)^2$$

$$\underset{\sim}{\hat{\eta}}_{CMLE} = \arg\max_{\underset{\sim}{\eta}} l(\underset{\sim}{\eta})$$

CMLEs are obtained using some iterative procedure.

<u>Exact MLE formulation</u>

Let's look at the multivariate formulation.

$\underset{\sim}{X} = (X_1, \ldots, X_n)'$ realization from an n-dimensional multivariate normal

i.e. $\underset{\sim}{X} \sim N_n(\mu \underset{\sim}{1}_n, \Omega)$

$$\Omega = E(\underset{\sim}{X} - \mu\underset{\sim}{1}_n)(\underset{\sim}{X} - \mu\underset{\sim}{1}_n)'$$

$$= \sigma^2 \begin{pmatrix} (1+\theta^2) & \theta & 0 & 0 & \cdots & 0 \\ \theta & (1+\theta^2) & \theta & 0 & \cdots & 0 \\ & \ddots & \ddots & \ddots & & \\ & & & & \theta & \theta \\ 0 & & & \cdots & \theta & (1+\theta^2) \end{pmatrix}.$$

(c)

Likelihood f^n

$$L(\underset{\sim}{\eta}) = (2\pi)^{-n/2} |\Omega|^{-1/2} \exp\left(-\frac{1}{2}(\underset{\sim}{x} - \mu \underset{\sim}{1}_n)' \Omega^{-1} (\underset{\sim}{x} - \mu \underset{\sim}{1}_n)\right)$$

log likelihood f^n

$$l(\underset{\sim}{\eta}) = -\frac{n}{2}\log 2\pi - \frac{1}{2}\log|\Omega| - \frac{1}{2}(\underset{\sim}{x} - \mu \underset{\sim}{1}_n)' \Omega^{-1} (\underset{\sim}{x} - \mu \underset{\sim}{1}_n)$$

Consider a factorization of Ω as

$$\Omega = A D A' \quad \text{---} \ (*)$$

$$A = \begin{pmatrix} 1 & & & & \\ \frac{\theta}{1+\theta^2} & 1 & & \bigcirc & \\ 0 & \frac{\theta(1+\theta^2)}{1+\theta^2+\theta^4} & 1 & & \\ & & \ddots & \ddots & \\ & \bigcirc & & \underset{\uparrow}{\_\_} & 1 \end{pmatrix}$$

$$\left(\frac{\theta(1+\theta^2+\cdots+\theta^{2(n-2)})}{1+\theta^2+\cdots+\theta^{2(n-1)}}\right)$$

$$D = \sigma^2 \operatorname{diag}\left(1+\theta^2, \frac{1+\theta^2+\theta^4}{1+\theta^2}, \frac{1+\theta^2+\theta^4+\theta^6}{1+\theta^2+\theta^4}, \dots\dots, \frac{1+\theta^2+\theta^4+\cdots+\theta^{2n}}{1+\theta^2+\cdots+\theta^{2(n-1)}}\right)$$

Using $(*)$ likelihood f^n is

$$L(\underset{\sim}{\eta}) = (2\pi)^{-n/2} |ADA'|^{-1/2} \exp\left(-\frac{1}{2}(\underset{\sim}{x} - \mu \underset{\sim}{1}_n)'(A')^{-1} D^{-1} A^{-1} (\underset{\sim}{x} - \mu \underset{\sim}{1}_n)\right)$$

Note that $|\Omega| = |ADA'| = |A||D||A'| = |D| = \prod_{t=1}^{n} d_{tt}$

$$d_{tt} = \sigma^2 \frac{1+\theta^2+\cdots+\theta^{2t}}{1+\theta^2+\cdots+\theta^{2(t-1)}}$$

Let $\quad \underset{\sim}{x}^0 = A^{-1}\left(\underset{\sim}{x} - \mu \underset{\sim}{1}_n\right)$

i.e. $A \underset{\sim}{x}_0 = \underset{\sim}{x} - \underset{\sim}{\mu}$

$$x_1^0 = x_1 - \mu$$

$$\vdots$$

$$x_t^0 = (x_t - \mu) - \frac{\theta(1+\theta^2+\cdots+\theta^{2(t-2)})}{1+\theta^2+\cdots+\theta^{2(t-1)}} x_{t-1}^0$$

$$L(\underset{\sim}{\eta}) = (2\pi)^{-n/2}\left(\prod_1^n d_{tt}\right)^{-1/2} \exp\left(-\frac{1}{2}\sum_{t=1}^n x_t^{0^2}/d_{tt}\right)$$

$$\hat{\eta}_{EMLE} = \underset{\underset{\sim}{\eta}}{\arg\max}\ l(\underset{\sim}{\eta}) \quad ; \ l(\underset{\sim}{\eta}) \text{ is the log likelihood}$$

Once again iterative optimization techniques are used to obtain EMLE using the data $(x_1, \ldots, x_n)$

MLE of Gaussian MA(q)

$$X_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}$$

$$\epsilon_t \overset{i.i.d.}{\sim} N(0, \sigma^2)$$

$$\underset{\sim}{\eta} = (\mu, \theta_1, \ldots, \theta_q, \sigma^2)' \text{ - parameter vector}$$

Conditional MLE formulation

Consider the likelihood conditional on the assumption that

$$\epsilon_0 = 0, \ \epsilon_{-1} = \epsilon_{-2} = \cdots = \epsilon_{-(q-1)} = 0$$

(at the expected value of $\epsilon_0, \epsilon_{-1}, \ldots, \epsilon_{-(q-1)}$)

Conditional likelihood (conditioned on $\underset{\sim}{\epsilon_0} = (\epsilon_0, \epsilon_{-1}, \ldots, \epsilon_{-(q-1)})' = \underset{\sim}{0}$

$$L_c(\underset{\sim}{\eta}) = f_{X_n, \ldots, X_1 | \underset{\sim}{\epsilon_0} = \underset{\sim}{0}} (x_n, \ldots, x_1; \underset{\sim}{\eta} | \underset{\sim}{\epsilon_0} = \underset{\sim}{0})$$

$$= f_{X_n | X_{n-1}, \ldots X_1, \underset{\sim}{\epsilon_0} = \underset{\sim}{0}} \; f_{X_{n-1}, \ldots X_1 | \underset{\sim}{\epsilon_0} = \underset{\sim}{0}}$$

$$\vdots$$

$$= f_{X_1 | \underset{\sim}{\epsilon_0} = \underset{\sim}{0}} \left(\prod_{t=2}^{n} f_{X_t | X_{t-1}, \ldots, X_1, \underset{\sim}{\epsilon_0} = \underset{\sim}{0}} \right)$$

Note that $X_1 | \underset{\sim}{\epsilon_0} = \underset{\sim}{0} \sim N(\mu, \sigma^2)$

ϵ_t given $x_t, x_{t-1}, \ldots \underset{\sim}{\epsilon_0}$ can be expressed as

$$\epsilon_t = x_t - \mu - \theta_1 \epsilon_{t-1} - \cdots - \theta_q \epsilon_{t-q}$$

$\forall t \geq 2$

$$X_t | X_{t-1}, \ldots, X_1, \underset{\sim}{\epsilon_0} = \underset{\sim}{0} \sim N(\mu + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}, \sigma^2)$$

Conditional log likelihood

$$l_c(\underset{\sim}{\eta}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{n}\left(x_t - \mu - \theta_1\epsilon_{t-1} - \cdots - \theta_q\epsilon_{t-q}\right)^2$$

$$l_c(\underset{\sim}{\eta}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{n}\epsilon_t^2$$

$$\hat{\underset{\sim}{\eta}}_{CMLE} = \arg\max_{\underset{\sim}{\eta}} l_c(\underset{\sim}{\eta})$$

Iterative methods used to obtain CMLEs. Newton Raphson

Exact MLE formulation

$\underset{\sim}{X} = (X_1, \ldots, X_n)'$ — $\{X_t\}$ is Gaussian MA(q)

$\Rightarrow \underset{\sim}{X} \sim N_n(\underset{\sim}{\mu}, \Omega)$

$\underset{\sim}{\mu} = E(\underset{\sim}{X}) = \mu \underset{\sim}{1}_n$

$$\Omega = Cov(\underset{\sim}{X}) = E(\underset{\sim}{X} - \underset{\sim}{\mu})(\underset{\sim}{X} - \underset{\sim}{\mu})'$$

$$= \begin{pmatrix} \gamma_0 & \gamma_1 & & \gamma_q & & 0 \\ \gamma_1 & \gamma_0 & \gamma_1 & \ddots & \ddots & \\ & \ddots & \ddots & \ddots & & \gamma_q \\ \gamma_q & & \ddots & \ddots & \ddots & \\ & \ddots & & \ddots & \ddots & \gamma_1 \\ 0 & & \gamma_q & & \gamma_1 & \gamma_0 \end{pmatrix}$$

$$\gamma_K = \begin{cases} \sigma^2(\theta_K + \theta_{K+1}\theta_1 + \cdots + \theta_q\theta_{q-K}), & K = 0, \ldots, q \\ 0, & o/w. \end{cases}$$

Exact log likelihood f^n

$$l(\underset{\sim}{\eta}) = -\frac{n}{2}\log 2\pi - \frac{1}{2}\log|\Omega| - \frac{1}{2}(\underset{\sim}{x} - \underset{\sim}{\mu})'\Omega^{-1}(\underset{\sim}{x} - \underset{\sim}{\mu}).$$

$$\hat{\underset{\sim}{\eta}}_{EMLE} = \arg\max_{\underset{\sim}{\eta}} l(\underset{\sim}{\eta})$$

<u>Note</u>: Similar to MA(1), decomposition of Ω as $\Omega = ADA'$ leads to simplification of $l(\underset{\sim}{\eta})$, expressing it explicitly in terms of the parameters $\mu, \theta_1, \ldots, \theta_q$ & σ^2.

<u>Note</u>: Iterative optimization techniques are used to obtain the EMLEs using $l(\underset{\sim}{\eta})$.

<u>Remark</u>: Conditional LSE for MA model

Consider an invertible MA(1) $(|\theta| < 1)$

$$X_t = \mu + \epsilon_t + \theta\epsilon_{t-1} \; ; \; \underset{\sim}{\eta} = (\mu, \theta)'$$

Ordinary LSE of μ, θ which is defined as

$$\hat{\eta}_{OLS} = \arg\min_{\underset{\sim}{\eta}} \underbrace{\sum_{t=2}^{n} (x_t - \mu - \theta\epsilon_{t-1})^2}_{\sum_{t=2}^{n} \epsilon_t^2}$$

is not feasible since ϵ_{t-1} is not observable

Let us put a condition that ϵ_0 is given (e.g $\epsilon_0 = 0$ It's expected value)

$$x_1 = \mu + \epsilon_1 + \theta\epsilon_0 \; ; \quad \epsilon_1 = (x_1 - \mu) - \theta\epsilon_0$$

$$x_2 = \mu + \epsilon_2 + \theta\epsilon_1 \; ; \quad \epsilon_2 = (x_2 - \mu) - \theta\epsilon_1$$

$$\epsilon_2 = (x_2 - \mu) - \theta\big((x_1 - \mu) - \theta\epsilon_0\big)$$

i.e. $\epsilon_2 = (x_2 - \mu) - \theta(x_1 - \mu) + \theta^2\epsilon_0$

$\vdots$

$$x_{t-1} = \mu + \epsilon_{t-1} + \theta\epsilon_{t-2}$$

$$\epsilon_{t-1} = (x_{t-1} - \mu) - \theta\epsilon_{t-2}$$

$$= (x_{t-1} - \mu) - \theta\big((x_{t-2} - \mu) - \theta\epsilon_{t-3}\big)$$

$\vdots$

$$\epsilon_{t-1} = \sum_{i=1}^{t-1} (x_i - \mu)(-\theta)^{t-1-i} \qquad (\epsilon_0 = 0).$$

Conditional LSE of μ & θ is defined as

$$\hat{\underset{\sim}{\eta}}_{CLSE} = \underset{\underset{\sim}{\eta}}{\arg\min} \sum_{t=2}^{n} \left(x_t - \mu - \theta\sum_{i=1}^{t-1}(x_i - \mu)(-\theta)^{t-1-i}\right)^2$$

<u>Remark</u>: CLSE for MA(q) can be framed in a similar manner.

MLE for Gaussian ARMA(p, q)

$$X_t = c + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}$$

$$\epsilon_t \overset{i.i.d.}{\sim} N(0, \sigma^2)$$

$\underset{\sim}{\eta} = (c, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)'$ - model parameters

σ^2 - noise parameter

Conditional MLE formulation

Consider the likelihood, conditioned on p initial values of x and q initial values of ϵ

$$\underset{\sim}{x_0} = (x_0, x_{-1}, x_{-2}, \ldots, x_{-(p-1)})'$$

and $$\underset{\sim}{\epsilon_0} = (\epsilon_0, \epsilon_{-1}, \ldots, \epsilon_{-(q-1)})'$$

Given $\underset{\sim}{x_0}$ and $\underset{\sim}{\epsilon_0}$ and using $\epsilon_t = x_t - c - \sum_{i=1}^{p} \phi_i x_{t-i} - \sum_{i=1}^{q} \theta_i \epsilon_{t-i}$, we can write $\epsilon_1, \ldots, \epsilon_n$ in terms of x_t s, $c, \phi_1, \ldots \phi_p, \theta_1, \ldots, \theta_q$

Note that

$$X_1 \mid \underset{\sim}{x_0}, \underset{\sim}{\epsilon_0} \sim N\left(c + \sum_{i=1}^{p} \phi_i x_{1-i} + \sum_{i=1}^{q} \theta_i \epsilon_{1-i}, \sigma^2\right)$$

In general,

$\forall\, t \geq 2$

$$X_t \mid X_{t-1}, \ldots, X_1, \underset{\sim}{X_0}, \underset{\sim}{\epsilon_0} \equiv X_t \mid X_{t-1}, \ldots, X_{t-p}, \epsilon_{t-1}, \ldots, \epsilon_{t-q}$$

$$\sim N\left(c + \sum_{i=1}^{p} \phi_i x_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i},\ \sigma^2\right)$$

Conditional likelihood f^n is given by

$$L_c(\underset{\sim}{\eta}) = f_{X_n, \ldots, X_1 \mid \underset{\sim}{X_0}, \underset{\sim}{\epsilon_0}}(x_n, \ldots, x_1; \underset{\sim}{\eta} \mid \underset{\sim}{x_0}, \underset{\sim}{\epsilon_0})$$

$$= f_{X_n \mid X_{n-1}, \ldots X_1, \underset{\sim}{X_0}, \underset{\sim}{\epsilon_0}}\ f_{X_{n-1}, \ldots X_1 \mid \underset{\sim}{X_0}, \underset{\sim}{\epsilon_0}}$$

$$\vdots$$

$$L_c(\underset{\sim}{\eta}) = f_{X_1 \mid \underset{\sim}{X_0}, \underset{\sim}{\epsilon_0}} \prod_{t=2}^{n} f_{X_t \mid X_{t-1}, \ldots X_1, \underset{\sim}{X_0}, \underset{\sim}{\epsilon_0}}$$

Conditional log likelihood f^n is given by

$$l_c(\underset{\sim}{\eta}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{n}\left(x_t - c - \sum_{i=1}^{p}\phi_i x_{t-i} - \sum_{i=1}^{q}\theta_i \epsilon_{t-i}\right)^2$$

i.e. $l_c(\underset{\sim}{\eta}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{n}\epsilon_t^2$

$$\hat{\underset{\sim}{\eta}}_{CMLE} = \underset{\underset{\sim}{\eta}}{\arg\max}\ l_c(\underset{\sim}{\eta}).$$

Note: Choice of initial values in CMLE formulation

Option I: Set initial xs and ϵs equal to their expected values

i.e. $x_s = \dfrac{c}{1-\phi_1-\cdots-\phi_p}$; $s = 0, -1, \ldots, -(p-1)$

& $\epsilon_s = 0$; $s = 0, -1, \ldots, -(q-1)$

and use $\epsilon_t = x_t - c - \sum \phi_i x_{t-i} - \sum \theta_i \epsilon_{t-i}$

for $t = 1, \ldots, n$ to write $\epsilon_1, \ldots, \epsilon_n$

in terms of xs and $c, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q$

Option II: Set ϵs at their expected values & set xs at their actual values

i.e. start with $(x_1, \ldots, x_p)$ as initial set of xs

and set $\epsilon_p = \epsilon_{p-1} = \cdots = \epsilon_{p-(q-1)} = 0$

& $\epsilon_t = x_t - c - \sum_{i=1}^{p} \phi_i x_{t-i} - \sum_{i=1}^{q} \theta_i \epsilon_{t-i}$

$t = p+1, p+2, \ldots, n$

Note that under this option the conditional log likelihood changes to

$$l_c(\underset{\sim}{\eta}) = -\frac{n-p}{2}\log 2\pi - \frac{n-p}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=p+1}^{n}\epsilon_t^2$$

Option iii : Set ϵs at the expected values

$$\epsilon_0 = \epsilon_{-1} = \cdots = \epsilon_{-(q-1)} = 0$$

and set xs at their "backforecasted" values

Back forcasting is a techniques for forcasting in backward direction.

Large Sample asymptotic distⁿ of MLE

Let $\{X_t\}$ be a Causal and invertible ARMA(p,q)

$$\phi(B) X_t = \theta(B) \epsilon_t$$

$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$$

$$\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$$

$$\underset{\sim}{\beta} = (\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)'$$

<u>Asymptotic distⁿ result:</u>

$$\sqrt{n}\left(\hat{\underset{\sim}{\beta}}_{MLE} - \underset{\sim}{\beta}\right) \xrightarrow{\mathcal{L}} N_{p+q}\left(\underset{\sim}{0}, V(\underset{\sim}{\beta})\right)$$

$$V(\underset{\sim}{\beta}) = \sigma^2 \begin{bmatrix} E\,\underset{\sim}{U}_t \underset{\sim}{U}_t' & E\,\underset{\sim}{U}_t \underset{\sim}{V}_t' \\ E\,\underset{\sim}{V}_t \underset{\sim}{U}_t' & E\,\underset{\sim}{V}_t \underset{\sim}{V}_t' \end{bmatrix}^{-1}$$

where, $\underset{\sim}{U}_t = (U_t, U_{t-1}, \ldots, U_{t-(p-1)})'$

$\underset{\sim}{V}_t = (V_t, V_{t-1}, \ldots, V_{t-(q-1)})'$

$\{U_t\}$ & $\{V_t\}$ are AR processes (stationary) given by

$$\phi(B) U_t = \epsilon_t$$

$$\theta(B) V_t = \epsilon_t$$

<u>Note</u>: If $\underbrace{p=0}_{MA}$, then $V(\underset{\sim}{\beta}) = \sigma^2 (E\,\underset{\sim}{V}_t \underset{\sim}{V}_t')^{-1}$.

If $\underbrace{q=0}_{AR}$, then $V(\underset{\sim}{\beta}) = \sigma^2 (E\,\underset{\sim}{U}_t \underset{\sim}{U}_t')^{-1}$

Example : AR(p)

$$\phi(B)X_t = \epsilon_t$$

$$\underset{\sim}{\phi} = (\phi_1, \ldots, \phi_p)'$$

$$V(\underset{\sim}{\phi}) = \sigma^2 (E\, \underline{U}_t \underline{U}_t')^{-1}$$

U_t is $\ni$ $\phi(B) U_t = \epsilon_t$

$$E\, \underline{U}_t \underline{U}_t' = \Gamma_p = Cov(\underset{\sim}{U}_t) \; ; \; U_t = \begin{pmatrix} U_t \\ U_{t-1} \\ \vdots \\ U_{t-(p-1)} \end{pmatrix}.$$

$$= \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{p-1} \\ & \ddots & & \vdots \\ & & \ddots & \gamma_1 \\ & & & \gamma_0 \end{pmatrix}$$

$$V(\underset{\sim}{\phi}) = \sigma^2 \Gamma_p^{-1}$$

$$\sqrt{n}\,(\hat{\underset{\sim}{\phi}}_{MLE} - \underset{\sim}{\phi}) \xrightarrow{\mathcal{L}} N_p(\underset{\sim}{0}, \sigma^2 \Gamma_p^{-1})$$

i.e. $\hat{\underset{\sim}{\phi}}_{MLE} \sim N_p(\underset{\sim}{\phi}, \frac{1}{n}\sigma^2 \Gamma_p^{-1})$ for large n

Use Galbraith's formula to write elements of Γ_p^{-1}.

AR(1): $\sqrt{n}\,(\hat{\phi}_{MLE} - \phi) \xrightarrow{\mathcal{L}} N\left(0, \sigma^2 \left(\frac{\sigma^2}{1-\phi^2}\right)^{-1}\right)$

i.e. $N(0, 1-\phi^2)$

Random sampling from stationary time series

Let $X_1, \ldots, X_n$ be a sample of size n from a stationary time series with

(i) $E X_t = \mu \quad \forall t$

(ii) $\gamma_h = \mathrm{Cov}(X_t, X_{t+h}) = E(X_t - \mu)(X_{t+h} - \mu) \quad \forall t$

& (iii) $\sum_h |\gamma_h| < \infty$

Estimation of μ

$\bar{X}_n = \frac{1}{n} \sum_{t=1}^{n} X_t$ is an unbiased estimator for μ

$$E \bar{X}_n = \mu$$

$$V \bar{X}_n = \frac{1}{n} \sum_{|h| \le n} \left(1 - \frac{|h|}{n}\right) \gamma_h$$

Some important asymptotic results:

Result 1: $E(\bar{X}_n - \mu)^2 \longrightarrow 0$ as $n \to \infty$

i.e. $\bar{X}_n \xrightarrow{m.s.} \mu$

Random sampling from stationary time series

Let $X_1, \ldots, X_n$ be a sample of size n from a stationary time series with

(i) $E X_t = \mu \quad \forall t$

(ii) $\gamma_h = \operatorname{Cov}(X_t, X_{t+h}) = E(X_t - \mu)(X_{t+h} - \mu) \quad \forall t$

& (iii) $\sum_h |\gamma_h| < \infty$

Estimation of μ

$\bar{X}_n = \frac{1}{n} \sum_{t=1}^{n} X_t$ is an unbiased estimator for μ

$$E \bar{X}_n = \mu$$

$$V \bar{X}_n = \frac{1}{n} \sum_{|h| \le n} \left(1 - \frac{|h|}{n}\right) \gamma_h$$

Some important asymptotic results:

Result 1: $E(\bar{X}_n - \mu)^2 \longrightarrow 0$ as $n \to \infty$

i.e. $\bar{X}_n \xrightarrow{m.s.} \mu$

Pf: $$n E(\bar{X}_n - \mu)^2 = n V(\bar{X}_n) = \sum_{h=-n}^{n} \left(1 - \frac{|h|}{n}\right) \gamma_h$$

i.e. $$n E(\bar{X}_n - \mu)^2 = \left| \gamma_0 + \left(1 - \frac{1}{n}\right) 2\gamma_1 + \left(1 - \frac{2}{n}\right) 2\gamma_2 + \cdots + \left(\frac{n-(n-1)}{n}\right) 2\gamma_{n-1} \right|$$

$$\le |\gamma_0| + \left(\frac{n-1}{n}\right) 2|\gamma_1| + \left(\frac{n-2}{n}\right) 2|\gamma_2| + \cdots + \frac{1}{n} 2|\gamma_{n-1}|$$

$$< |\gamma_0| + 2|\gamma_1| + \cdots + 2|\gamma_{n-1}|$$

$$\le 2 \sum_{h=0}^{n-1} |\gamma_h| \to 2 \sum_{0}^{\infty} |\gamma_h| < \infty$$

$$\Rightarrow \quad E(\bar{X}_n - \mu)^2 = O\left(\frac{1}{n}\right)$$

$$\Rightarrow \quad E(\bar{X}_n - \mu)^2 \to 0 \quad \text{as } n \to \infty$$

i.e. $\bar{X}_n \xrightarrow{m.s.} \mu$

Remark : as $\bar{X}_n \xrightarrow{m.s.} \mu$; we also have $\bar{X}_n \xrightarrow{p} \mu$

Result 2 : $\displaystyle\lim_{n \to \infty} n\, V(\bar{X}_n) = \sum_{h=-\infty}^{\infty} \gamma_h$

Pf :
$$n\, V(\bar{X}_n) = n\, E(\bar{X}_n - \mu)^2$$
$$= \sum_{h=-n}^{n} \left(1 - \frac{|h|}{n}\right) \gamma_h$$
$$= \sum_{h=-n}^{n} \gamma_h - \sum_{h=-n}^{n} \frac{|h|}{n} \gamma_h$$

$$\lim_{n \to \infty} \sum_{h=-n}^{n} \gamma_h = \sum_{-\infty}^{\infty} \gamma_h$$

Also, $\displaystyle\sum_{h=-n}^{n} \frac{|h|}{n} \gamma_h = 2 \sum_{h=0}^{n} \frac{h \gamma_h}{n}$

Now $\displaystyle\sum_{h=0}^{n} \frac{h \gamma_h}{n} = \sum_{h=0}^{N} \frac{h \gamma_h}{n} + \sum_{N+1}^{n} \frac{h}{n} \gamma_h$

$$\lim_{n \to \infty} \frac{1}{n} \underbrace{\sum_{h=0}^{N} h \gamma_h}_{\text{finite qty}} = 0$$

& $\displaystyle\left| \sum_{N+1}^{n} \frac{h}{n} \gamma_h \right| \le \sum_{N+1}^{n} \left| \frac{h}{n} \gamma_h \right| \le \sum_{N+1}^{n} |\gamma_h|$

Since $\sum_{0}^{\infty} |\gamma_h| < \infty$, $\forall\, \epsilon > 0$ $\exists$ an N_0 (for large n)

$\ni \forall\, N \geq N_0$ $\sum_{N+1}^{\infty} |\gamma_h| < \epsilon$

$\Rightarrow \left| \sum_{N+1}^{n} \frac{h}{n} \gamma_h \right| < \epsilon$ for large n

$\Rightarrow \lim_{n \to \infty} \sum_{h=-n}^{n} \frac{|h|}{n} \gamma_h = 0$

Hence $\lim_{n \to \infty} n V(\bar{X}_n) = \sum_{h=-\infty}^{\infty} \gamma_h$

Example 1

AR(1) $X_t = \phi X_{t-1} + \epsilon_t$; $|\phi| < 1$, $\epsilon_t \sim WN(0, \sigma^2)$

$$\gamma_h = \frac{\sigma^2}{1-\phi^2} \phi^{|h|}$$

$\Rightarrow \sum |\gamma_h| < \infty$ (as $|\phi| < 1$)

$$n V(\bar{X}_n) = \sum_{h=-\infty}^{\infty} \frac{\sigma^2}{1-\phi^2} \cdot \phi^{|h|} \quad \text{for large } n$$

$$= \frac{\sigma^2}{1-\phi^2} \left(1 + 2\phi + 2\phi^2 + \cdots \right)$$

$$= \frac{\sigma^2}{1-\phi^2} \left(1 + \frac{2\phi}{1-\phi} \right) = \frac{\sigma^2}{1-\phi^2} \cdot \frac{1+\phi}{1-\phi}$$

$$= \frac{\sigma^2}{(1-\phi)^2} \; ; \; V(\bar{X}_n) \approx \frac{1}{n} \frac{\sigma^2}{(1-\phi)^2} \text{ for large } n$$

Example 2 :

X_t is stationary ARMA(p, q)

$$\phi(B) X_t = \theta(B) \epsilon_t$$

$$X_t = \phi(B)^{-1} \theta(B) \epsilon_t$$

$$\text{i.e. } X_t = \Psi(B) \epsilon_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$$

If $\sum |\psi_j| < \infty$, then $\sum |\gamma_h| < \infty$

then $\quad n V(\bar{X}_n) \simeq \sum_{h=-\infty}^{\infty} \gamma_h \quad$ for large n

Remark : Note that ACGF of $\{X_t\}$ is

$$g_X(z) = \sum_{h=-\infty}^{\infty} \gamma_h z^h$$

$$\Rightarrow \sum_{-\infty}^{\infty} \gamma_h = \lim_{n \to \infty} n V(\bar{X}_n) = g_X(1)$$

MA(1)

$X_t = \theta(B) \epsilon_t$

$g_X(z) = \sigma^2 \theta(z) \theta(z^{-1})$

$g_X(1) = \sigma^2 (1+\theta)^2$

(i) $X_t \sim AR(1)$; $\quad \phi(B) X_t = \epsilon_t$

$$\text{ACGF} \quad g_X(z) = \frac{\sigma^2}{\phi(z)\phi(z^{-1})} = \frac{\sigma^2}{(1-\phi z)(1-\phi z^{-1})}$$

$$\Rightarrow \lim_{n \to \infty} n V(\bar{X}_n) = \sum_{h=-\infty}^{\infty} \gamma_h = g_X(1) = \frac{\sigma^2}{(1-\phi)^2}.$$

(ii) $X_t \sim MA(\infty)$ $\quad X_t = \Psi(B) \epsilon_t = \sum_{0}^{\infty} \psi_j \epsilon_{t-j}$

with $\sum |\psi_j| < \infty$

then $\sum_{-\infty}^{\infty} \gamma_h = g_X(1) = \sigma^2 (\Psi(1))^2 = \sigma^2 (1 + \psi_1 + \psi_2 + \cdots)^2$

Distribution of $\bar{X}_n$

Case 1: Gaussian time series

Suppose $\{X_t\}$ is a Gaussian time series.

i.e. $\forall\, n$, $\underset{\sim}{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \sim N_n(\mu \underset{\sim}{1}_n, \Gamma_n)$

$$\Gamma_n = \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{n-1} \\ & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & \gamma_2 \\ & & & \ddots & \gamma_1 \\ & & & & \gamma_0 \end{pmatrix}.$$

[Note that if $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$ then

(i) $\underset{\sim}{X} - \underset{\sim}{\mu} \sim N_p(\underset{\sim}{0}, \Sigma)$

(ii) $\underset{q\times p}{A}\underset{\sim}{X} + \underset{q\times 1}{\underset{\sim}{b}} \sim N_q(A\underset{\sim}{\mu} + \underset{\sim}{b}, A\Sigma A')$

(i) & (ii) follows from the def^n of N_p dist^n.]

$$\bar{X}_n = \frac{1}{n}\sum_{t=1}^{n} X_t = 1_n' \underset{\sim}{X}/n \; ; \; \underset{\sim}{X} = (X_1, \ldots, X_n)'$$

Using (ii) above,

$$\bar{X}_n = \left(\frac{1}{n} 1_n'\right) \underset{\sim}{X}$$

$$\sim N_1\left(\frac{1}{n} 1_n' \cdot \mu \underset{\sim}{1}_n \,,\, \frac{1}{n^2} \underset{\sim}{1}_n' \Gamma_n \underset{\sim}{1}_n\right)$$

i.e. $\bar{X}_n \sim N_1\left(\mu, \frac{1}{n}\sum_{h=-n}^{n}\left(1 - \frac{|h|}{n}\right)\gamma_h\right)$

i.e. $\sqrt{n}\left(\bar{X}_n - \mu\right) \sim N_1\left(0, \sum_{|h|\le n}\left(1-\frac{|h|}{n}\right)\gamma_h\right)$

Let $v = \sum_{|h|\le n}\left(1-\frac{|h|}{n}\right)\gamma_h$

$$P\left(\frac{\sqrt{n}\,|\bar{X}_n-\mu|}{v^{1/2}} \le \tau_{\alpha/2}\right) = 1-\alpha \quad -(*)$$

$\tau_{\alpha/2}$ is a pt $\ni$ $P(Z > \tau_{\alpha/2}) = \alpha/2$; $Z\sim N(0,1)$

$(*) \Rightarrow$ the $100(1-\alpha)\%$ Confidence interval for μ in such a Case is

$\bar{X}_n \mp \tau_{\alpha/2}\, v^{1/2}$ (provided v is known)

Note that v is usually unknown as $\{\gamma_h\}$ is unknown and we use estimate of v as

$$\hat{v} = \sum_{|h|\le n}\left(1-\frac{|h|}{n}\right)\hat{\gamma}_h$$

Note $\quad \frac{\sqrt{n}\,(\bar{X}_n-\mu)}{\sqrt{\hat{v}}} \xrightarrow{L} N(0,1)$

$\xrightarrow{L}$: Convergence in law

So, the asymptotic $100(1-\alpha)\%$ Confidence interval for μ is $\bar{X}_n \mp \tau_{\alpha/2}\sqrt{\hat{v}}$

Case 2 : Non-Gaussian linear process

We only have asymptotic dist^n result for such non-Gaussian linear processes.

Let $\{X_t\}$ be a covariance stationary linear time series

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j \epsilon_{t-j} \quad ; \quad \epsilon_t \sim WN(0, \sigma^2)$$

where $\sum_j |\psi_j| < \infty$ & $\sum_j \psi_j \neq 0$, then

$$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{L} N\left(0, \sum_{h=-\infty}^{\infty} \gamma_h\right)$$

i.e $\sqrt{n}(\bar{X}_n - \mu) \overset{asym}{\sim} N\left(0, \left(\sum \psi_j\right)^2 \sigma^2\right)$.

Example : $\{X_t\}$ is non Gaussian stationary AR(1)

$$X_t = \delta + \phi X_{t-1} + \epsilon_t \quad ; \quad \epsilon_t \sim WN(0, \sigma^2)$$

$$|\phi| < 1$$

$$\delta = \mu(1-\phi) \quad ; \quad \mu = \frac{\delta}{1-\phi}$$

$$(1-\phi B) X_t = \delta + \epsilon_t$$

$$X_t = (1-\phi B)^{-1} \delta + (1-\phi B)^{-1} \epsilon_t$$

i.e $X_t = \left(\frac{\delta}{1-\phi}\right) + \sum_{j=0}^{\infty} \phi^j \epsilon_{t-j}$

i.e. $X_t = \mu + \sum_{j=0}^{\infty} \phi^j \epsilon_{t-j}$

For AR(1), $\sum_{-\infty}^{\infty} \gamma_h = g_X(1) = \frac{\sigma^2}{(1-\phi)^2}$.

Using the asymptotic result, we have

$$\sqrt{n}\left(\bar{X}_n - \mu\right) \xrightarrow{d} N\left(0, \frac{\sigma^2}{(1-\phi)^2}\right)$$

Estimation of γ_h / ρ_h

$$\hat{\gamma}_h = \frac{1}{n}\sum_{t=1}^{n-h} (X_t - \bar{X}_n)(X_{t+h} - \bar{X}_n)$$

i.e. $$\hat{\gamma}_h = \frac{1}{n}\sum_{t=h+1}^{n} (X_t - \bar{X}_n)(X_{t-h} - \bar{X}_n)$$

$$\hat{\rho}_h = \hat{\gamma}_h / \hat{\gamma}_0$$

Asymptotic Result for $dist^n$

Suppose $\{X_t\}$ is covariance stationary linear process

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j \epsilon_{t-j}$$

$\epsilon_t \sim WN(0, \sigma^2)$, $\sum_j |\psi_j| < \infty$ & $E(\epsilon_t^4) < \infty$

Then for each h, we have

$$\underset{\sim}{\hat{\rho}(h)} \overset{asym}{\sim} N_h\left(\underset{\sim}{\rho(h)} , \frac{1}{n} W \right)$$

i.e. $$\sqrt{n}\left(\underset{\sim}{\hat{\rho}(h)} - \underset{\sim}{\rho(h)} \right) \xrightarrow{\mathcal{L}} N_h(\underset{\sim}{0}, W)$$

Where $$\underset{\sim}{\hat{\rho}(h)} = \left(\hat{\rho}_1, \ldots, \hat{\rho}_h \right)'$$

$$\underset{\sim}{\rho(h)} = \left(\rho_1, \ldots, \rho_h \right)'$$

$$W = ((w_{ij}))$$

$$w_{ij} = \sum_{k=-\infty}^{\infty} \Big(\rho(k+i)\rho(k+j) + \rho(k-i)\rho(k+j) + 2\rho(i)\rho(j)\rho^2(k) - 2\rho(i)\rho(k)\rho(k+j) - 2\rho(j)\rho(k)\rho(k+i) \Big)$$

Alt form

$$= \sum_{k=1}^{\infty} \Big(\rho(k+i) + \rho(k-i) - 2\rho(i)\rho(k) \Big) \Big(\rho(k+j) + \rho(k-j) - 2\rho(j)\rho(k) \Big)$$

The above is called the "Bartlett's formula".

Application of the above result

Example 1: Let $\{X_t\}$ be an i.i.d. $\underset{WN}{(0, \sigma^2)}$ sequence

$$\rho_h = 0 \;\forall\; |h| > 0$$

Applying Bartlett's formula

$$w_{ii} = \sum_{k=1}^{\infty} \Big(\rho(k+i) + \rho(k-i) - 2\rho(i)\rho(k) \Big)^2 = \rho(0)^2 = 1$$

$$W_{ij} = \sum_{K=1}^{\infty} \left(\rho(K+i) + \rho(K-i) - 2\rho(i)\rho(K) \right) \left(\rho(K+j) + \rho(K-j) - 2\rho(j)\rho(K) \right)$$

$$= 0 \quad \forall\ i \neq j$$

i.e. $W_{ij} = \begin{cases} 1, & i = j \\ 0, & \text{o/w} \end{cases}$

$$\sqrt{n}\left(\hat{\underset{\sim}{\rho}}(h) - \underset{\sim}{\rho}(h) \right) \xrightarrow{L} N_h(\underset{\sim}{0}, I_h)$$

i.e. $\hat{\underset{\sim}{\rho}}(h) \overset{asym}{\sim} N_h\left(\underset{\sim}{\rho}(h), \frac{1}{n} I_h \right)$

So, for large n, $\hat{\rho}_1, \ldots, \hat{\rho}_h$ are approximately independent and identically distributed univariate normal r.v.s with mean 0 and variance $\frac{1}{n}$.

Example 2 :

$$X_t = \epsilon_t + \theta \epsilon_{t-1} \ ; \ \epsilon_t \sim WN(0, \sigma^2)$$

$$W_{ii} = \sum_{K=1}^{\infty} \left(\rho(K+i) + \rho(K-i) - 2\rho(i)\rho(K) \right)^2$$

$$W_{11} = \sum_{K=1}^{\infty} \left(\rho(K+1) + \rho(K-1) - 2\rho(1)\rho(K) \right)^2$$

$$= \underbrace{\left(\rho(0) - 2\rho(1)\rho(1) \right)^2}_{K=1} + \underbrace{\left(\rho(1) \right)^2}_{K=2} + 0 \cdots$$

i.e. $W_{11} = 1 + 4\rho(1)^4 - 4\rho(1)^2 + \rho(1)^2$

$$W_{11} = 1 - 3\rho(1)^2 + 4\rho(1)^4$$

$$W_{22} = \sum_{K=1}^{\infty} \left(\rho(K+2) + \rho(K-2) - 2\rho(2)\rho(K) \right)^2$$

$$W_{22} = \underset{\substack{\uparrow \\ K=1}}{\rho(1)^2} + \underset{\substack{\uparrow \\ K=2}}{\rho(0)^2} + \underset{\substack{\uparrow \\ K=3}}{\rho(1)^2} = 1 + 2\rho(1)^2$$

$\forall i \geq 2$

$$W_{ii} = \sum_{K=1}^{\infty} \left(\rho(K+i) + \rho(K-i) - 2\rho(i)\rho(K) \right)^2$$

i.e. $W_{ii} = \underset{\substack{\downarrow \\ K=i-1}}{\rho(1)^2} + \underset{\substack{\downarrow \\ K=i}}{\rho(0)^2} + \underset{\substack{\downarrow \\ K=i+1}}{\rho(1)^2}$

$$W_{ii} = 1 + 2\rho(1)^2 \quad \forall i \geq 2$$

Further $W_{ij} \neq 0$ for $i \neq j$.

$$\sqrt{n}\left(\hat{\rho}_1 - \rho_1\right) \xrightarrow{L} N\left(0,\ 1 - 3\rho_1^2 + 4\rho_1^4\right)$$

i.e. $\hat{\rho}_1 \overset{asym}{\sim} N\left(\rho_1,\ \frac{1}{n}\left(1 - 3\rho_1^2 + 4\rho_1^4\right)\right)$

& $\forall i > 1$

$$\sqrt{n}\,\hat{\rho}_i \xrightarrow{L} N\left(0,\ 1 + 2\rho_1^2\right) \quad \left(\begin{array}{l}\text{Note that} \\ \rho_i = 0\ \forall i > 1\end{array}\right)$$

Forecasting in stationary time series: Best Linear Predictor (BLP)

$\{X_t\}$ - Covariance stationary time series with mean μ and ACVF $\{\gamma_h\}$

Given information upto time n, $\{X_1, \ldots, X_n\}$ problem is to predict X_{n+h} for some $h > 0$.

BLP approach: Find the linear combination of $X_n, \ldots, X_1$ that provides the "best" forecast of X_{n+h}

"best": w.r.t. minimum mean square prediction error

Def^n: Best linear Predictor (BLP)

BLP of X_{n+h} in terms of $(X_n, X_{n-1}, \ldots, X_1)$ denoted by $P_{(X_n, \ldots, X_1)} X_{n+h} = P_n X_{n+h}$ is the linear f^n

$$a_0^* + a_1^* X_n + \cdots + a_n^* X_1 \text{ if}$$

$$E\left(X_{n+h} - P_{(X_n, \ldots, X_1)} X_{n+h}\right)^2 \text{ is minimum}$$

among all such linear functions.

Derivation of BLP

Let $\underset{\sim}{a} = (a_0, a_1, \ldots, a_n)'$ and

$$S(\underset{\sim}{a}) = E\left(X_{n+h} - a_0 - \sum_{i=1}^{n} a_i X_{n+1-i}\right)^2$$

$$\underset{\sim}{a}_{BLP} = \arg\min_{\underset{\sim}{a}} S(\underset{\sim}{a})$$

BLP eq^{n}s :

$$\frac{\partial S(\underset{\sim}{a})}{\partial a_j} = 0 \qquad j = 0, 1, \ldots, n$$

$\frac{\partial S}{\partial a_0} = 0$ gives

$$E\left(X_{n+h} - a_0 - \sum_{i=1}^{n} a_i X_{n+1-i}\right) = 0 \qquad -\text{(i)}$$

$\frac{\partial S}{\partial a_j} = 0$ gives

$j = 1, \ldots, n$ $$E\left(X_{n+h} - a_0 - \sum_{i=1}^{n} a_i X_{n+1-i}\right) X_{n+1-j} = 0 \qquad -\text{(ii)}$$

$$\text{(i)} \Rightarrow \mu - a_0 - \sum_{i=1}^{n} a_i \mu = 0$$

$$\text{i.e. } a_0 = \mu\left(1 - \sum_{i=1}^{n} a_i\right)$$

$$\text{(ii)} \Rightarrow E\left(X_{n+h} X_{n+1-j}\right) - a_0 \mu - \sum_{1}^{n} a_i E\left(X_{n+1-i} X_{n+1-j}\right) = 0$$

$j = 1(1)n$

$$\text{i.e. } E\left(X_{n+h} X_{n+1-j}\right) - \mu^2\left(1 - \sum_{1}^{n} a_i\right) - \sum_{1}^{n} a_i E\left(X_{n+1-i} X_{n+1-j}\right) = 0$$

$j = 1(1)n$

$$\text{i.e. } \left(E\left(X_{n+h} X_{n+1-j}\right) - \mu^2\right) - \sum_{1}^{n} a_i \left(E\left(X_{n+1-i} X_{n+1-j}\right) - \mu^2\right) = 0$$

$j = 1(1)n$

$$\text{i.e. } \gamma_{h+j-1} = \sum_{i=1}^{n} a_i \gamma_{i-j}; \qquad j = 1(1)n$$

Thus the BLP prediction eqns are

$$a_0 = \mu\left(1 - \sum_{1}^{n} a_i\right) \qquad -\text{(iii)}$$

$$\gamma_{h+j-1} = \sum_{i=1}^{n} a_i \gamma_{i-j}; \quad j = 1(1)n \qquad -\text{(iv)}$$

Let $\underset{\sim}{\gamma}_n(h) = (\gamma_h, \gamma_{h+1}, \ldots, \gamma_{h+n-1})'$

$\underset{\sim}{a}_n = (a_1, \ldots, a_n)'$

$\Gamma_n = ((\gamma_{i-j}))$

(iv) is $\underset{\sim}{\gamma}_n(h) = \Gamma_n \underset{\sim}{a}_n$

BLP is $P_n X_{n+h} = \mu(1 - \sum a_i) + \sum_{i=1}^{n} a_i X_{n+1-i}$

$$= \mu + \sum_{i=1}^{n} a_i (X_{n+1-i} - \mu)$$

$\underset{\sim}{a}_n \ni \underset{\sim}{\gamma}_n(h) = \Gamma_n \underset{\sim}{a}_n$

<u>Note</u> (i) $E(X_{n+h} - P_n X_{n+h}) = 0$

(ii) $E(X_{n+h} - P_n X_{n+h})^2$

$$= E(X_{n+h} - \mu - \underset{\sim}{a}_n' \underset{\sim}{Y}_n)^2$$

$(\underset{\sim}{Y}_n = (X_n - \mu, \ldots, X_1 - \mu)')$

$$= E((X_{n+h} - \mu) - \underset{\sim}{a}_n' \underset{\sim}{Y}_n)((X_{n+h} - \mu) - \underset{\sim}{a}_n' \underset{\sim}{Y}_n)'$$

$$= E(X_{n+h} - \mu)^2 + \underset{\sim}{a}_n' E \underset{\sim}{Y}_n \underset{\sim}{Y}_n' \underset{\sim}{a}_n - 2E(X_{n+h} - \mu)\underset{\sim}{a}_n' \underset{\sim}{Y}_n$$

$$= \gamma_0 + \underset{\sim}{a}_n' \Gamma_n \underset{\sim}{a}_n - 2 \underset{\sim}{a}_n' \underset{\sim}{\gamma}_n(h)$$

$$= \gamma_0 + \underset{\sim}{a}_n' \underset{\sim}{\gamma}_n(h) - 2 \underset{\sim}{a}_n' \underset{\sim}{\gamma}_n(h) \quad \left(\begin{array}{l} \underset{\sim}{a}_n \text{ is } \ni \\ \underset{\sim}{\gamma}_n(h) = \Gamma_n \underset{\sim}{a}_n \end{array}\right)$$

$$= \gamma_0 - \underset{\sim}{a}_n' \underset{\sim}{\gamma}_n(h)$$

This is the mean square prediction error corresponding to BLP.

Note (iii) $E\left(X_{n+h} - P_n X_{n+h}\right) X_j = 0 \quad \forall\, j = 1(1)n$

Note (iv) If $\mu = 0$, then $a_0 = 0$, so we can start with prediction eq^n without a constant

Remark: sp case - one step ahead prediction for zero mean

$$P_n X_{n+1} = \sum_{i=1}^{n} \beta_i X_{n+1-i}$$

$\underset{\sim}{\beta}_n = (\beta_1, \ldots, \beta_n)'$ is $\ni$

$$\Gamma_n \underset{\sim}{\beta}_n = \underset{\sim}{\gamma}_n(1)$$

$\underset{\sim}{\gamma}_n(1) = (\gamma_1, \ldots, \gamma_n)' \qquad \Gamma_n = ((\gamma_{i-j}))$

$$\underset{\sim}{\beta}_n^{(BLP)} = \Gamma_n^{-1} \underset{\sim}{\gamma}_n(1)$$

use estimates $\hat{\gamma}_i$ to get

$$\hat{\underset{\sim}{\beta}}_n^{BLP} = \hat{\Gamma}_n^{-1} \hat{\underset{\sim}{\gamma}}_n(1)$$

Predicted X_{n+1} : $\hat{X}_{n+1}^{BLP} = \left(\hat{\Gamma}_n^{-1} \hat{\underset{\sim}{\gamma}}_n(1)\right)' \underset{\sim}{X}_n$

Example: Prediction for AR(2)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t$$

Prediction of X_2 using X_1

$$P_{X_1} X_2 = \beta_1 X_1$$

β_1 is $\ni$ $E(X_2 - \beta X_1)^2$ is minimized w.r.t. β

$$g(\beta) = E(X_2 - \beta X_1)^2$$

$$\frac{\partial g}{\partial \beta} = 0 \Rightarrow E(X_2 - \beta X_1) X_1 = 0$$

i.e. $\gamma_1 = \beta \gamma_0$

$$\Rightarrow \beta = \frac{\gamma_1}{\gamma_0} = \rho_1$$

$$\frac{\partial^2 g}{\partial \beta^2} = +\gamma_0 \qquad P_{X_1} X_2 = \rho_1 X_1$$

Prediction of X_3 using X_2 & X_1

$$P_{(X_2, X_1)} X_3 = \beta_1^* X_2 + \beta_2^* X_1$$

β_1^*, β_2^* are $\ni$ $g(\beta_1, \beta_2) = E\left(X_3 - \beta_1 X_2 - \beta_2 X_1\right)^2$ is minimized w.r.t. β_1 & β_2

BLP eq^{n}s : $\dfrac{\partial g}{\partial \beta_1} = 0 \Rightarrow E\left(X_3 - \beta_1 X_2 - \beta_2 X_1\right) X_2 = 0$

i.e. $\gamma_1 = \beta_1 \gamma_0 + \beta_2 \gamma_1$

$$\frac{\partial g}{\partial \beta_2} = 0 \Rightarrow E\left(X_3 - \beta_1 X_2 - \beta_2 X_1\right) X_1 = 0$$

i.e. $\gamma_2 = \beta_1 \gamma_1 + \beta_2 \gamma_0$

$$\begin{pmatrix} \beta_1^* \\ \beta_2^* \end{pmatrix}_{BLP} = \begin{pmatrix} \gamma_0 & \gamma_1 \\ \gamma_1 & \gamma_0 \end{pmatrix}^{-1} \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \leftarrow \text{as } \underline{\text{expected}}$$

Notice the similarity of BLP eq^{n}s with Yule-Walker eq^{n}s.

Further,

$$\forall n \geq 2 ; \quad P_{(X_n, \ldots, X_1)} X_{n+1} = \phi_1 X_n + \phi_2 X_{n-1}$$

i.e. $\beta_1^* = \phi_1$, $\beta_2^* = \phi_2$; $\beta_3^* = \cdots = \beta_n^* = 0$

Prediction in a general setup

Suppose Y and $W_n, \dots, W_1$ are any random variables with finite 2^{nd} order joint moments

$$E(Y) = \mu \quad ; \quad E(W_i) = \mu_{W_i}$$

$Cov(Y, W_i)$ $\forall i$ $Cov(W_i, W_j)$ $\forall i, j$ and $V(Y)$ are all finite.

Consider linear predictor of Y based on $(W_n, \dots, W_1)$

$$P_{(W_n, \dots, W_1)} Y = P(Y|W) = a_0^* + a_1^* W_n + \dots + a_n^* W_1$$

The above BLP is $\exists$ mean square prediction error of LP is minimum

i.e. $E(Y - P(Y|W))^2$ is min w.r.t. all possible linear predictors.

Derivation of BLP

Let $\underset{\sim}{a} = (a_0, a_1, \dots, a_n)'$

$$S(\underset{\sim}{a}) = E\left(Y - a_0 - \sum_{i=1}^{n} a_i W_{n+1-i}\right)^2$$

$$\underset{\sim}{a}_{BLP} = \underset{\underset{\sim}{a}}{\arg\min}\ S(\underset{\sim}{a})$$

BLP eq^{n}s

$$\frac{\partial S(\underset{\sim}{a})}{\partial a_j} = 0 \; ; \quad j = 0, 1, \ldots, n$$

$\dfrac{\partial S(\underset{\sim}{a})}{\partial a_0} = 0$ gives

$$E\left(Y - a_0 - \sum_{i=1}^{n} a_i W_{n+1-i}\right) = 0$$

i.e. $a_0 = \mu_Y - \underset{\sim}{a}_n' \underset{\sim}{\mu}_{\underset{\sim}{W}}$

$$\underset{\sim}{\mu}_{\underset{\sim}{W}} = (\mu_{W_n}, \ldots, \mu_{W_1})'$$

$\dfrac{\partial S(\underset{\sim}{a})}{\partial a_j} = 0$ gives

$$E\left(Y - a_0 - \sum_{1}^{n} a_i W_{n+1-i}\right) W_{n+1-j} = 0$$

$$j = 1, \ldots, n$$

i.e. $E(Y W_{n+1-j}) - a_0 \mu_{W_{n+1-j}} - \sum_{1}^{n} a_i E(W_{n+1-i} W_{n+1-j}) = 0$

$$E(Y W_{n+1-j}) - (\mu_Y - \underset{\sim}{a}_n' \underset{\sim}{\mu}_{\underset{\sim}{W}}) \mu_{W_{n+1-j}} - \sum_{1}^{n} a_i E(W_{n+1-i} W_{n+1-j}) = 0$$

~~[illegible]~~,

$$\left(E(Y W_{n+1-j}) - \mu_Y \mu_{W_{n+1-j}}\right)$$

$$- \sum_{i=1}^{n} a_i \left(E(W_{n+1-i} W_{n+1-j}) - \mu_{W_{n+1-i}} \mu_{W_{n+1-j}} \right) = 0.$$

i.e. $\mathrm{Cov}(Y, W_{n+1-j}) - \sum_{i=1}^{n} a_i \,\mathrm{Cov}(W_{n+1-i}, W_{n+1-j}) = 0$

$$j = 1, \ldots, n$$

Applications

(i) Estimation of missing values.

Let $X_t = \phi X_{t-1} + \epsilon_t$; $|\phi| < 1$, $\epsilon_t \sim WN(0, \sigma^2)$

Suppose X_2 is missing and we wish to use X_1 and X_3 to estimate X_2

Frame BLP of X_2 using X_1, X_3

$$P_{(X_1, X_3)} X_2 = \alpha_1 X_1 + \alpha_2 X_3$$

$$\left(Y = X_2 \;\; ; \;\; \underset{\sim}{W} = (X_1, X_3)' \right)$$

$$\operatorname{cov}(\underset{\sim}{W}) = \begin{pmatrix} \gamma_0 & \gamma_2 \\ \gamma_2 & \gamma_0 \end{pmatrix} = \frac{\sigma^2}{1-\phi^2} \begin{pmatrix} 1 & \phi^2 \\ \phi^2 & 1 \end{pmatrix}$$

$$\underset{\sim}{\gamma} = \begin{pmatrix} \operatorname{cov}(Y, X_1) \\ \operatorname{cov}(Y, X_3) \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ \gamma_1 \end{pmatrix} = \frac{\sigma^2}{1-\phi^2} \begin{pmatrix} \phi \\ \phi \end{pmatrix}$$

BLP eqns $\quad \Gamma \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \underset{\sim}{\gamma}$

i.e $\begin{pmatrix} 1 & \phi^2 \\ \phi^2 & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} \phi \\ \phi \end{pmatrix}$.

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}_{BLP} = \frac{1}{1-\phi^4} \begin{pmatrix} 1 & -\phi^2 \\ -\phi^2 & 1 \end{pmatrix} \begin{pmatrix} \phi \\ \phi \end{pmatrix}.$$

i.e $\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}_{BLP} = \frac{1}{1+\phi^2} \begin{pmatrix} \phi \\ \phi \end{pmatrix}$

$$P_{(X_1, X_3)} X_2 = \frac{\phi}{1+\phi^2}(X_1 + X_3)$$

missing value estimate of X_2 using BLP

minimum mean square prediction error

$$E\left(X_2 - P_{(X_1, X_3)} X_2\right)^2 = V(X_2) - \underline{\alpha}' \underset{\sim}{\gamma}$$

$$= \frac{\sigma^2}{1-\phi^2} - \frac{1}{1+\phi^2}(\phi, \phi) \frac{\sigma^2}{1-\phi^2}\begin{pmatrix}\phi \\ \phi\end{pmatrix}$$

$$= \frac{\sigma^2}{1+\phi^2}$$

(ii) Back forecasting

use $(X_1, \ldots X_n)$ to back forecast $X_0, X_{-1}, \ldots$ required for conditional MLE initialization of AR / ARMA models.

$$\underbrace{P_{(X_1, \ldots X_n)}}_{\underset{\sim}{W}} X_0 \leftarrow Y$$

(iii) Defining partial auto correlation function (PACF) - a major tool in model identification

Partial Auto Correlation function (PACF)

PACF, $\alpha(k)$, at lag k, is the correlation between X_1 and X_{k+1}, adjusting for the intervening observations $(X_2, \ldots, X_k)$.

<u>Def</u>^{n}: PACF $\alpha(.)$ of a stationary process $\{X_t\}$ is defined by

$$\alpha(1) = \text{Corr}^n(X_2, X_1) = \rho(1)$$

and $$\alpha(k) = \text{Corr}^n\left(X_{k+1} - P_{(X_k, \ldots, X_2)} X_{k+1},\; X_1 - P_{(X_k, \ldots, X_2)} X_1\right)$$

$k \geq 2$

$\alpha(k)$: PACF at lag k

$P_{(\,)} X_{k+1}$ & $P_{(\,)} X_1$ can be found using the BLP approach

PACF of standard prob models

PACF of AR(1)

$\{X_t\}$ covariance stationary AR(1)

$$X_t = \phi X_{t-1} + \epsilon_t \;;\; |\phi| < 1,\; \epsilon_t \sim WN(0, \sigma^2)$$

$$\alpha(1) = \rho(1) = \phi$$

$$\alpha(2) = \text{Corr}\left(X_3 - P_{X_2} X_3,\; X_1 - P_{X_2} X_1\right)$$

$$f(\beta) = E\left(X_3 - \beta X_2\right)^2$$

BLP eq^n: $E(X_3 - \beta X_2) X_2 = 0$

$$\beta_{BLP} = \rho_1 = \phi$$

$$\tilde{f}(\alpha) = E\left(X_1 - \alpha X_2\right)^2$$

BLP eq^n: $E(X_1 - \alpha X_2) X_2 = 0$

$$\alpha_{BLP} = \rho_1 = \phi$$

$$\Rightarrow P_{X_2} X_3 = \phi X_2 \quad \& \quad P_{X_2} X_1 = \phi X_2$$

$$\alpha_{\cancel{2}}(2) = \text{Corr}\left(X_3 - \phi X_2, X_1 - \phi X_2\right)$$

$$= \text{Corr}\left(\epsilon_3, X_1 - \phi X_2\right)$$

$$= \frac{\text{Cov}(\epsilon_3, X_1 - \phi X_2)}{\left[V(\epsilon_3)\; V(X_1 - \phi X_2)\right]^{1/2}}$$

$$= 0 \quad \text{as } \text{Cov}(\epsilon_t, X_{t-j}) = 0 \quad \forall j > 0$$

$$\alpha(3) = \text{Corr}\left(X_4 - P_{(X_2, X_3)} X_4, X_1 - P_{(X_2, X_3)} X_1\right).$$

$$= \text{Corr}\left(X_4 - \phi X_3, X_1 - \alpha_{1(BLP)} X_2 - \alpha_{2(BLP)} X_3\right)$$

$$= \text{Corr}\left(\epsilon_4, X_1 - \alpha_{1(BLP)} X_2 - \alpha_{2(BLP)} X_3\right)$$

$$= 0$$

$\forall K \geq 2 \quad \alpha(K) = 0.$

Remark: PACF of AR(1) cuts off after lag 1.

PACF for AR(2)

$\{X_t\}$ is covariance stationary AR(2)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t$$

$$\alpha(1) = \rho(1)$$

$$\alpha(2) = \mathrm{Corr}\left(X_3 - P_{X_2} X_3 \, , X_1 - P_{X_2} X_1 \right)$$

$$P_{X_2} X_3 = \rho(1) X_2 \qquad P_{X_2} X_1 = \rho(1) X_2$$

$$\alpha(2) = \mathrm{Corr}\left(X_3 - \rho(1) X_2 \, , X_1 - \rho(1) X_2 \right)$$

$$= \frac{\mathrm{Cov}\left(X_3 - \rho(1) X_2 \, , X_1 - \rho(1) X_2 \right)}{\left[V\left(X_3 - \rho(1) X_2 \right) V\left(X_1 - \rho(1) X_2 \right) \right]^{1/2}}$$

$$V\left(X_3 - \rho(1) X_2 \right) = \gamma_0 + \rho(1)^2 \gamma_0 - 2\rho(1)\gamma_1$$

$$= V\left(X_1 - \rho(1) X_2 \right)$$

$$\mathrm{Cov}\left(X_3 - \rho(1) X_2 \, , X_1 - \rho(1) X_2 \right)$$

$$= \gamma_2 - \rho(1)\gamma_1 - \rho(1)\gamma_1 + \rho(1)^2 \gamma_0 .$$

$$= \gamma_2 - 2\rho(1)\gamma_1 + \rho(1)^2 \gamma_0 \neq 0$$

$$\alpha(2) \neq 0 .$$

$$\alpha(3) = \mathrm{Corr}\left(X_4 - P_{(X_2, X_3)} X_4 \, , X_1 - P_{(X_2, X_3)} X_1 \right)$$

$$= \mathrm{Corr}\left(X_4 - \phi_1 X_3 - \phi_2 X_2 \, , X_1 - \alpha_{1(BLP)} X_2 - \alpha_{2(BLP)} X_3 \right)$$

$$= \mathrm{Corr}\left(\epsilon_4 \, , X_1 - \alpha_{1(BLP)} X_2 - \alpha_{2(BLP)} X_3 \right)$$

$$= 0 .$$

Further $\forall K \geq 3$; $\alpha(K) = 0$.

<u>Remark</u> : The above pattern holds true for a stationary AR(p) and for such a model

$$\alpha(K) = 0 \quad \forall K > p$$

<u>Remark</u> : For a stationary AR(p)

$\forall K > p$; $\hat{\alpha}(K) \overset{asym}{\sim} N\left(0, \frac{1}{n}\right)$ for large n

estimator of $\alpha(K)$

PACF for MA(1)

$$X_t = \epsilon_t + \theta\epsilon_{t-1} \; ; \quad \epsilon_t \sim WN(0, \sigma^2)$$

$$\alpha(1) = \rho_1 = \frac{\theta}{1+\theta^2}.$$

$$\alpha(2) = Corr\left(X_3 - P_{X_2}X_3 \, , \, X_1 - P_{X_2}X_1\right)$$

$$P_{X_2}X_3 = \rho_1 X_2 \quad \& \quad P_{X_2}X_1 = \rho_1 X_2$$

$$= \theta^* X_2 \text{, say} \qquad\qquad = \theta^* X_2$$

$$\theta^* = \frac{\theta}{1+\theta^2}.$$

$$\alpha(2) = Corr\left(X_3 - \theta^* X_2 \, , \, X_1 - \theta^* X_2\right)$$

$$= \frac{Cov\left(X_3 - \theta^* X_2 \, , \, X_1 - \theta^* X_2\right)}{\left[V\left(X_3 - \theta^* X_2\right) V\left(X_1 - \theta^* X_2\right)\right]^{1/2}}$$

$$V(X_3 - \theta^* X_2) = \gamma_0 + \theta^{*2}\gamma_0 - 2\theta^*\gamma_1$$

$$= \sigma^2(1+\theta^2) + \left(\frac{\theta}{1+\theta^2}\right)^2 (1+\theta^2)\sigma^2 - 2\left(\frac{\theta}{1+\theta^2}\right)\theta\sigma^2$$

$$= \sigma^2\left((1+\theta^2) - \frac{\theta^2}{1+\theta^2}\right)$$

$$= \sigma^2\left(\frac{1+\theta^2+\theta^4}{1+\theta^2}\right) = V(X_1 - \theta^* X_2)$$

$$Cov(X_3 - \theta^* X_2, X_1 - \theta^* X_2)$$

$$= \underset{\gamma_2}{0} - \theta^*\gamma_1 - \theta^*\gamma_1 + \theta^{*2}\gamma_0$$

$$= -2\left(\frac{\theta}{1+\theta^2}\right)\theta\sigma^2 + \left(\frac{\theta}{1+\theta^2}\right)^2\sigma^2(1+\theta^2)$$

$$= -\frac{\theta^2}{1+\theta^2}\cdot\sigma^2$$

$$\Rightarrow \alpha(2) = -\frac{\theta^2}{1+\theta^2+\theta^4}.$$

in general, $\alpha(k) = -\dfrac{(-\theta)^k}{1+\theta^2+\cdots+\theta^{2k}}$

<u>Remark</u>: PACF of MA(1) does not cut off but tails of

<u>Remark</u>: MA(q) process has a similar behavior of PACF.

<u>Remark</u>: As ARMA(p, q) has MA part, PACF of ARMA process also does not cut off.

Remark: To sum up the behavior of ACF & PACF

Model	ACF	PACF
AR(1)	decays exponentially (tails off)	Single spike (cuts off)
MA(1)	cuts off (after 1 spike)	tails off
AR(p)	tails off	cuts off (after p spikes)
MA(q)	cuts off (after q spikes)	tails off
ARMA(p,q)	tails off	tails off

The above table is to be used for model identification

Model order estimation

Using penalized log likelihood criteria or the information theoretic criteria

Akaike information Criterion (AIC)

General form of AIC:

$$AIC(K) = -2 \log \hat{L} + 2K$$

K: # of parameters in the model

ARMA(p, q) model

$$AIC(p, q) = -2 \log \hat{L} + 2(p+q+1)$$

$$(\hat{p}, \hat{q}) = \underset{\substack{p \in \{0,1,\ldots,P\} \\ q \in \{0,1,\ldots Q\}}}{\arg\min} \; AIC(p, q)$$

Bayesian Information Criterion (BIC)

General form of BIC:

$$BIC(K) = -2 \log \hat{L} + (\log n) K$$

ARMA(p, q) model

$$BIC(p, q) = -2 \log \hat{L} + (\log n)(p+q+1)$$

$$(\hat{p}, \hat{q}) = \underset{\substack{p \in \{0,1,\ldots,P\} \\ q \in \{0,1,\ldots,Q\}}}{\arg\min} \; BIC(p, q)$$

Frequency Domain Analysis

Aim : To study the frequency (corresponding to periodic component) properties of time series and identify dominant frequencies that drive the time series

Tool : Spectral density function

Def^n: Spectral density

Suppose that $\{X_t\}$ is a stationary zero mean time series with autocovariance f^n $\gamma(\cdot)$ satisfying $\sum_h |\gamma(h)| < \infty$. The spectral density of $\{X_t\}$ is the function $f(\cdot)$ defined by

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \gamma(h) \; ; \quad -\infty < \lambda < \infty$$

Frequency Domain Analysis

Aim: To study the frequency (corresponding to periodic component) properties of time series and identify dominant frequencies that drive the time series

Tool: Spectral density function

Def^n: Spectral density

Suppose that $\{X_t\}$ is a stationary zero mean time series with autocovariance f^n $\gamma(\cdot)$ satisfying $\sum_h |\gamma(h)| < \infty$. The spectral density of $\{X_t\}$ is the function $f(\cdot)$ defined by

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \gamma(h) \; ; \quad -\infty < \lambda < \infty$$

Important properties

(1) Alternate form

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) \left(\cos \lambda h - i \sin \lambda h \right)$$

$$= \frac{1}{2\pi} \gamma(0) \left(\cos 0 - i \sin 0 \right) + \frac{1}{2\pi} \left(\sum_{h=1}^{\infty} \gamma(h) \left(\cos \lambda h + \cos(-\lambda h) - i \cancel{\sin(\lambda h)} - i \cancel{\sin(-\lambda h)} \right) \right)$$

i.e. $$f(\lambda) = \frac{1}{2\pi} \left(\gamma(0) + 2 \sum_{h=1}^{\infty} \gamma(h) \cos \lambda h \right)$$

(2) $f(\cdot)$ is even

$$f(\lambda) = \frac{1}{2\pi} \left(\gamma(0) + 2 \sum_{h=1}^{\infty} \gamma(h) \cos \lambda h \right)$$

$$= \frac{1}{2\pi} \left(\gamma(0) + 2 \sum_{h=1}^{\infty} \gamma(h) \cos(-\lambda h) \right)$$

$$= f(-\lambda) \quad \forall \lambda$$

(3) $f(\lambda)$ is periodic with period 2π

$$f(\lambda) = \frac{1}{2\pi}\left(\gamma(0) + 2\sum_{h=1}^{\infty}\gamma(h)\cos\lambda h\right)$$

Note that $\cos((\lambda+2\pi k)h) = \cos\lambda h$ for any integer k & h

$$f(\lambda+2\pi k) = \frac{1}{2\pi}\left(\gamma(0) + 2\sum_{h=1}^{\infty}\gamma(h)\cos\left((\lambda+2\pi k)h\right)\right)$$

$$= \frac{1}{2\pi}\left(\gamma(0) + 2\sum_{h=1}^{\infty}\gamma(h)\cos\lambda h\right)$$

$$= f(\lambda) \quad \text{for any integer } k$$

Hence, it suffices to confine attention to the values of f on the interval $-\pi$ to π as if we know the values of $f(\lambda)$ $\forall\, \lambda \in [-\pi, \pi]$, we can infer the values of $f(\lambda)$ for any λ.

(4) $f(\lambda) \geq 0 \quad \forall\, \lambda$

For any positive integer N, define

$$f_N(\lambda) = \frac{1}{2\pi N} E\left(\left|\sum_{r=1}^{N} X_r e^{-ir\lambda}\right|^2\right)$$

Note that $f_N(\lambda) \geq 0 \quad \forall\, N$

$$f_N(\lambda) = \frac{1}{2\pi N} E\left(\left| \sum_{r=1}^{N} X_r e^{-ir\lambda} \right|^2 \right)$$

$$= \frac{1}{2\pi N} E\left(\sum_{r=1}^{N} X_r e^{-ir\lambda} \right)\left(\sum_{s=1}^{N} X_s e^{is\lambda} \right)$$

$$= \frac{1}{2\pi N} E\left(X_1 e^{-i\lambda} + X_2 e^{-2i\lambda} + \cdots + X_N e^{-Ni\lambda} \right)$$

$$\left(X_1 e^{i\lambda} + X_2 e^{2i\lambda} + \cdots + X_N e^{Ni\lambda} \right)$$

$$f_N(\lambda) = \frac{1}{2\pi N} \sum_{|h|<N} (N - |h|)\, e^{-ih\lambda}\, \gamma(h)$$

[You may recall a result from Lecture #

$$N V(\bar{X}_N) = \frac{1}{N} \sum_{|h|<N} (N - |h|)\, \gamma(h) \quad \text{with } \sum |\gamma(h)| < \infty$$

$$\longrightarrow \sum_{h=-\infty}^{\infty} \gamma(h) \quad \text{as } N \to \infty \;]$$

Proceeding exactly along the same lines of the proof of this result, it follows that

$$f_N(\lambda) \longrightarrow \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \gamma(h) = f(\lambda)$$

as $N \to \infty$

Thus, $f_N(\lambda) \geq 0 \quad \forall N \; \forall \lambda$

and $f_N(\lambda) \to f(\lambda)$ as $N \to \infty$

$$\Rightarrow f(\lambda) \geq 0 \quad \forall \lambda$$

(5) Inversion formula

Let $\{\gamma(h)\}$ be an absolutely summable sequence of auto covariances associated with a covariance stationary process and let the spectral density of the process $\{X_t\}$ be

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \gamma(h).$$

Then

$$\gamma(k) = \int_{-\pi}^{\pi} f(\lambda) e^{i\lambda k} \, d\lambda$$

Consider,

$$\int_{-\pi}^{\pi} e^{i\lambda k} f(\lambda) \, d\lambda = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\sum_{h=-\infty}^{\infty} \gamma(h) e^{-i\lambda h} \right) e^{i\lambda k} \, d\lambda$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{h=-\infty}^{\infty} \gamma(h) e^{i\lambda(k-h)} \, d\lambda$$

$$= \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) \int_{-\pi}^{\pi} e^{i\lambda(k-h)} \, d\lambda \quad - (*)$$

Note that

$$\int_{-\pi}^{\pi} e^{i\lambda(k-h)} \, d\lambda$$

$$= \int_{-\pi}^{\pi} \left(\cos(\lambda(k-h)) + i \sin(\lambda(k-h)) \right) d\lambda$$

$$= \begin{cases} 2\pi, & \text{if } h = k \\ 0, & \text{o/w (i.e. } h \neq k) \end{cases}$$

$\Rightarrow (*) = \gamma(k)$

Characterization of spectral density function

A real valued function $f(\cdot)$ defined on $[-\pi, \pi]$ is the spectral density of a stationary process with absolutely summable auto covariance function iff

(i) $f(\lambda) = f(-\lambda) \quad \forall \lambda$

(ii) $f(\lambda) \geq 0 \quad \forall \lambda$

& (iii) $\int_{-\pi}^{\pi} f(\lambda)\, d\lambda < \infty$

Alternate characterization of ACVF

An absolutely summable f^n $\gamma(\cdot)$ defined on the set of integers is the ACVF of a stationary process iff it is

(i) even and

(ii) $f(\lambda) = \frac{1}{2\pi} \sum_{-\infty}^{\infty} e^{-ih\lambda} \gamma(h) \geq 0 \quad \forall \lambda \in [-\pi, \pi]$

in which case $f(\lambda)$ is the corresponding spectral density function.

Spectral density of standard models

(I) White noise process

Let $X_t \sim WN(0, \sigma^2)$

$$\gamma(h) = \begin{cases} \sigma^2, & h = 0 \\ 0, & o/w \end{cases}$$

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h)\, e^{-i\lambda h} = \frac{1}{2\pi}\gamma(0) = \frac{\sigma^2}{2\pi} \quad \forall \lambda$$

Alternatively, given $f(\lambda) = \frac{\sigma^2}{2\pi}$

$$\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda} f(\lambda)\, d\lambda = \frac{\sigma^2}{2\pi} \int_{-\pi}^{\pi} e^{ih\lambda}\, d\lambda$$

$$= \frac{\sigma^2}{2\pi} \begin{cases} 2\pi, & \text{if } h = 0 \\ 0, & o/w \end{cases}$$

i.e. $\gamma(h) = \begin{cases} \sigma^2, & h = 0 \\ 0, & o/w. \end{cases}$

II AR(1) process

$$X_t = \phi X_{t-1} + \epsilon_t; \quad \epsilon_t \sim WN(0, \sigma^2)$$

$$\gamma(h) = \frac{\sigma^2}{1-\phi^2}\, \phi^{|h|}$$

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \left(\frac{\sigma^2}{1-\phi^2} \phi^{|h|} \right)$$

$$= \frac{1}{2\pi} \frac{\sigma^2}{1-\phi^2} \left(1 + \sum_{h=1}^{\infty} \phi^h e^{-ih\lambda} + \sum_{h=1}^{\infty} \phi^h e^{ih\lambda} \right)$$

$$= \frac{\sigma^2}{2\pi(1-\phi^2)} \left(1 + \frac{\phi e^{-i\lambda}}{1-\phi e^{-i\lambda}} + \frac{\phi e^{i\lambda}}{1-\phi e^{i\lambda}} \right)$$

i.e. $f(\lambda) = \frac{\sigma^2}{2\pi(1-\phi^2)}\left(\frac{1+\cancel{\phi^2}-\phi\cancel{e^{i\lambda}}-\phi\cancel{e^{i\lambda}}+\phi\cancel{e^{i\lambda}}-\cancel{\phi^2}+\phi\cancel{e^{i\lambda}}-\phi^2}{(1-\phi e^{-i\lambda})(1-\phi e^{i\lambda})}\right)$

$$= \frac{\sigma^2}{2\pi(\cancel{1-\phi^2})} \cdot \frac{(\cancel{1-\phi^2})}{(1-\phi e^{-i\lambda})(1-\phi e^{i\lambda})}$$

i.e. $f(\lambda) = \frac{\sigma^2}{2\pi} \cdot \frac{1}{(1-\phi e^{-i\lambda})(1-\phi e^{i\lambda})} = \frac{\sigma^2}{2\pi} \frac{1}{(1-2\phi\cos\lambda+\phi^2)}$.

$$f(\lambda) = \frac{\sigma^2}{2\pi} \cdot \frac{1}{\phi(e^{-i\lambda})\,\phi(e^{i\lambda})}$$

$\underline{\text{iii}}$ <u>MA(1) process</u>

$$X_t = \epsilon_t + \theta\epsilon_{t-1} \;;\; \epsilon_t \sim WN(0,\sigma^2)$$

$$\gamma(h) = \begin{cases} \sigma^2(1+\theta^2), & h=0 \\ \theta\sigma^2, & h=\pm 1 \\ 0, & o/w \end{cases}$$

$$f(\lambda) = \frac{1}{2\pi}\sum_{h=-\infty}^{\infty} e^{-ih\lambda}\gamma(h)$$

$$= \frac{1}{2\pi}\left(\sigma^2(1+\theta^2) + \sigma^2\theta e^{-i\lambda} + \sigma^2\theta e^{i\lambda}\right)$$

$$= \frac{\sigma^2}{2\pi}\left(1+\theta^2+2\theta\cos\lambda\right)$$

$$= \frac{\sigma^2}{2\pi}(1+\theta e^{-i\lambda})(1+\theta e^{i\lambda}) = \frac{\sigma^2}{2\pi}\theta(e^{-i\lambda})\,\theta(e^{i\lambda})$$

<u>Remark</u> : Note that

$$f(\lambda) = \frac{1}{2\pi}\sum_{h=-\infty}^{\infty}(e^{-i\lambda})^h\gamma(h) = \frac{1}{2\pi}\, g_X(e^{-i\lambda})$$

where $g_X(\cdot)$ is ACVF of $\{X_t\}$

$f(\cdot)$ of WN, AR(1) and MA(1) are accordingly in term of $g_X(\cdot)$

We can use the above connection between spectral density function and ACGF to obtain spectral densities of standard model for which we have earlier desired ACGF.

Spectral density $f^{\wedge}$ of AR(p)

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t \; ; \; \epsilon_t \sim WN(0, \sigma^2)$$

$$\Rightarrow \phi(B) X_t = \epsilon_t$$

$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$$

ACGF $\quad g_X(z) = \dfrac{\sigma^2}{\phi(z)\,\phi(z^{-1})}$

Spectral density

$$f_X(\lambda) = \frac{1}{2\pi} g_X(e^{-i\lambda})$$

i.e. $f_X(\lambda) = \dfrac{\sigma^2}{2\pi} \dfrac{1}{\phi(e^{-i\lambda})\,\phi(e^{i\lambda})}$

e.g. AR(1) $\quad X_t = \phi X_{t-1} + \epsilon_t$

$$f_X(\lambda) = \frac{\sigma^2}{2\pi} \frac{1}{(1-\phi e^{-i\lambda})(1-\phi e^{i\lambda})}$$

$$f_X(\lambda) = \frac{\sigma^2}{2\pi} \frac{1}{1 - \phi e^{-i\lambda} - \phi e^{i\lambda} + \phi^2}$$

Note that the spectral density f^n of AR(p) can be expressed in terms of roots of AR polynomial.

Suppose $n_1, \ldots, n_p$ are the roots of

$$n^p - \phi_1 n^{p-1} - \cdots - \phi_p = 0$$

then $$f_X(\lambda) = \frac{\sigma^2}{2\pi}\left[\prod_{j=1}^{p}(e^{i\lambda} - n_j)(e^{-i\lambda} - n_j)\right]^{-1}$$

Spectral density of MA (q)

$$Y_t = \sum_{j=0}^{q}\theta_j \epsilon_{t-j} \quad ; \; \epsilon_t \sim WN(0, \sigma^2)$$

ACGF

$$g_y(z) = \sigma^2 \, \theta(z)\, \theta(z^{-1}) \; ; \; \theta(B) = \theta_0 + \theta_1 B + \cdots + \theta_q B^q$$

$\theta_0 = 1$

Spectral density f^n

$$f_y(\lambda) = \frac{\sigma^2}{2\pi}\, \theta(e^{-i\lambda})\, \theta(e^{i\lambda})$$

$$= \frac{\sigma^2}{2\pi}\left(\sum_{j=0}^{q}\theta_j e^{-ij\lambda}\right)\left(\sum_{j=0}^{q}\theta_j e^{ij\lambda}\right)$$

if $m_1, m_2, \ldots, m_q$ are the roots of

$$m^q + \theta_1 m^{q-1} + \cdots + \theta_q = 0$$

then $$f_y(\lambda) = \frac{\sigma^2}{2\pi}\left(\prod_{j=1}^{q}(e^{-i\lambda} - m_j)(e^{i\lambda} - m_j)\right)$$

Spectral density of ARMA(p, q)

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}$$

$$\epsilon_t \sim WN(0, \sigma^2)$$

$$\phi(B) X_t = \theta(B) \epsilon_t$$

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$$

$$\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$$

ACGF : $g_X(z) = \sigma^2 \dfrac{\theta(z)\theta(z^{-1})}{\phi(z)\phi(z^{-1})}$.

Spectral density f^n :

$$f_X(\lambda) = \frac{\sigma^2}{2\pi} \cdot \frac{\theta(e^{-i\lambda})\,\theta(e^{i\lambda})}{\phi(e^{-i\lambda})\,\phi(e^{i\lambda})}.$$

If $n_1, \ldots, n_p$ are the roots of

$$n^p - \phi_1 n^{p-1} - \cdots - \phi_p = 0$$

& $m_1, \ldots, m_q$ are the roots of

$$m^q + \theta_1 m^{q-1} + \cdots + \theta_q = 0$$

then

$$f_X(\lambda) = \frac{\sigma^2}{2\pi} \frac{\left(\prod_{j=1}^{q} (e^{i\lambda} - m_j)(e^{-i\lambda} - m_j)\right)}{\left(\prod_{j=1}^{q} (e^{i\lambda} - n_j)(e^{-i\lambda} - n_j)\right)}.$$

Result: Suppose $\{X_t\}$ be a zero mean stationary process with absolutely summable ACVF $\gamma_X(\cdot)$ and $\{a_j\}_{j=-\infty}^{\infty}$ be an absolutely summable seq, then the spectral density f^n of the filtered process

$$Y_t = \sum_{j=-\infty}^{\infty} a_j X_{t-j} \quad \text{is given by}$$

$$f_Y(\lambda) = (2\pi)^2 f_X(\lambda) \, f_a(\lambda) \, f_a^*(\lambda)$$

where, $f_X(\cdot)$: spectral density of input process $\{X_t\}$

$$f_a(\lambda) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} a_j e^{-i\lambda j}$$

$f_a^*(\lambda)$: conjugate of $f_a(\lambda)$

Pf: $E\,Y_t = 0$

$$\begin{aligned}
\gamma_Y(h) &= E\,Y_t Y_{t+h} \\
&= E\left(\sum_{j=-\infty}^{\infty} a_j X_{t-j}\right)\left(\sum_{k=-\infty}^{\infty} a_k X_{t+h-k}\right) \\
&= E\left(\sum_j \sum_k a_j a_k X_{t-j} X_{t+h-k}\right) \\
&= \sum_j \sum_k a_j a_k E\left(X_{t-j} X_{t+h-k}\right)
\end{aligned}$$

i.e. $\gamma_y(h) = \sum_j \sum_k a_j a_k \gamma_x(h-k+j)$

$$f_y(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \gamma_y(h)$$

$$= \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \left(\sum_j \sum_k a_j a_k \gamma_x(h-k+j) \right)$$

$$= \frac{1}{2\pi} \sum_j \sum_k a_j a_k \sum_h e^{-ih\lambda} \gamma_x(h-k+j)$$

$$= \frac{1}{2\pi} \sum_j \sum_k a_j a_k e^{-i\lambda k} e^{ij\lambda} \sum_h e^{-i\lambda(h-k+j)} \gamma_x(h-k+j)$$

$$= \frac{1}{2\pi} \sum_j \sum_k a_j a_k e^{-ik\lambda} e^{ij\lambda} \sum_{h'=-\infty}^{\infty} e^{-i\lambda h'} \gamma_x(h')$$

$h' = h-k+j)$

$$= \sum_j a_j e^{ij\lambda} \sum_k a_k e^{-ik\lambda} \frac{1}{2\pi} \sum_{h'=-\infty}^{\infty} e^{-ih'\lambda} \gamma_x(h')$$

i.e. $f_y(\lambda) = (2\pi)^2 f_a(\lambda) f_a^*(\lambda) f_x(\lambda)$

Remark: The above result can be used to derived spectral densities of AR, MA, ARMA processes.

Remark: $\sum_k a_k e^{-ik\lambda}$ is called transfer f^n of the filter with $\{a_k\}$ as filter coeffs

$\left|\sum_k a_k e^{-ik\lambda}\right|^2$ is called power transfer f^n of the filter.

Spectral density of MA(q) thro filtering result

$$X_t = \sum_{j=0}^{q} \theta_j \epsilon_{t-j} \quad ; \quad \epsilon_t \sim WN(0, \sigma^2)$$

$\{\epsilon_t\}$ is the input seq & $\{X_t\}$ is the output process after linear filtering.

$$\Rightarrow \quad f_X(\lambda) = (2\pi)^2 f_\theta(\lambda) f_\theta^*(\lambda) \; f_\epsilon(\lambda)$$

$$= \left(\sum_{j=0}^{q} \theta_j e^{-ij\lambda}\right)\left(\sum_{j=0}^{q} \theta_j e^{ij\lambda}\right) \frac{\sigma^2}{2\pi}$$

i.e. $f_X(\lambda) = \dfrac{\sigma^2}{2\pi} \theta(e^{-i\lambda}) \, \theta(e^{i\lambda})$

as obtained earlier.

MA(α) : $\quad X_t = \sum_{j=0}^{\alpha} \theta_j \epsilon_{t-j}$

$$f_X(\lambda) = \frac{\sigma^2}{2\pi}\left(\sum_0^{\alpha} \theta_j e^{-ij\lambda}\right)\left(\sum_0^{\alpha} \theta_j e^{ij\lambda}\right)$$

$$= \frac{\sigma^2}{2\pi} \theta(e^{-i\lambda}) \quad \theta(e^{i\lambda})$$

Spectral density of AR(p) thro filtering result

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t \; ; \; \epsilon_t \sim WN(0, \sigma^2)$$

$$\phi(B) X_t = \epsilon_t \quad ; \quad \phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$$

$$\epsilon_t = \sum_{j=0}^{p} \tilde{\phi}_j X_{t-j} \; ;$$

with $\tilde{\phi}_0 = 1$, $\tilde{\phi}_j = -\phi_j \;\; \forall j \geq 1$

using the filtering result

$$f_\epsilon(\lambda) = f_X(\lambda) \left(\sum_0^{p} \tilde{\phi}_j e^{ij\lambda}\right)\left(\sum_0^{p} \tilde{\phi}_j e^{-ij\lambda}\right)$$

i.e. $\dfrac{\sigma^2}{2\pi} = f_X(\lambda)\left(\sum_0^p \tilde{\phi}_j e^{ij\lambda}\right)\left(\sum_0^p \tilde{\phi}_j e^{-ij\lambda}\right)$

$$\Rightarrow \quad f_X(\lambda) = \frac{\sigma^2}{2\pi}\left[\left(\sum_0^p \tilde{\phi}_j e^{i\lambda j}\right)\left(\sum_0^p \tilde{\phi}_j e^{-ij\lambda}\right)\right]^{-1}$$

i.e. $f_X(\lambda) = \dfrac{\sigma^2}{2\pi} \cdot \dfrac{1}{\phi(e^{i\lambda})\,\phi(e^{-i\lambda})}$

Spectral density of ARMA(p,q) thro filtering argument

Let the ARMA(p,q) model be

$$\sum_0^p \phi_j X_{t-j} = \sum_{j=0}^q \theta_j \epsilon_{t-j} \; ; \; \epsilon_t \sim WN(0,\sigma^2)$$

$$\phi(B) X_t = \theta(B)\epsilon_t$$

Spectral density of l.h.s

$$f_X(\lambda)\left(\sum_0^p \phi_j e^{ij\lambda}\right)\left(\sum_0^p \phi_j e^{-ij\lambda}\right)$$

spectral density of r.h.s.

$$f_\epsilon(\lambda)\left(\sum_0^q \theta_j e^{ij\lambda}\right)\left(\sum_0^q \theta_j e^{-ij\lambda}\right)$$

$$\Rightarrow \quad f_X(\lambda) = \frac{\sigma^2}{2\pi}\,\frac{\left(\sum_0^q \theta_j e^{ij\lambda}\right)\left(\sum_0^q \theta_j e^{-ij\lambda}\right)}{\left(\sum_0^p \phi_j e^{ij\lambda}\right)\left(\sum_0^p \phi_j e^{-ij\lambda}\right)}$$

i.e. $f_X(\lambda) = \dfrac{\sigma^2}{2\pi}\,\dfrac{\theta(e^{i\lambda})\,\theta(e^{-i\lambda})}{\phi(e^{i\lambda})\,\phi(e^{-i\lambda})}$

Estimation of spectral density

Non-parametric approach:

$$\hat{f}_X(\lambda) = \frac{1}{2\pi} \sum_{h=-(n-1)}^{n-1} e^{-ih\lambda} \hat{\gamma}(h)$$

where, $$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (X_t - \bar{X}_n)(X_{t+h} - \bar{X}_n)$$

Parametric approach:

Suppose $\{X_t\}$ is a Gaussian, stationary and invertible ARMA(p, q)

$$f_X(\lambda) = \frac{\sigma^2}{2\pi} \frac{\left(\sum_{j=0}^{q} \theta_j e^{-ij\lambda}\right)\left(\sum_{j=0}^{q} \theta_j e^{ij\lambda}\right)}{\left(\sum_{j=0}^{p} \phi_j e^{-ij\lambda}\right)\left(\sum_{j=0}^{p} \phi_j e^{ij\lambda}\right)}.$$

Obtain parameter estimates (say MLE under Gaussian) as $\hat{\phi}_1, \ldots, \hat{\phi}_p, \hat{\theta}_1, \ldots, \hat{\theta}_q$ and $\hat{\sigma}^2$ and hence

$$\hat{f}_X(\lambda) = \frac{\hat{\sigma}^2}{2\pi} \frac{\left(\sum_{j=0}^{q} \hat{\theta}_j e^{-ij\lambda}\right)\left(\sum_{j=0}^{q} \hat{\theta}_j e^{ij\lambda}\right)}{\left(\sum_{j=0}^{p} \hat{\phi}_j e^{-ij\lambda}\right)\left(\sum_{j=0}^{p} \hat{\phi}_j e^{ij\lambda}\right)}$$

Spectral distribution function

Spectral representation theorem: A function $\gamma(\cdot)$ defined on the set of integers is the ACVF of a stationary process iff there exists a f^n $F(\cdot)$ which is right continuous, non decreasing, bounded on $[-\pi, \pi]$ with $F(-\pi) = 0$ $\ni$

$$\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda}\, dF(\lambda) \quad \forall \text{ integer } h$$

<u>Note</u>: Spectral representation theorem ensures that there exists such an F for every stationary process. $F(\cdot)$ is called the spectral distribution function.

<u>Note</u>: F defined through the spectral representation theorem is a generalized distribution function on $[-\pi, \pi]$ in the sense that

$$G(\lambda) = \frac{F(\lambda)}{F(\pi)}$$ is a proper distribution function on $[-\pi, \pi]$.

<u>Note</u>: Note that $F(\pi) = \gamma(0) = V(X_1)$; from

$$\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda}\, dF(\lambda)$$

we have

$$\frac{\gamma(h)}{F(\pi)} = \frac{\gamma(h)}{\gamma(0)} = \rho(h) = \int_{-\pi}^{\pi} e^{ih\lambda}\, d\left(\frac{F(\lambda)}{F(\pi)}\right)$$

i.e. $$\rho(h) = \int_{-\pi}^{\pi} e^{ih\lambda}\, dG(\lambda)$$

The above is the spectral representation of $\rho(h)$.

Note: If $F(\lambda)$ is $\ni$ it can be expressed as.

$$F(\lambda) = \int_{-\pi}^{\lambda} f(y)\,dy \quad \forall \lambda \in [-\pi, \pi] \; : f(\cdot) \geq 0$$

i.e. $F(\cdot)$ is a generalised continuous dist^n f^n

(in other words $G(\cdot)$ is dist^n f^n corresponding to a continuous random variable)

$f(\cdot)$ in the above representation is the spectral density function and the associated time series is said to have a continuous spectrum

Note: If $F(\cdot)$ is a generalized discrete distribution function (i.e. $G(\cdot)$ is a proper ~~dist~~ discrete distribution function), increasing only by jumps, then the associated time series is said to have a discrete spectrum

Note: If $F(\cdot)$ is a generalized mixed distribution function ($G(\cdot)$ is a proper mixed distribution function), then we have a mixed spectrum; i.e. time series will have a continuous spectrum part and a discrete spectrum part.

Example 1 : Discrete spectrum

$$X_t = A\cos\omega t + B\sin\omega t$$

A & B are uncorrelated random variables with mean 0 and variance 1; $\omega \in (0,\pi)$ is fixed

$\gamma_X(h) = \cos(\omega h)$ is not absolutely summable and hence we can't talk about spectral density

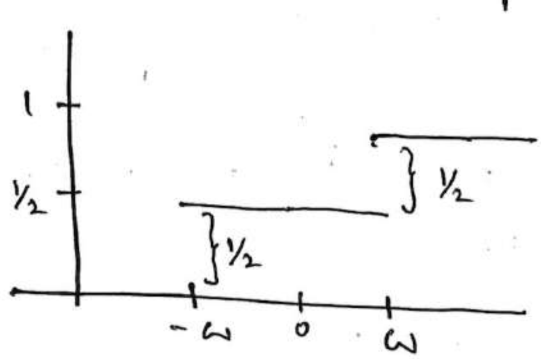
By spectral representation theorem,

$$\gamma(h) = \cos(\omega h) = \int_{-\pi}^{\pi} e^{ih\lambda}\, dF(\lambda)$$

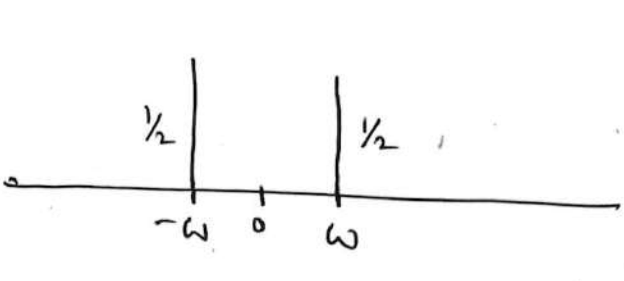
$$= \frac{1}{2}e^{-ih\omega} + \frac{1}{2}e^{ih\omega}$$

$$\Rightarrow F(\lambda) = \begin{cases} 0, & \lambda < -\omega \\ \frac{1}{2}, & -\omega \le \lambda < \omega \\ 1, & \lambda \ge \omega \end{cases} \leftarrow \text{spectral dist}^n\ f^n \text{ of } \{X_t\}$$

$F(\pi) = 1$; $F(\cdot)$ is a proper dist^n f^n



Discrete spectrum

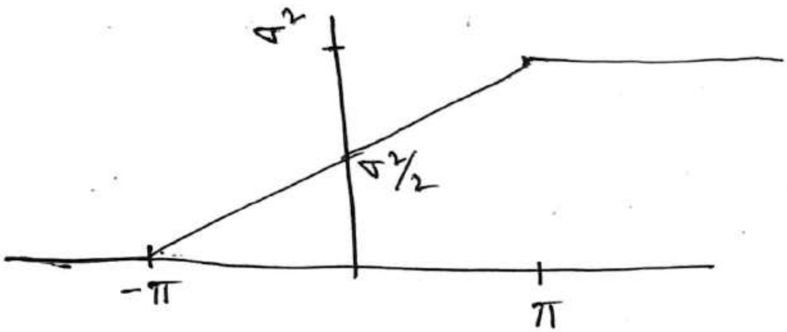


<u>Example 2</u> Continuous spectrum

$X_t \sim WN(0, \sigma^2)$

$$f_X(\lambda) = \frac{\sigma^2}{2\pi} \quad \forall \lambda \in [-\pi, \pi]$$

$$F_X(\lambda) = \int_{-\pi}^{\lambda} \frac{\sigma^2}{2\pi} \, d\lambda = \frac{\sigma^2}{2\pi}(\lambda + \pi)$$



$F_X(\lambda)$ is the spectral distⁿ fⁿ of $\{X_t\}$.

<u>Note</u>: If $\{X_t\}$ and $\{Y_t\}$ are uncorrelated stationary processes with ACVF $\gamma_X(\cdot)$ and $\gamma_Y(\cdot)$ and spectral distⁿ functions $F_X(\cdot)$ and $F_Y(\cdot)$ then spectral distⁿ function of $Z_t = X_t + Y_t$ is $F_Z(\lambda) = F_X(\lambda) + F_Y(\lambda)$.