Some operators:

1. Back shift operator: (B)

$$B^h X_t = X_{t-h} \qquad h \in \mathbb{Z}$$

eg: $B^2 X_{10} = X_8$

2. Difference operator : (∇)

$\nabla \equiv I - B$ where I stands for identity operator.

$$\nabla X_t \equiv (I - B) X_t = X_t - X_{t-1}$$

$$\Rightarrow \nabla^h X_t \equiv (I-B)^h X_t \equiv \sum_{k=0}^{h} \binom{h}{k} (-1)^{h-k} (B)^{h-k} X_t.$$

$X_1\ X_2\ X_3\ X_4\ X_5\ X_6.$

$\nabla \rightarrow \quad \nabla X_2\ \nabla X_3\ \nabla X_4\ \nabla_5\ \nabla X_6$

$\nabla^2 \rightarrow \quad \nabla^2 X_3\ \nabla^2 X_4\ \nabla^2 X_5\ \nabla^2 X_6.$

$$\nabla^2 X_3 = \nabla X_3 - \nabla X_2 = X_3 - X_2 - X_2 + X_1 = X_3 - 2X_2 + X_1$$

$$= (I - 2B + B^2) X_3 = (I - B)^2 X_3.$$

(3) Seasonal difference : (∇_S)

$$\nabla_S = (I - B^S) \neq (I - B)^S = \nabla^S \quad \text{in general.}$$

$x_1\ x_2\ x_3\ x_4 \ \vdots\ x_5\ x_6\ x_7\ x_8 \ \vdots\ x_9\ x_{10}\ x_{11}\ x_{12} \ \vdots\ x_{13}\ x_{14}\ x_{15}\ x_{16} \ \vdots \ \ldots$

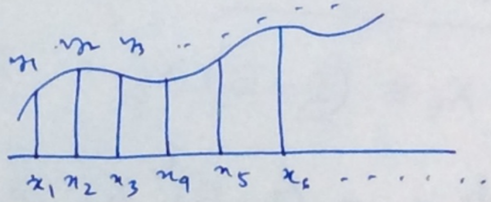
$$\nabla_4 X_6 = (I - B^4) X_6 = X_6 - X_2 .$$

application: is to remove seasonal effect from the data. or identify the seasonality

(4)

$x_1\ \ x_2\ \ x_3\ \ x_4\ \ x_5\ \ \ldots\ldots$

$f(x) \rightarrow y_1\ \ y_2\ \ y_3\ \ y_4\ \ y_5\ \ \ldots\ldots$



devided difference: $\dfrac{\nabla Y_2}{\nabla X_2}\ ,\ \dfrac{\nabla Y_3}{\nabla X_3}\ ,\ \dfrac{\nabla Y_4}{\nabla X_4}\ \ldots\ldots$

if $\nabla X_i = X_i - X_{i-1} = h$ common differences are same.

$$\frac{\nabla Y_i}{\nabla X_i} = \frac{Y_i - Y_{i-1}}{h} = \frac{f(x_i) - f(x_{i-1})}{h} = \frac{f(x_{i-1}+h) - f(x_{i-1})}{h}.$$

let $h \rightarrow 0$. and $h > 0$ it will approximate right side derivative.

<u>Linear process:</u> A time series is said to be a linear process if it has following representation.

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}$$

$\forall t \in \mathbb{Z}$, $Z_t \sim WN(0, \sigma^2)$, $\{\psi_j\}$ are <u>absolutely summable.</u>

i.e. $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j B^j Z_t = \mu + \left(\sum_{j=-\infty}^{\infty} \psi_j B^j\right) Z_t = \mu + \Psi(B) Z_t$$

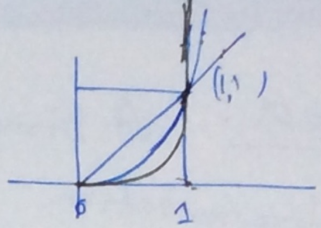
where $\Psi(B) = \sum_{j=-\infty}^{\infty} \psi_j B^j$

Here $\{X_t\}$ is a linear <u>process.</u>

why do we need $\{\psi_j\}$ ~~to~~ to be absolutely summable?

If we do not have an absolutely summable series then different rearrangements of the series may lead to different limits with sometimes even may not exist also.

$\log_e 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} \cdots\cdots$



but $\left(1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} \cdots\cdots\right) - \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \cdots\cdots\right)$ limit does not exists.

$\downarrow \infty \qquad \downarrow \infty$

$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} \cdots\cdots \longrightarrow \infty$.

$\sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$.

Let $\{X_t\}$ be a linear process. Find its expectation if exists. ?

$Z_t \sim WN$

$E(Z_t) = 0$

$V(Z_t) = \sigma^2$

$\Rightarrow E(Z_t^2) = \sigma^2$

$\Rightarrow E(|Z_t|^2) = \sigma^2 < \infty$

$\Rightarrow E|Z_t| < \infty$

$$E|X_t| = E\left|\mu + \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}\right|$$

$$\leq E\left(|\mu| + \left|\sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}\right|\right) \quad \text{triangular inequity}$$

$$\leq |\mu| + E\left(\sum_{j=-\infty}^{\infty} |\psi_j| |Z_{t-j}|\right) \quad \text{triangular inequity}$$

$$= |\mu| + \sum_{j=-\infty}^{\infty} |\psi_j| \left(E|Z_{t-j}|\right) \quad \text{sum law of expectation.}$$

$$\leq |\mu| + k \sum_{j=-\infty}^{\infty} |\psi_j| < \infty.$$

$$E(X_t) = \mu + \left(E \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}\right) = \mu. < \infty$$

If $E(|X|^r) < \infty$

then $E(|X|^s) < \infty$ when $0 < s \leq r$

HW

Th. Let $\{X_t\}$ be a (weakly) stationary time series with $E(X_t) = 0 \;\forall t$.

and $\gamma_X(h) = \operatorname{Cov}(X_t, X_{t+h})$ exists. $\{ \Leftrightarrow E|X_t X_{t+h}| < \infty \}$.

If $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ then $Y_t = \sum_{j=-\infty}^{\infty} \psi_j X_{t-j}$ has

covariance as $\gamma_Y(h) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k \gamma_X(h+k-j)$

and $E(Y_t) = 0 \;\forall t$.

HW $X_t \equiv WN(0, \sigma^2)$.

$Y_t \sim AR(1)$. find $\gamma_Y(h)$

$$E(|Y_t|) = E\left|\sum_{j=-\infty}^{\infty} \psi_j X_{t-j}\right| \le \sum_{j=-\infty}^{\infty} |\psi_j| \, E|X_{t-j}| < \infty$$

Hence $E(Y_t) = E\left(\sum_{j=-\infty}^{\infty} \psi_j X_{t-j}\right) = 0.$

te. $\gamma_Y(h) = \text{cov}(Y_t, Y_{t+h}) = E(Y_t Y_{t+h}).$

$$E|Y_t Y_{t+h}| = E\left|\left(\sum_{j=-\infty}^{\infty} \psi_j X_{t-j}, \sum_{k=-\infty}^{\infty} \psi_k X_{t+h-k}\right)\right|$$

$$= E\left|\left(\sum_{j=-\infty}^{\infty}\sum_{k=-\infty}^{\infty} \psi_j \psi_k X_{t-j} X_{t+h-k}\right)\right|$$

$$\le \sum_{j=-\infty}^{\infty}\sum_{k=-\infty}^{\infty} |\psi_j \psi_k| \, E|X_{t-j}||X_{t+h-k}|$$ as. $E|X_t X_{t+h}| < \infty$

$$\le M \sum_{j=-\infty}^{\infty}\sum_{k=-\infty}^{\infty} |\psi_j||\psi_k| \le M \sum_{j=-\infty}^{\infty} |\psi_j| \sum_{j=-\infty}^{\infty} |\psi_k|$$

all are finite.

M is $\max_h \gamma_X(h)$.

$$\gamma_Y(h) = \operatorname{cov}(Y_{t+h}, Y_t)$$

$$= \operatorname{cov}\left(\sum_{j=-\infty}^{\infty} \psi_j X_{t+h-j}, \sum_{k=-\infty}^{\infty} \psi_k X_{t-k}\right)$$

$$= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k \operatorname{cov}(X_{t+h-j}, X_{t-k})$$

$$= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k \gamma_X(h-j+k)$$

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

$X_t \sim AR(1)$, $AR(2)$, $ARMA(1,1)$ ...

$Z_t \sim MA(1)$, $MA(2)$.

MA(q) process when $q \in \mathbb{N}$. (Moving average of order q).

MA(q) is a linear process with $\psi_0 = 1$.

$$\psi_j = \begin{cases} \theta_j & 1 \le j \le q \\ 0 & \text{o.w.} \end{cases}$$

$$X_t = W_t + \sum_{j=1}^{q} \theta_j W_{t-j}$$

$$= \left(I + \sum_{j=1}^{q} \theta_j B^j \right) W_t = \Theta_q(B) W_t.$$

where $W_t \sim WN(0, \sigma^2)$.

Where

$$\Theta_q(z) = 1 + \theta_1 z + \theta_2 z^2 \dots + \theta_q z^q$$

X_t is weakly stationary.

$\text{cov}(X_t, X_s) = 0$ if $|t-s| > q$.

As a particular case. X_t and X_s are independent if for $|t-s| > q$.

W_ts iid or independent white noise.

Auto regressive process of order (p). AR(p).

$$W_t \sim WN(0, \sigma^2).$$

$$X_t = \sum_{j=1}^{p} \varphi_j X_{t-j} + W_t.$$

$$\Rightarrow \left(X_t - \sum_{j=1}^{p} \varphi_j X_{t-j}\right) = W_t.$$

$$\Rightarrow \left(I - \sum_{j=1}^{p} \varphi_j B^j\right) X_t = W_t.$$

$$\underline{\Phi_p = 1 - \sum_{j=1}^{p} \varphi_j z^j}$$

$$\Rightarrow \Phi_p(B)\, X_t = W_t.$$

$$\Rightarrow X_t = \frac{1}{\Phi_p(B)} W_t.$$

AR(1) with $\varphi \neq 0$ and $|\varphi| < 1$ is weakly stationary.

AR(1) is a linear process.

AR(1) is equivalent to <u>MA(∞) process</u>. (To prove).

$X_t = \phi X_{t-1} + W_t \quad \phi \neq 0 \; |\phi| < 1$

$\Rightarrow X_t - \phi B X_t = W_t.$

$\Rightarrow X_t = \left(\frac{1}{1 - \phi B} \right) W_t.$

$\Rightarrow X_t = \sum_{j=0}^{\infty} (\phi B)^j W_t.$

$\Rightarrow \cancel{X_t} = \sum_{j=0}^{\infty} \cancel{\phi^j} \cancel{W}_t$

$\Rightarrow X_t = \sum_{j=0}^{\infty} \phi^j W_{t-j} \rightarrow$ AR(1) process.

$X \overset{d}{=} Y$ "x is distributionally same as Y"

$\Rightarrow F_X(a) = F_Y(a) \quad \forall a. \Leftrightarrow$ X and Y have same c.d.f.

Ex. $X \sim N(0,1)$ $\qquad X \overset{d}{=} Y$ but $\underline{\underline{X \neq Y}}$

$N(0,1)$

(11)

$X \sim U(0,1)$

$Y = 1 - X \sim U(0,1)$ $\quad X \overset{d}{=} Y$ but $X \neq Y$.

MGF, direct pdf/cdf, L_2 convergence. Can be use to show the equality of distribution.

L_2 convergence ✓ $\Rightarrow$ L_1 convergence ✓ (convergence in prob) $\Rightarrow$ convergence in distribution ✓

$\not\Leftarrow$ $\qquad \not\Leftarrow$

$$X_t = \varphi X_{t-1} + w_t.$$
$$= \varphi (w_{t-1} + \varphi X_{t-2}) + w_t.$$
$$= w_t + \varphi w_{t-1} + \varphi^2 X_{t-2}$$
$$= w_t + \varphi w_{t-1} + \varphi^2 w_{t-2} + \varphi^3 X_{t-3}.$$
$$= w_t + \varphi w_{t-1} + \varphi^2 w_{t-2}$$
$$\vdots$$
$$= \left(\sum_{j=0}^{k} \varphi^j w_{t-j}\right) + \varphi^{k+1} X_{t-(k+1)}.$$

$$\left(X_t - \sum_{j=0}^{k} \varphi^j w_{t-j}\right) = \varphi^{k+1} X_{t-(k+1)}.$$

$$E\left[\left(X_t - \sum_{j=0}^{k} \varphi^j Z_{t-j}\right)^2\right] = E\left(\varphi^{2(k+1)} X^2_{t-(k+1)}\right).$$

$$\lim_{k\uparrow\infty} E\left[\left(X_t - \sum_{j=0}^{k} \varphi^j Z_{t-j}\right)^2\right] = \lim_{k\uparrow\infty} \varphi^{2k+2} E\left(X^2_{t-(k+1)}\right) \to 0$$

as $\underline{|\varphi|<1}$

$$\lim_{k\uparrow\infty} \sum_{j=0}^{k} \varphi^j Z_{t-j} \xrightarrow{d} X_t .$$

$MA(\infty)$

$\underline{\underline{AR(1)}}$.

L_2 convergence $\Rightarrow$ L_1 convergence $\xrightarrow{\text{as } n\uparrow\infty}$ $\Rightarrow$ Convergence in dis

Convergence in prob.

$$E(X_n - Y)^2 \to 0 \Rightarrow E|X_n - Y| \to 0 \Rightarrow X_n \xrightarrow{d} Y$$

$$\Updownarrow \qquad\qquad \Updownarrow$$

$$P(|X_n - Y| > \epsilon) \to 0 \qquad F_{X_n}(a) = F_Y(a) \ \forall a$$

where F_Y is continuous.

$$P(X = c) = 1$$

ARMA(p, q)

A time series $\{X_t\}$ is said to follow ARMA(p, q) process if it is weakly stationary and it satisfies

$$X_t - \sum_{j=1}^{p} \phi_j X_{t-j} = Z_t + \sum_{j=1}^{q} \theta_j Z_{t-j} \quad \forall t, \quad Z \sim WN$$

$$\Rightarrow \left(I - \sum_{j=1}^{p} \phi_j B^j\right) X_t = \left(I + \sum_{j=1}^{q} \theta_j B^j\right) Z_t$$

$$\Rightarrow \Phi_p(B) X_t = \Theta_q(B) Z_t.$$

$\Phi_p(n)$ and $\Theta_q(n)$ don't have common root.

ARMA(1,1)

$$X_t - \varphi X_{t-1} = Z_t + \theta Z_{t-1}$$

$|\varphi| < 1$ $\varphi \neq 0$.
$\varphi + \theta \neq 0$

$Z_t \sim WN(0, \sigma^2)$.

$\Rightarrow$ $$(I - \varphi B) X_t = (I + \theta B) Z_t.$$

$\Rightarrow$ $$X_t = \frac{(I + \theta B)}{(I - \varphi B)} Z_t.$$

$\Rightarrow$ $$X_t = (1 + \theta B)\left(\sum_{j=0}^{\infty} (\varphi B)^j\right) Z_t.$$

Hence X_t is a linear process.

$$\equiv \Psi(B) \, Z_t$$

$\gamma(0) =$
$\gamma(1) =$
$\gamma(h) =$

$$\begin{cases} \psi_0 = 1 \\ \psi_j = (\theta + \varphi)\varphi^{j-1} \quad \forall \, j \geq 1 \end{cases}$$

$$X_t = Z_t + (\theta + \varphi) \sum_{j=1}^{\infty} \varphi^{j-1} Z_{t-j}$$

$$\gamma(0) = \sigma^2 \sum_{j=0}^{\infty} \psi_j^2$$

$$\gamma(h) = E(X_{t+h} X_t) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+|h|}$$

$$= \sigma^2 \left[1 + (\theta+\varphi)^2 \sum_{j=1}^{\infty} \varphi^{2j-2} \right]$$

$$= \sigma^2 \left[1 + \frac{(\theta+\varphi)^2}{1-\varphi^2} \right]$$

$$\gamma(1) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+1}$$

$$= \sigma^2 \left[(\theta+\varphi) + \frac{(\theta+\varphi)^2 \varphi}{1-\varphi^2} \right]$$

$$\gamma(h) = \varphi^{h-1} \gamma(1) \quad \underline{\text{(H.W)}}$$

r MA(q) process.

$\gamma(h) \neq 0$ for $\underline{|h| \leq q}$ $\qquad \underline{\gamma(h) = \sigma^2 \sum_{j=0}^{q-|h|} \theta_j \theta_{j+|h|}}$

0 otherwise.

For ~~AR~~ AR (p) process.

$\gamma(h) =$ it decays exponentially depending on the value of φ_j's.

For ARMA process we also see the similar ~~th~~ property of $\gamma(h)$.

(ACF) Auto correlation function ~~ρ_x~~ $\frac{\gamma_x(h)}{\gamma_x(0)} = \rho_x(h)$.

Can be used as a classifier between MA and <u>AR or ARMA</u>.

Partial Correlation Coefficient.

$$\begin{pmatrix} Y \\ Z \\ \hdashline \underset{\sim}{X} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_y \\ \mu_z \\ \hdashline \underset{\sim}{\mu_x} \end{pmatrix}, \begin{pmatrix} \Sigma_{yz} & \Sigma_{yzx}^T \\ \Sigma_{yzx} & \Sigma_x \end{pmatrix} \right)$$

$E(Y) = \mu_y$

$E(Z) = \mu_z$

$E(\underset{\sim}{X}) = \underset{\sim}{\mu_x}$

$$D\begin{pmatrix} Y \\ Z \end{pmatrix} = \begin{pmatrix} \sigma_{yy} & \sigma_{yz} \\ \sigma_{yz} & \sigma_{zz} \end{pmatrix} = \underset{2\times 2}{\Sigma_{yz}}.$$

$$\underset{\sim}{\sigma_{yx}} = \operatorname{cov}(Y, \underset{\sim}{X}) = \begin{pmatrix} \sigma_{yx_1} \\ \sigma_{yx_2} \\ \vdots \\ \sigma_{yx_p} \end{pmatrix} \qquad \underset{\sim}{\sigma_{zx}} = \operatorname{cov}(Z, \underset{\sim}{X}) = \begin{pmatrix} \sigma_{zx_1} \\ \sigma_{zx_2} \\ \vdots \\ \sigma_{zx_p} \end{pmatrix}.$$

$$\underset{p\times 2}{\Sigma_{yzx}} = (\underset{\sim}{\sigma_{yx}} \; \underset{\sim}{\sigma_{zx}}) \qquad \underset{2\times p}{\Sigma_{yzx}^T} = \begin{pmatrix} \underset{\sim}{\sigma_{yx}}^T \\ \underset{\sim}{\sigma_{zx}}^T \end{pmatrix}. \qquad \Sigma_x \text{ is } p\times p.$$

Partial correlation coefficient between Y and Z after removing the impact of $\underset{\sim}{X}$. is defined as.

$$\rho_{yz\cdot \underset{\sim}{x}} = \frac{\operatorname{cov}\left(Y - E(Y|\underset{\sim}{x}),\; Z - E(Z|\underset{\sim}{x})\right)}{\sqrt{V\left(Y - E(Y|\underset{\sim}{x})\right) V\left(Z - E(Z|\underset{\sim}{x})\right)}}.$$

$$= \frac{\operatorname{cov}\left(e_{Y|X}, e_{Z|X}\right)}{\sqrt{V(e_{Y|X})\, V(e_{Z|X})}} = \frac{\sigma_{yz\cdot x}}{\sqrt{\sigma_{yy\cdot x}\, \sigma_{zz\cdot x}}}.$$

The result will hold even for non-normal distribution under the <u>best-linear prediction</u> (instead of $E(Y|X)$ and $E(Z|X)$.) for welldefined mean vector and covariance matrix. <u>except the distribution</u>.

For multivariate Normal distribution as defined above.

$$\begin{pmatrix} Y \\ Z \end{pmatrix} \Big| X = \underset{\sim}{x} \sim N\left(\begin{pmatrix} \mu_y \\ \mu_z \end{pmatrix} + \Sigma_{yzx}^{T} \Sigma_x^{-1} (\underset{\sim}{x} - \underset{\sim}{\mu_x}), \; \Sigma_{yz} - \Sigma_{yzx}^{T} \Sigma_x^{-1} \Sigma_{yzx} \right).$$

In particular $\begin{pmatrix} U \\ V \end{pmatrix} \sim N(\mu_u, \mu_v, \sigma_u^2, \sigma_v^2, \rho)$

$$V | U = u \sim N\left(\mu_v + \rho \frac{\sigma_v}{\sigma_u}(u - \mu_u), \; (1-\rho^2)\sigma_v^2\right)$$

$$\rho \frac{\sigma_v}{\sigma_u} = \frac{\rho \sigma_v \sigma_u}{\sigma_u^2} = \mathrm{cov}(U,V)(\mathrm{Var}(U))^{-1}$$

$$\sigma_v^2 - \rho^2 \sigma_v^2 = \sigma_v^2 - (\rho \sigma_u \sigma_v)(\sigma_u^2)^{-1}(\rho \sigma_u \sigma_v).$$

For a non-normal distribution the conditional distribution will NOT be normal. But the Variance - Covariance structure will be same for best linear prediction.

Error part:

$$\begin{pmatrix} Y \\ Z \end{pmatrix} - E\left(\begin{pmatrix} Y \\ Z \end{pmatrix} \middle| X = x\right) = \begin{pmatrix} Y \\ Z \end{pmatrix} - \left[\begin{pmatrix} \mu_y \\ \mu_z \end{pmatrix} + \Sigma_{YZx}^T \Sigma_x^{-1} (\underset{\sim}{x} - \underset{\sim}{\mu_x})\right]$$

$$\sim N\left(\underset{\sim}{0}, \; \Sigma_{YZ} - \Sigma_{YZx}^T \Sigma_x^{-1} \Sigma_{YZx}\right)$$

$$\mathrm{cov}(Y - E(Y|X), \; Z - E(Z|X)) = \mathrm{cov}(e_{Y|X}, e_{Z|X})$$

$$= \sigma_{YZ} - \underset{\sim}{\sigma}_{YX}^T \Sigma_X^{-1} \underset{\sim}{\sigma}_{ZX} = \sigma_{YZ} - \underset{\sim}{\sigma}_{ZX}^T \Sigma_x^{-1} \underset{\sim}{\sigma}_{Yx} = \sigma_{YZ\cdot x}$$

$$\mathrm{Var}(Y - E(Y|X)) = \sigma_{YY} - \underset{\sim}{\sigma}_{Yx}^T \Sigma_X^{-1} \underset{\sim}{\sigma}_{Yx} = \sigma_{YY\cdot x}$$

$$\mathrm{Var}(Z - E(Z|X)) = \sigma_{ZZ} - \underset{\sim}{\sigma}_{Zx}^T \Sigma_X^{-1} \underset{\sim}{\sigma}_{Zx} = \sigma_{ZZ\cdot x}$$

$$\rho_{YZ\cdot \underset{\sim}{x}} = \frac{\sigma_{YZ\cdot x}}{\sqrt{\sigma_{YY\cdot x}\, \sigma_{ZZ\cdot x}}} = \frac{\sigma_{YZ} - \underset{\sim}{\sigma}_{Yx}^T \Sigma_X^{-1} \underset{\sim}{\sigma}_{Zx}}{\sqrt{(\sigma_{YY} - \underset{\sim}{\sigma}_{Yx}^T \Sigma_X^{-1} \underset{\sim}{\sigma}_{Yx})(\sigma_{ZZ} - \underset{\sim}{\sigma}_{Zx}^T \Sigma_x^{-1} \underset{\sim}{\sigma}_{Zx})}}$$

Partial correlation coefficient will not change even though we ~~rearran~~ rearrange $(Y\, Z\, \underset{\sim}{X})$ to $(Y, \underset{\sim}{X}, Z)$ as follows.

$$\begin{pmatrix} Y \\ \cdots \\ \underset{\sim}{X} \\ \cdots \\ Z \end{pmatrix} \quad E\begin{pmatrix} Y \\ \underset{\sim}{X} \\ Z \end{pmatrix} = \begin{pmatrix} \mu_y \\ \underset{\sim}{\mu}_x \\ \mu_z \end{pmatrix} \quad D\begin{pmatrix} Y \\ \underset{\sim}{X} \\ Z \end{pmatrix} = \left(\begin{array}{c|c|c} \sigma_{YY} & \underset{\sim}{\sigma}_{Yx}^{T} & \sigma_{YZ} \\ \hline \underset{\sim}{\sigma}_{Yx} & \Sigma_{x} & \underset{\sim}{\sigma}_{Zx} \\ \hline \sigma_{YZ} & \underset{\sim}{\sigma}_{Zx}^{T} & \sigma_{ZZ} \end{array}\right)$$

This representation will be useful for time series data. and we will define Partial autocorrelation coefficient for time series.

PACF

Partial auto correlation coefficient (PACF) for a (weakly) Stationary timeseries with covariance function $\gamma_X(h)$ for lag h can be defined as in the following way.

Consider $(X_0 \,\vdots\, X_1\, X_2 \cdots X_h \,\vdots\, X_{h+1})$ then the correlation between

$(X_0 - \hat{X}_0(X_1, X_2 \cdots, X_h))$ and ~~X~~ h A V I X_{h+1}

$(X_{h+1} - \hat{X}_{h+1}(X_1 X_2 \cdots X_h))$ is defined as PACF of lag$(h+1)$ where $\hat{X}_0$ and $\hat{X}_{n+1}$ are best linear predictors.

PACF of order $(h+1)$ is denoted as $\alpha(h+1)$.

Let $E(X_t) = \mu$ which can be zero also.

$$
D\begin{pmatrix} X_0 \\ X_1 \\ \vdots \\ X_h \\ X_{h+1} \end{pmatrix}
=
\left(\begin{array}{c|ccc|c}
\gamma(0) & \gamma(1) & \cdots & \gamma(h) & \gamma(h+1) \\ \hline
\gamma(1) & \gamma(0) & \gamma(1) \cdots & \gamma(h-1) & \gamma(h) \\
\vdots & & \ddots & & \gamma(h-1) \\
\vdots & & \gamma(0) & & \vdots \\
\gamma(h) & \gamma(h-1) & \cdots & \gamma(1)\ \gamma(0) & \gamma(1) \\ \hline
\gamma(h+1) & \gamma(h) & \cdots & \gamma(1) & \gamma(0)
\end{array}\right)
=
\left(\begin{array}{c|c|c}
\gamma(0) & \underset{\sim}{\gamma}_h^T(1) & \gamma(h+1) \\ \hline
\underset{\sim}{\gamma}_h(1) & \Gamma_h & \underset{\sim}{\tilde{\gamma}}_h(1) \\ \hline
\gamma(h+1) & \underset{\sim}{\tilde{\gamma}}_h(1) & \gamma(0)
\end{array}\right)
$$

$$\underset{\sim}{\gamma}_h(1) = \left(\gamma(1)\ \gamma(2) \cdots \gamma(h)\right)$$

$$\underset{\sim}{\tilde{\gamma}}_h(1) = \left(\gamma(h)\ \gamma(h-1) \cdots \gamma(1)\right)$$

$$\text{PACF} = \alpha(h+1) = \frac{\gamma(h+1) - \underset{\sim}{\gamma}_h^T(1)\,\Gamma_h^{-1}\,\underset{\sim}{\tilde{\gamma}}_h(1)}{\gamma(0) - \underset{\sim}{\tilde{\gamma}}_h(1)\,\Gamma_h^{-1}\,\underset{\sim}{\tilde{\gamma}}_h(1)}$$

$$= \frac{\gamma(h+1) - \underset{\sim}{\gamma}_h^T(1)\,\Gamma_h^{-1}\,\underset{\sim}{\tilde{\gamma}}_h(1)}{\gamma(0) - \underset{\sim}{\gamma}_h^T(1)\,\Gamma_h^{-1}\,\underset{\sim}{\gamma}_h(1)} \in (-1, 1)$$

as it is correlation.

Claim.

$$\underset{\sim}{\tilde{\gamma}}_n^T(1)\,\Gamma_n^{-1}\,\underset{\sim}{\tilde{\gamma}}_n(1) = \underset{\sim}{\gamma}_n(1)\,\Gamma_n^{-1}\,\underset{\sim}{\gamma}_n(1)$$

$$\underset{\sim}{\tilde{\gamma}}_n(1) = R\,\underset{\sim}{\gamma}_n(1)$$

$$\begin{pmatrix} c \\ b \\ a \end{pmatrix} = \underbrace{\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}}_{R_{3\times 3}} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

$$\Rightarrow \underset{\sim}{\tilde{\gamma}}_n^T(1)\,\Gamma_n^{-1}\,\underset{\sim}{\tilde{\gamma}}_n(1)$$

$$= \underset{\sim}{\gamma}_n^T(1)\left(R^T\,\Gamma_n^{-1}\,R\right)\underset{\sim}{\gamma}_n(1)$$

$$= \underset{\sim}{\gamma}_n^T(1)\,\Gamma_n^{-1}\,\underset{\sim}{\gamma}_n(1)$$

$$\Gamma_n = ((\gamma(|i-j|)))$$

When prediction is constructed through Durbin Levinson algorithm then coefficients will turn up to be in the form of PACF and their magnitude will bounded by 1. Recurration will be possible only because of this result.

$$a_1\ a_2 \cdots \,|\, \cdots a_{n-1}\ a_n$$

$$[n/2]$$

(1) ACF can distinguish between (AR/ARMA) and MA.

(2) PACF can distinguish between AR and (ARMA/MA).

Estimation of moments of a weakly stationary time series.

Let $\{X_t\}$ be a weakly stationary time series. with $E(X_t) = \mu$.
and $\mathrm{cov}(X_t, X_{t+h}) = \gamma_x(h)$. we are interested to estimat.
$\mu, \gamma_x(h), \rho_x(h)$

(*) We can estimate μ $\hat{=} \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{X}$

$\bar{X}$ is an unbiased estimator of μ. $E(\bar{X}) = \frac{1}{n} n\mu = \mu$.

(*) Is $\bar{X}$ a consistent estimator of μ? $\lim_{n\to\infty} P\left(|\bar{X} - \mu| > \epsilon\right) \downarrow 0$

(i) Moment condition $\to$??
(ii) Chebyshev's inequility.

$$\mathrm{Var}(\bar{X}) = E\left[(\bar{X} - \mu)^2\right] = E\left[\frac{1}{n}\sum(X_i - \mu)\right]^2 = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{cov}(X_i, X_j).$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_x(|i-j|) = \frac{1}{n^2}\mathrm{Var}(\underset{\sim}{1}^T \underset{\sim}{X}) = \frac{1}{n^2}\underset{\sim}{1}^T \Gamma_n \underset{\sim}{1}$$

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

$$= \frac{1}{n}\sum_{h=-n}^{n}\left(1 - \frac{|h|}{n}\right)\gamma_x(h)$$

$\gamma(0) \longrightarrow n$
$\gamma(1) \longrightarrow 2(n-1)$
$\gamma(2) \longrightarrow 2(n-2)$
$\vdots$

$$V(\bar{X}) = \frac{1}{n}\sum_{h=-n}^{n}\left(1-\frac{|h|}{n}\right)\gamma(h).$$

$$< \frac{1}{n}\sum_{h=-n}^{n}\left(1-\frac{|h|}{n}\right)|\gamma(h)|.$$

$$< \frac{1}{n}\sum_{h=-n}^{n}|\gamma(h)|. \quad < \infty \quad \underline{\underline{\text{assume}}}.$$

If $\lim_{n\uparrow\infty} \sum_{h=-n}^{n} |\gamma(h)| < \infty$ then. $V(\bar{X}) = V(\bar{X}-\mu) \downarrow 0$ as $n\uparrow\infty$.

$$\lim_{n\uparrow\infty} P\left(|\bar{X}_n - \mu| < \epsilon\right) = 1.$$

$\underline{\sigma_n^2 = \text{Var.}\left(\sqrt{n}\,\bar{X}\right)} = \sum_{h=-\infty}^{\infty} \gamma(h)$ is known as long run variance of the timeseries.

provided. the series is absolutely summable.

$$\boxed{\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i}$$

If $\{X_t\}$ is a weakly stationary time series

$$\sqrt{n}\left(\bar{X}-\mu\right) \sim N\left(0, \sum_{|h|<n}\left(1-\frac{|h|}{n}\right)\gamma(h)\right).$$

for large n.

As a consequence 95% CI of μ can be obtained as:

$$\left(\bar{X} - 1.96\,\frac{\hat{\sigma}_n}{\sqrt{n}},\ \bar{X} + 1.96\,\frac{\hat{\sigma}_n}{\sqrt{n}}\right).$$

$$\sigma_n^2 = \sum_{|h|<n}\gamma(h) \overset{\text{approx}}{\approx} \sum_{|h|<[\sqrt{n}]}\underline{\underline{\hat{\gamma}(h)}}.$$

If h is large then $\hat{\gamma}(h)$ will have less number of observations to estimate.

$$x_8 \ \cdots$$

$$h=2 \quad \begin{pmatrix}x_0\\x_2\end{pmatrix}\begin{pmatrix}x_1\\x_3\end{pmatrix}\begin{pmatrix}x_2\\x_3\end{pmatrix}\begin{pmatrix}x_3\\x_5\end{pmatrix}\begin{pmatrix}x_4\\x_6\end{pmatrix} \begin{matrix}x_5\\x_7\end{matrix} \quad \begin{matrix}x_6\\x_8\end{matrix} \quad x_7 \ \cdots$$

If the lag is increasing then the number of paired observations will decreasing. Hence $\cancel{\hat{\gamma}(h)}\ \hat{\sigma}_n^2$ is not be a stable estimator if we take h up to $\underline{<n}$.

Estimation of autocovariance Matrix.

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{i=1}^{n-|h|} (X_{t+|h|} - \bar{X})(X_t - \bar{X}).$$ estimated autocovariance.

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}.$$ estimated auto correlation.

We devide the ~~p/m~~ product term by n instead of $(n-|h|)$.

If we devide it by $(n-|h|)$ then $\hat{\Gamma}$ may not be pd or psd.

We can estimate Γ. i.e. autocovariance matrix. as follows.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \quad \text{and} \quad \tilde{X}_t = X_t - \bar{X}$$

$$M = \begin{pmatrix} 0 & 0 & \cdots & 0 & \tilde{x}_1 & \tilde{x}_2 & \cdots & \tilde{x}_n \\ 0 & 0 & \cdots & \tilde{x}_1 & \tilde{x}_2 & \cdots & \tilde{x}_n & 0 \\ 0 & 0 & \cdots & \tilde{x}_1 \; \tilde{x}_2 & \tilde{x}_3 & \cdots \tilde{x}_n & 0 & 0 \\ 0 & 0 & \cdots & & & & & \\ \vdots & & & & & & & \\ & & & & & & \cdots & 0 \\ \tilde{x}_1 & \tilde{x}_2 & \tilde{x} & \cdots & \tilde{x}_n & 0 & \cdots & 0 \end{pmatrix}$$

$n \times (2n-1)$.

$$\hat{\Gamma} = \frac{1}{n} M M^T.$$

H.W. $\hat{\Gamma}$ is a psd. matrix.?

$$\begin{cases} \underset{\sim}{a} \neq \underset{\sim}{0} \\ a^T \hat{\Gamma} a = \frac{1}{n} a^T M M^T a \geq 0 \end{cases}$$

Let $\hat{\underset{\sim}{\rho}} = (\hat{\rho}(1), \hat{\rho}(2), \cdots, \hat{\rho}(h))$

Bartlett's formula.

Consider a linear process with $\sum_{i=-\infty}^{\infty} |\psi_i| < \infty$ and also

$$\sum_{h=0}^{\infty} h \, |\psi_h|^2 < \infty \quad \text{then.}$$

We can have a large sample test to check whether h th order autocorrelation is significant or not.

$$\sqrt{n}\,(\hat{\underset{\sim}{\rho}} - \underset{\sim}{\rho}) \sim N_h(\underset{\sim}{0}, W).$$

$$W = ((w_{ij})) = \sum_{k=1}^{\infty} \{\rho(k+i) + \rho(k-i) - 2\rho(i)\rho(k)\} \otimes \{\rho(k+j) + \rho(k-j) - 2\rho(j)\rho(k)\}.$$

$$= \operatorname{Cov}\left(\sqrt{n}\,(\hat{\rho}(i) - \rho(i)), \sqrt{n}\,(\hat{\rho}(j) - \rho(j))\right).$$

Ex: WN process. $w_{ij} = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{o/w.} \end{cases}$

Ex $x_t \sim MA(1)$. $w_{ii} = \begin{cases} 1 - 3\rho^2(1) + 4\rho^4(1). & \text{if } i=1 \\ 1 + 2\rho^2(1) & \text{if } i>1 \end{cases}$

$$w_{12} = 2\rho(1)(1 - \rho^2(1)).$$

Causality : A linear process $\{X_t\}$ in causal as a function of WN $\{W_t\}$ if we have the representation

$$X_t = \left(I + \sum_{i=1}^{\infty} \varphi_i B^i\right) W_t$$

with $\sum_{i=1}^{\infty} |\varphi_i| < \infty$.

Ex. $X_t = \varphi X_{t-1} + W_t$ i.e. AR(1) process.

if $|\varphi| < 1$ then X_t is causal. $\Phi_1(z) = 1 - z\varphi$.

if $|\varphi| < 1$ then the solution of $\Phi_1(z) = 0$ lies out of the unit circle.

for ARMA (1, 1). $X_t = W_t + (\theta + \varphi) \sum_{j=1}^{\infty} \varphi^{j-1} W_{t-j}$ $\left|\frac{1}{\varphi}\right| > 0$. $\varphi \neq 0$.

Similarly for AR(p) process Φ if the roots of $\Phi_p(z) = 0$ lies out of the unit circle. $|z| \leq 1$ then it is also causal.

If $\varphi > 1$ for an AR(1) process.

$$(1-\varphi B) X_t = W_t .$$

$$\Rightarrow \quad X_t = \frac{1}{1-\varphi B} W_t .$$

$$= -\sum_{j=1}^{\infty} \varphi^{-j} B^{-j} W_t .$$

$$= -\sum_{j=1}^{\infty} \varphi^{-j} W_{t+j}$$ → belongs to future.

Hence not Causal.

$$\frac{1}{1-z\varphi} = -\sum_{j=1}^{\infty} (\varphi z)^{-j}$$

$$\begin{cases} |\varphi| > 1 \\ |z| \geq 1. \end{cases}$$

Invertability. : A linear process is said to invertable. if

$$W_t = \left(1 + \sum_{j=1}^{\infty} \theta_j B^j\right) X_t . \quad \text{where } \sum_{j=1}^{\infty} |\theta_j| < \infty .$$

WN can be expressed in terms of past time series data.
Innovation method of prediction uses this idea.

Ex.1 MA(1) process is invertable. if $|\theta| < 1$.