TS

Review

— Introduction

~~Weak stationarity~~:

$Z(w,t)$, $w \in S$, $t \in I$. Assume $I = \mathbb{Z}$.

$Z(w_0,t)$ = sample function or realization

$Z(w, t_0)$ = r.v.

Simply write $Z(w_0,t) = Z_t$. The population that consists of all possible realizations = time series.

Mean function of the process : $\mu_t = E(Z_t)$.

Var " : $\sigma_t^2 = E[(Z_t - \mu_t)^2]$.

(Auto)Cov " : $\gamma(t_1,t_2) = E[(Z_{t_1} - \mu_{t_1})(Z_{t_2} - \mu_{t_2})]$

(Auto)Cor " : $\rho(t_1,t_2) = \dfrac{\gamma(t_1,t_2)}{\sigma_{t_1}\sigma_{t_2}}$

Let $T = I$, $t_{1,2} \in T$.

So $\gamma_Z(t_1,t_2) = Cov(Z_{t_1}, Z_{t_2})$

$= E[(Z_{t_1} - \mu_{t_1})(Z_{t_2} - \mu_{t_2})]$,

$\rho_Z(t_1,t_2) = \dfrac{\gamma_Z(t_1,t_2)}{\sqrt{\gamma_Z(t_1,t_1)\gamma_Z(t_2,t_2)}}$.

Cross-~~covariance~~ function: $\gamma_{XY}(t_1,t_2) = E[(X_{t_1} - \mu_{X_{t_1}})(Y_{t_2} - \mu_{Y_{t_2}})]$

Cross-corr : $\rho_{XY}(t_1,t_2) = \dfrac{\gamma_{XY}(t_1,t_2)}{\sqrt{\gamma_{XY}(t_1,t_1)\gamma_{XY}(t_2,t_2)}}$ (X).

$\rho_{XY}(t_1,t_2) = \dfrac{\gamma_{XY}(t_1,t_2)}{\sqrt{\gamma_X(t_1,t_1)\gamma_Y(t_2,t_2)}}$.

* Weak stationarity ; finite 2nd moment, const. μ and Cov. thr time.

1. $E(|X_t|^2) < \infty$ $\forall t \in T$.
2. $E(X_t) = m$ $\forall t \in T$
3. $\gamma_X(t_1,t_2) = \gamma_X(t_1+t, t_2+t)$ $\forall t, t_1, t_2 \in T$.

① $E|X_t|^2 < \infty \quad \forall t \in T$

② $E(X_t) = m \quad \forall t \in T$

③ $\gamma_x(t_1, t_2) = \gamma_x(t_1+k, t_2+k) \quad \forall k, t_1, t_2 \in T.$

~~Let~~ $\cancel{k = t_2 - t_1}$ ~~Then:~~

$\cancel{\gamma_x(k) = \gamma_x(t_1 + t_2 - t_1, t_2 + t_2 - t_1)}$

$\cancel{= \gamma_x(t_2, 2t_2 - t_1)}$

So ③ $\Leftrightarrow \gamma_x(t_1, t_2) = \gamma_x(t_1+k, t_2+k)$

$= Cov(X_{t_1}, X_{t_2}) = Cov(X_{t_1+k}, X_{t_2+k})$

Let $t_1 = t-k$, $t_2 = t$.

Then $\gamma_x(t_1, t_2) = \gamma_x(t-k, t)$

$= \gamma_x(t-k+k, t+k) \quad (\because ③)$

$= \gamma_x(t, t+k)$

$=: \gamma_x(k) =: \gamma_k.$

So Weak Stationarity $\Rightarrow$ ③ $\gamma_x(k)$ is t_1 & t_2-invariant.

<u>Def'n</u> $\gamma_k = \gamma_x(k)$, $\gamma_0 = Var(X_t)$,

$\rho_k = \frac{\gamma_k}{\gamma_0}$.

<u>Note</u> $\gamma_x(k) = Cov(X_{t+k}, X_t)$

$= \gamma_x(t+k, t) \quad \forall t, k \in T,$

i.e. $\gamma_k = \gamma_x(1+k, 1) = \gamma_x(2+k, 2)$

$= \gamma_x(3+k, 3) = \cdots$.

* We call γ_k a ACVF, ρ_k a ACF, k a lag.

<u>Properties</u>

① $\gamma_0 = Var(X_t)$, $\rho_0 = 1$.

② $|\gamma_k| \leq \gamma_0$; $|\rho_k| \leq 1$.

<u>Why</u> $|\rho_k| \leq 1 \quad (\because$ corr$)$

$\Rightarrow \left|\frac{\gamma_k}{\gamma_0}\right| \leq 1 \Rightarrow |\gamma_k| \leq \gamma_0. \checkmark$.

③ γ_k and ρ_k are even functions, i.e. $\gamma_k = \gamma_{-k}$; $\rho_k = \rho_{-k}$.

④ For any set of time points $t_1, \cdots, t_n$ and any real #s $\alpha_1, \cdots, \alpha_n$, $\begin{cases} \sum_i \sum_j \alpha_i \alpha_j \gamma_{|t_i - t_j|} \geq 0 \\ \sum_i \sum_j \alpha_i \alpha_j \rho_{|t_i - t_j|} \geq 0. \end{cases}$

Why ④ Let ~~$Y=\sum\sum$~~ $Y = \sum_{i=1}^{n} \alpha_i Z_{t_i}$.

Then $Var(Y) = \sum\sum \alpha_i \alpha_j \underbrace{Cov(Z_{t_i}, Z_{t_j})}_{:= \gamma_{|t_i - t_j|}}$.

Divide the inequality by γ_0. ✓

* PACF

$$PACF = Corr(Z_t, Z_{t+k} \mid Z_{t+1}, \cdots, Z_{t+k-1}).$$

$$= \rho_{z-\hat{z}}(t, t+k)$$

$$\left.\begin{aligned} \hat{Z}_{t+k} &= \textstyle\sum_{i=1}^{k-1} \alpha_i Z_{t+k-i}, \\ \hat{Z}_t &= \textstyle\sum_{i=1}^{k-1} \beta_i Z_{t+i}. \end{aligned}\right\} \alpha_i = \beta_i,\ i = 1, \cdots, k-1.$$

"Investigate the correlation btwn Z_t and Z_{t+k} after their mutual linear dependency on the intervening variables $Z_{t+1}, \cdots, Z_{t+k-1}$ has been removed."

Def'n A PACF btwn Z_t and Z_{t+k} is

$$P_k = \rho_{z-\hat{z}}(t, t+k) = \frac{Cov(Z_t - \hat{Z}_t, Z_{t+k} - \hat{Z}_{t+k})}{\sqrt{Var(Z_t - \hat{Z}_t)\, Var(Z_{t+k} - \hat{Z}_{t+k})}}$$

$$= \frac{\gamma_{z-\hat{z}}(t, t+k)}{\sqrt{\gamma_{z-\hat{z}}(t,t)\, \gamma_{z-\hat{z}}(t+k, t+k)}}$$

α_i's are the mean squared linear reg coeffs obtained from minimizing $\gamma_{z-\hat{z}}(t+k, t+k)$

$= E[(Z_{t+k} - \sum_{i=1}^{k-1} \alpha_i Z_{t+k-i})^2]$

$\rightsquigarrow$ (differentiation; routine...)

$\to \gamma_i = \alpha_1 \gamma_{i-1} + \cdots + \alpha_{k-1} \gamma_{i-k+1}$

$\Rightarrow \rho_i = \alpha_1 \rho_{i-1} + \cdots + \alpha_{k-1} \rho_{i-k+1} = \sum_{j=1}^{k-1} \alpha_j \rho_{i-j}$.

In matrix form, $(\rho_0 = 1)$

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_{k-1} \end{bmatrix}_{(k-1)\times 1} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} \\ & & \vdots & & \\ \rho_{k-2} & \rho_{k-3} & \rho_{k-4} & \cdots & \rho_0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{k-1} \end{bmatrix}$$

$\{Z_t\}$ a zero-mean (weakly) stationary process.

OR, to calculate PACF at lag k, consider:

$Z_{t+k} = \phi_{k1} Z_{t+k-1} + \phi_{k2} Z_{t+k-2} + \cdots + \phi_{kk} Z_t + e_{t+k}$,

where ϕ_{ki} denotes the i^{th} regression param.,

e_{t+k} is an error term w/ mean 0 and uncorrelated w/ Z_{t+k-j} for $j = 1, \cdots, k$.

If not zero-mean $(\mu \neq 0)$, use $\dot{Z}_{t+k} = Z_{t+k} - \mu$.

ϕ_{kk} is our PACF P_k.

From $Z_{t+k} = \sum_{i=1}^{k} \phi_{ki} Z_{t+k-i} + e_{t+k}$,

multiply both by Z_{t+k-j} and take $E(\cdot)$:

$\gamma_j = \sum_{i=1}^{k} \phi_{ki} \gamma_{j-i} + 0$.

Divide both by γ_0:

$\rho_j = \sum_{i=1}^{k} \phi_{ki} \rho_{j-i} \quad , j = 1, \cdots, k$

$$\Rightarrow \begin{bmatrix} \rho_1 \\ \vdots \\ \rho_k \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-2} \\ & & \vdots & & \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & 1 \end{bmatrix}}_{A} \begin{bmatrix} \phi_{k1} \\ \vdots \\ \phi_{kk} \end{bmatrix}$$

By Cramer's rule, $\phi_{kk} = \dfrac{\det A_{-k}}{\det A}$,

where A_{-k} is ~~the~~ matrix w/ k^{th} column replaced w/ $(\rho_1, \cdots, \rho_k)^T$.

~~Some defn~~

$\phi_{00} := \rho_0 = 1, \quad \phi_{11} = \rho_1,$

$$\phi_{22} = \frac{\det\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\det\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}, \quad \phi_{33} = \frac{\det\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\det\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & \rho_0 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}, \text{ etc.}$$

* White Noise

Def'n A white noise process $\{a_t\}$ is a sequence of uncorrelated r.v.s from a fixed dist. w/ const. mean $E(a_t) = \mu_a$ ~~usually~~ (usually 0), and variance $Var(a_t) = \sigma_a^2$, and $\gamma_k = Cov(a_t, a_{t+k}) = 0 \ \forall k \neq 0$.

Unless mentioned otherwise, $a_t \overset{iid}{\sim} N(0, \sigma_a^2)$.

$\{a_t\}$ is weakly stationary because ① finite 2^{nd} moment, ② const. mean, and
③ $\gamma_k = Cov(a_t, a_{t+k}) = \begin{cases} 0 & \text{if } k \neq 0, \\ 1 & \text{if } k = 0 \end{cases}$

i.e. γ_k is a function of k only and t-invariant.

· Statistics

Assume weakly stationary process.

→ Mean: $\bar{Z} = \frac{1}{n}\sum_{t=1}^{n} Z_t = \hat{\mu}_a$, i.e. time avg of n obs.

Properties

① Unbiased: $E(\bar{Z}) = \mu_a$.

② Consistent:

$$Var(\bar{Z}) = \frac{1}{n^2}\sum_{t=1}^{n}\sum_{s=1}^{n} Cov(Z_t, Z_s) = \frac{\gamma_0}{n^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\rho_{|t-s|}$$

$$= \frac{\gamma_0}{n^2}\sum_{k=-(n-1)}^{n-1}(n-|k|)\rho_k \quad (k = |t-s|)$$

e.g. $n=3 \to \begin{matrix} 1+\rho_1+\rho_2 \\ +\rho_1+1+\rho_1 \\ +\rho_2+\rho_1+1 \end{matrix} \to \begin{matrix} 1:3 \\ \rho_1:4 \\ \rho_2:2 \end{matrix}$

$n=4 \to \begin{matrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 3 & 2 & 1 & 0 \end{matrix} \to \begin{matrix} 1:4 \\ \rho_1:6 \\ \rho_2:4 \\ \rho_3:2 \end{matrix}$

So $\lim_{n\to\infty} Var(\bar{Z}) = 0$, i.e. $Var(\bar{z}) \to 0$.

Sufficient cond: $\rho_k \xrightarrow{k\to\infty} 0$.

Again: $Var(\bar{Z})$

$$= \frac{1}{n^2}\sum_t\sum_s Cov(Z_t, Z_s)$$

$$= \frac{\gamma_0}{n^2}\sum_{t=1}^{n}\sum_{s=1}^{n} \rho_{|t-s|}$$

$$= \frac{\gamma_0}{n^2}\sum_{t=1}^{n}\left[\rho_{|t-1|}+\rho_{|t-2|}+\cdots+\rho_{|t-n|}\right]$$

$\begin{matrix} 1+\rho_1+\rho_2+\cdots+\rho_{n-1} \\ +\rho_1+1+\rho_1+\cdots+\rho_{n-2} \\ +\cdots \\ +\rho_{n-1}+\rho_{n-2}+\rho_{n-3}+\cdots+1 \end{matrix}$

$$= \frac{\gamma_0}{n^2}\Big[1+\rho_1+\rho_2+\cdots+\rho_{n-1} + \rho_1+1+\rho_1+\cdots+\rho_{n-2} + \cdots + \rho_{n-1}+\rho_{n-2}+\cdots+1\Big]$$

$\begin{bmatrix} 0 & 1 & 2 & \cdots & n-1 \\ 1 & 0 & 1 & \cdots & n-2 \\ & & \vdots & & \\ n-1 & n-2 & n-3 & \cdots & 0 \end{bmatrix}$

$$= \frac{\gamma_0}{n^2}\Big[n + 2(n-1)\rho_1 + 2(n-2)\rho_2 + 2(n-3)\rho_3 + \cdots + 2(n-n+1)\rho_{n-1}\Big]$$

$$= \frac{\gamma_0}{n^2}\left[\sum_{k=-(n-1)}^{n-1}(n-|k|)\rho_k\right].$$

→ ACVF

$$\hat{\gamma}_k = \frac{1}{n}\sum_{t=1}^{n-k}(Z_t-\bar{Z})(Z_{t+k}-\bar{Z}) \Rightarrow \hat{\gamma}_0 = \frac{1}{n}\sum_{t=1}^{n}(Z_t-\bar{Z})^2$$

Properties:

- Biased + asymptotically biased

$$E(\hat{\gamma}_k) \approx \gamma_k - \frac{k}{n}\gamma_k - \left(\frac{n-k}{n}\right)Var(\bar{Z}).$$

→ ACF

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=1}^{n-k}(Z_t-\bar{Z})(Z_{t+k}-\bar{Z})}{\sum_{t=1}^{n}(Z_t-\bar{Z})^2}, \quad k=0,\ldots,n-1.$$

<u>Properties</u>

- $\hat{\rho}_{-k} = \dfrac{\sum_{t=k+1}^{n}(Z_t-\bar{Z})(Z_{t+k}-\bar{Z})}{\sum_{t=1}^{n}(Z_t-\bar{Z})^2} = \dfrac{\sum_{t=1}^{n-k}(Z_t-\bar{Z})(Z_{t+k}-\bar{Z})}{\sum(Z_t-\bar{Z})} = \hat{\rho}_k.$

That is, $\hat{\rho}_k$ is **even wrt k**, i.e. symmetric around k.

- $\hat{\rho}_k \sim N(0, \frac{1}{N})$, $k \neq 0$, where N = length of $\{Z_t\}$.

We usually take max$|k| << N$.

We expect 95% of $\hat{\rho}_k$ will lie w/i $\pm 2/\sqrt{N}$,

$2 \approx$ qnorm(.975).

→ PACF

$\hat{\phi}_{kk} = \hat{\rho}_1$ if $k=1$.

$$\hat{\phi}_{kk} = \frac{\det \hat{A}_k}{\det \hat{A}}, \quad \hat{A} = \begin{bmatrix} 1 & \hat{\rho}_1 & \hat{\rho}_2 & \cdots & \hat{\rho}_{k-1} \\ \hat{\rho}_1 & 1 & \hat{\rho}_1 & \cdots & \hat{\rho}_{k-2} \\ & & \vdots & & \\ \hat{\rho}_{k-1} & \hat{\rho}_{k-2} & \hat{\rho}_{k-3} & \cdots & 1 \end{bmatrix}_{k \times k}$$

Quenouille: H_0: underlying process is a white noise.

$\Rightarrow Var(\hat{\phi}_{kk}) \approx \frac{1}{n}$ ∴ Limits: $\pm 2/\sqrt{n}$.

<u>Recall</u> $\{a_t\}$ is a white noise if $Cov(a_t, a_{t+k}) = 0$ $\forall k \neq 0$, $E(a_t) = \mu_a$, and $Var(a_t) = \sigma_a^2$ $\forall t \in T$.

— Correlation in simple reg.

$y_t = \alpha + \beta x_t + e_t$

$\gamma_{XY}(t,t) = Cov(X_t, Y_t)$

$= Cov(x_t, \alpha + \beta x_t + e_t)$

$= 0 + \beta\sigma_{x_t}^2 + 0 = \beta\sigma_{x_t}^2 = \beta\sigma_x^2.$

Also, $\rho_{XY}(t,t) = \dfrac{\gamma_{XY}(t,t)}{\sqrt{\gamma_X(t,t)\gamma_Y(t,t)}} = \dfrac{\beta\sigma_x^{\cancel{2}}}{\cancel{\sigma_x}\sigma_y}.$

~~$Cov(X_t, Y_t) = \rho_{XY}(0)\sigma_X\sigma_Y$~~

$\Rightarrow \beta = \rho_{XY}(t,t)\cdot\frac{\sigma_Y}{\sigma_X} = \rho_{XY}(0,0)\cdot\frac{\sigma_Y}{\sigma_X} \propto \rho_{XY}(0,0).$ ⇐ "$\rho_{XY}(0)$" by instructor.

* If $\{X_t\}$ and $\{Y_t\}$ are stationary, indep. of each other, then $\sqrt{n}\, r_{xy}(k)$ is asymptotically normal w/ $\mu = 0$ and $\sigma^2 = 1 + 2\sum_{j=1}^{\infty} \rho_x(j)\rho_y(j)$.

(+ no serial corr.) $r_{xy}(k) \sim N(0, \frac{1}{n})$ ← size of sample.

— Approaches to Time Series Analysis

- Time Domain Analysis
  - → Use ACF and PACF to study the evolution of a time series through parametric models
  - → Focus on modelling some future values of a time series as a parametric function of the current and past values
- Frequency domain analysis
  - → Use spectral functions to study nonparametric decomposition of a time series into its different frequency components
  - → Assume that the primary characteristics of interest in time series analysis relate to periodic or systematic sinusoidal variations found naturally in most data.

(TDA > FDA over short samples)

— Decomposition of TS

- trend + seasonal variation + cyclical variation + irregular fluctuations.

= long-term change in the mean level
+ variation exhibited in every periods
+ variation w/i a fixed period due to some other physical cause
+ noise.

— Steps to time series modelling.

① Plot the time series and check for trend, seasonal and other cyclic components, any apparent sharp changes in behaviour, as well as any outlying observations.

② Remove trend, and seasonal components to get residuals.

③ Choose a model to fit the residuals.

④ Forcast residuals and then invert the transformation carried out in ②.

— Stationary time series models

* AR(p) process : $\phi_p(B)\dot{z}_t = a_t$, $\dot{z}_t = z_t - \mu$.

$\phi_p(B) = 1 - \sum_{j=1}^{p} \phi_j B^j$; $\phi_k \neq 0$ if $k \leq p$, $\phi_k = 0$ if $k > p$.

$\Rightarrow \phi_p(B)\dot{z}_t = a_t$

$= \dot{z}_t - \phi_1 \dot{z}_{t-1} - \phi_2 \dot{z}_{t-2} - \cdots - \phi_p \dot{z}_{t-p} = a_t$.

$\Rightarrow \dot{z}_t = \phi_1 \dot{z}_{t-1} + \phi_2 \dot{z}_{t-2} + \cdots + \phi_p \dot{z}_{t-p} + a_t$.

Note ① If $\sum_{j=1}^{p} |\phi_j| < \infty$, then the process is invertible. (that is, AR(p) → MA(∞) possible).

② Roots of $\phi_p(x) = 0$ lie outside of the unit circle ⟺ stationary.

e.g. AR(1) of $\mu = 0$ process.

$\phi_1(B) z_t = a_t$

$= z_t - \phi_1 z_{t-1} = a_t$

$\Rightarrow z_t = \phi_1 z_{t-1} + a_t$, $a_t \overset{iid}{\sim} N(0, \sigma_a^2)$.

Q. What is ρ_k?

A. $E(z_t) = 0$. Suppose $k \in T$.

$z_{t-k} z_t = \phi_1 z_{t-1} z_{t-k} + a_t z_{t-k}$

$\forall t \in T = \mathbb{Z}$.

Taking $E(\cdot)$ gives $\gamma_k = \phi_1 \gamma_{k-1}$.

So $\rho_k = \phi_1 \rho_{k-1}$ by dividing both by γ_0.

$k=1 \Rightarrow \rho_1 = \phi_1 \rho_0 = \phi_1$

$k=2 \Rightarrow \rho_2 = \phi_1 \rho_1 = \phi_1 \phi_1 = \phi_1^k$.

Inductively, $\therefore \rho_k = \phi_1^k$ //

Q. But why is $E(a_t Z_{t-k}) = 0$? ($k \in \mathbb{Z}$).

A. $Z_{t-k} = \phi_1 Z_{t-(k+1)} + a_{t-k}$,

$Z_{t-(k+1)} = \phi_1 Z_{t-(k+2)} + a_{t-(k+1)}$

$\Downarrow$

$$\begin{aligned} Z_{t-k} &= \phi_1 [\phi_1 Z_{t-(k+2)} + a_{t-(k+1)}] + a_{t-k} \\ &= \phi_1^2 Z_{t-(k+2)} + \phi_1 a_{t-(k+1)} + a_{t-k} \\ &= \cdots \text{ (induction)} \\ &= \phi_1^\infty Z_{t-(k+\infty)} + \sum_{i=0}^{\infty} \phi_1^i a_{t-(k+i)} \\ &= \sum_{i=0}^{\infty} \phi_1^i a_{t-(k+i)} \end{aligned}$$

$(\because \phi_1^\infty = 0)$

Or, equivalently, $\phi_p(B) Z_t = a_t \Rightarrow Z_t = \Theta_\infty(B) a_t$.

— 3 stages of Box-Jenkins Approach

$\boxed{\text{Start}}$
- TS realization
- Understand a problem
- Collect + plot data

$\downarrow$

① Identify a prelim time series model
→ perform differencing + transformations to transform data into stationary
→ Identify prelim ARMA(p,q) models using $\widehat{ACF}$ and $\widehat{PACF}$.

② Estimate the model parameters
→ MoM, MLE, Kalman Filter, etc.

③ Diagnose model adequacy
→ Examine if the res of the fitted model are approx. uncorrelated → Stop

- if the fitted model fails diagnostic test
- identify another model.

- if passes
- use model for and

* ARMA(p,q) : $\phi_p(B) Z_t = \Theta_q(B) a_t$,
$\phi_p(B) = 1 - \sum_{j=1}^{p} \phi_j B^j$, $\Theta_q(B) = 1 + \sum_{j=1}^{q} \theta_j B^j$.

* MA(∞)
$\{X_t\}$ is a MA(∞) process of $\{a_t\}$ if $\exists (\psi_j)_{j=0}^{\infty}$ w/ $\sum_{j=0}^{\infty} |\psi_j| < \infty$ s.t. $X_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}$, $t \in \mathbb{Z}$.

* As long as $\{X_t\}$ can be written in a form of MA(∞) process, we can calculate ACF of $\{X_t\}$.

THM MA(∞) process $\{X_t\}$ is stationary w/ mean 0 and ACF $\gamma_k = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+|k|}$.

· Special case : MA(q): $X_t = \Theta_q(B) a_t$, $a_t \overset{iid}{\sim} N(0, \sigma_a^2)$.
e.g. MA(2)
Q. Find γ_k's.

So $X_t = \Theta_2(B) a_t$
$= (1 + \theta_1 B + \theta_2 B^2) a_t$
$= a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2}$.

$\gamma_k := Cov(X_t, X_{t+k})$
$= Cov(a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2}, a_{t+k} + \theta_1 a_{t-1+k} + \theta_2 a_{t-2+k})$
$= Cov(a_t, a_{t+k}) + \theta_1 Cov(a_{t-1}, a_{t+k}) + \theta_2 Cov(a_{t-2}, a_{t+k})$
$+ \theta_1 Cov(a_t, a_{t-1+k}) + \theta_1^2 Cov(a_{t-1}, a_{t-1+k}) + \theta_1\theta_2 Cov(a_{t-2}, a_{t-1+k})$
$+ \sim\sim$.

$k=0 \to \gamma_0 = Cov(X_t, X_t) = Cov(a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2}, a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2})$
$= \sigma_a^2 + \theta_1^2 \sigma_a^2 + \theta_2^2 \sigma_a^2$

$k=1 \to \gamma_1 = Cov(X_t, X_{t+1}) = Cov(a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2}, a_{t+1} + \theta_1 a_t + \theta_2 a_{t-1})$
$= \theta_1 \sigma_a^2 + \theta_1 \theta_2 \sigma_a^2$

$k=2 \to \gamma_2 = Cov(X_t, X_{t+2}) = Cov(a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2}, a_{t+2} + \theta_1 a_{t+1} + \theta_2 a_t)$
$= \theta_2 \sigma_a^2$.

$k \geq 3 \to \gamma_k = 0$.

e.g. AR(2). Find γ_k. For that matter, AR(p).

$\Phi_2(B) x_t = a_t$, $a_t \overset{iid}{\sim} N(0, \sigma_a^2)$.

$\Rightarrow x_t = \Theta_\infty(B) a_t$

$\gamma_k = Cov(x_t, x_{t+k})$

$= Cov(\Theta_\infty(B) a_t, \Theta_\infty(B) a_{t+k})$

$\gamma_0 = Cov(\Theta_\infty(B) a_t, \Theta_\infty(B) a_t)$

$= Cov(a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots, a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots)$

$= \sigma_a^2 + \theta_1^2 \sigma_a^2 + \theta_2^2 \sigma_a^2 + \cdots$

$= \sigma_a^2 \sum_{j=0}^{\infty} \theta_j^2$, $\theta_0 = 1$.

---

t
t-1
t-2
⋮
t+k
t+k-1
t+k-2

Redo. $(\theta_0 = 1)$

$\gamma_k = Cov(x_t, x_{t+k}) = Cov(\Theta_\infty(B) a_t, \Theta_\infty(B) a_{t+k})$

$= Cov(\sum_{j=0}^{\infty} \theta_j a_{t-j}, \sum_{j=0}^{\infty} \theta_j a_{t+k-j})$

$= Cov(\sum_{j=0} \theta_j a_{t-j} + \sum_{j=0}^{\infty} \theta_j a_{t+k-j}, \sum_{j=0} \theta_j a_{t+k-j})$

$= Cov(\theta_0 a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots, \theta_0 a_{t+k} + \theta_1 a_{t+k-1} + \cdots)$

$= Cov(\theta_0 a_t + \theta_1 a_{t-1} + \cdots, \theta_0 a_{t-k} + \theta_1 a_{t-(k+1)} + \cdots)$

$= Cov(\theta_0 a_t + \theta_1 a_{t-1} + \cdots + \theta_k a_{t-k} + \theta_{k+1} a_{t-(k+1)} + \cdots, \theta_0 a_{t-k} + \cdots)$

$= Cov(\sum_{j=0}^{k-1} \theta_j a_{t-j} + \sum_{j=0} \theta_{k+j} a_{t-k-j}, \sum_{j=0}^{\infty} \theta_j a_{t-k-j})$

$= \sum_{j=0}^{\infty} \theta_j \theta_{j+k} \sigma_a^2$

---

e.g. AR(2). Find γ_k's.

So $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + a_t$.

$\Rightarrow \gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2}$, $k \geq 1$.

WAY1 $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + a_t$

$$\Rightarrow x_t x_t = \phi_1 x_t x_{t-1} + \phi_2 x_t x_{t-2} + a_t x_t$$

$$\Rightarrow (E(\cdot))\ \gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma_a^2,$$

$$\because E(a_t x_t) = E(a_t(\phi_1 x_{t-1} + \phi_2 x_{t-2} + a_t)) = 0 + 0 + \sigma_a^2.$$

So, $\begin{cases} \gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma_a^2 \\ \gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2}, \quad k \geq 1. \end{cases}$

$k=1 \Rightarrow \gamma_1 = \phi_1 \gamma_0 + \phi_2 \gamma_1 \rightarrow -\phi_1 \gamma_0 + (1-\phi_2)\gamma_1 = 0.$

$k=2 \Rightarrow \gamma_2 = \phi_1 \gamma_1 + \phi_2 \gamma_0$

That is, $\begin{bmatrix} 1 & -\phi_1 & -\phi_2 \\ -\phi_1 & 1-\phi_2 & 0 \\ -\phi_2 & -\phi_1 & 1 \end{bmatrix} \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} \sigma_a^2 \\ 0 \\ 0 \end{bmatrix}$

So $\left[\begin{array}{ccc|c} 1 & -\phi_1 & -\phi_2 & \sigma_a^2 \\ -\phi_1 & 1-\phi_2 & 0 & 0 \\ \phi_2 & \phi_1 & 1 & 0 \end{array}\right]$ will give $\gamma_{0,1,}$ and $_2$.

e.g. $x_t = \overbrace{0.7}^{\phi_1} x_{t-1} \overbrace{-0.12}^{\phi_2} x_{t-2} + a_t, \quad a_t \overset{iid}{\sim} N(0, 0.1).$

$$\Rightarrow \left[\begin{array}{ccc|c} 1 & -.7 & .12 & 0.1 \\ -.7 & 1.12 & 0 & 0 \\ .12 & .7 & 1 & 0 \end{array}\right] \longrightarrow \cdots \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & \sim \\ 0 & 1 & 0 & \sim \\ 0 & 0 & 1 & \sim \end{array}\right]$$

$\therefore \gamma_0 = \frac{250}{1239}, \ \gamma_1 = \frac{625}{4956}, \ \gamma_2 = -\frac{1115}{9912}.$

$$\gamma_3 = \phi_1 \gamma_2 + \phi_2 \gamma_1 = 0.7 \times -\frac{1115}{9912} + (-0.12) \times \frac{625}{4956}.$$

WAY2 Pre's.

WAY3 Markov chain

What about ϕ_{kk}'s?

e.g. $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + a_t$

$\Rightarrow \gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2}$

$\Rightarrow \rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}$

$$\Rightarrow \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix}$$

So in fact, if $x_t \sim AR(p)$, then $\phi_p = \phi_{pp}$!

e.g. $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + a_t$. Find ϕ_{11}.

$\Rightarrow$ ① Restrict i to a desired lag.

$$\begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix} \begin{bmatrix} \phi_{11} \\ \phi_{12} \end{bmatrix} \Rightarrow [\rho_1] = [1][\phi_{11}].$$

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

② $\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}$

$\Rightarrow \rho_0 = \phi_1 \rho_1 + \phi_2 \rho_2 = 1$

$\rho_1 = \phi_1 + \phi_2 \rho_1 \Rightarrow (1 - \phi_2)\rho_1 = \phi_1$

$\Rightarrow \rho_1 = \dfrac{\phi_1}{1-\phi_2} = \phi_{11}$. //

Summary;

$AR(p) \Rightarrow \gamma_k =$ exists $\forall k$, $\phi_{kk} = 0$ if $k > p$.

$MA(q) \Rightarrow \gamma_k = 0$ if $k > q$, ϕ_{kk} exists $\forall k$.

— Causal & Invertible process

stationary $\Leftrightarrow$ · Causal : $\{x_t\}$ can be written in terms of $\{a_s \mid s \leq t\}$.

· Invertible :

No restrictions on $\{\theta_i\}$ are required for a finite-order MA process to be stationary. The imposition of the invertible condition ensures that there's the unique MA process for a given set of ACFs.

$\sum|\phi_j| < \infty \Rightarrow$

Note AR(p) is always invertible ; unit circle.

" is stationary if roots outside of

MA(q) is invertible if ~~$\sum_{j=0}^{\infty}|\theta_j| < \infty$~~ roots outside of unit circle

" is always stationary

$\sum|\theta_j| < \infty \Rightarrow$

If AR(p) is stationary, then AR(p) $\Leftrightarrow$ MA(∞) $\exists$MA.

MA(q) is invertible, then MA(q) $\Leftrightarrow$ AR(∞) $\exists$AR.

* ARMA(p,q) : $\Phi_p(B)x_t = \Theta_q(B)a_t$

AR always invertible

MA always stationary (causal)

ARMA(p,q) is :

~~→ stationary if roots of $\Theta_q(B)$ outside of unit circle~~

~~→ invertible if roots of $\Phi_p(B)$ outside of unit circle.~~

~~→ stationary (causal) if $\exists(\phi_j)_{j=0}^{\infty}$ s.t. $\sum_{j=0}^{\infty}|\phi_j| < \infty$ and~~

$\psi_0 = 1$ → stationary if $\exists(\psi_j)_{j=0}^{\infty}$ s.t. $\sum|\psi_j| < \infty$ and $x_t = \Psi_\infty(B)a_t$ (notice $x_t = \frac{\Theta_q(B)}{\Phi_p(B)}a_t$), $\Psi_\infty(B) = \frac{\Theta(B)}{\Phi(B)}$, $|B| \leq 1$

$\pi_0 = 1$ → invertible if $\exists(\pi_j)_{j=0}^{\infty}$ s.t. $\sum|\pi_j| < \infty$ and $\Pi_\infty(B)x_t = a_t$ ($\frac{\Phi(B)}{\Theta(B)} = \Pi(B)$, $|B| \leq 1$).

AR always invertible — MA always stationary

i.e. ARMA(p, q) is : ($\Phi_p(B) z_t = \Theta_q(B) a_t$)

→ causal (stationary) iff all the roots of $\Phi_p(B)$ lie outside the unit circle ⟺ ①

→ invertible iff all the roots of $\Theta_q(B)$ lie outside the unit circle. ⟺ ②

① $\Psi_\infty(B) = \frac{\Theta(B)}{\Phi(B)}$ has coefficients $\sum|\psi_j| < \infty$ and $x_t = \Psi_\infty(B) a_t \iff \Psi_\infty(x)$ for $|x| \le 1$. ($\frac{\Theta(x)}{\Phi(x)}$; domain)

② $\Pi_\infty(B) = \frac{\Phi(B)}{\Theta(B)}$ has coefficiones $\sum|\pi_j| < \infty$ and $\Pi_\infty(B) x_t = a_t \iff \Pi_\infty(x)$, for $|x| \le 1$. ($\frac{\Phi(x)}{\Theta(x)}$)

— How to find ϕ_{11}, ϕ_{22}, etc.

① Find ρ_k's.

② Use the eqn:

$$\phi_{kk} = \frac{\det A_k}{\det A}$$

(k: fixed, variable; determines the size of A_{kk}.)

$k=0 \to \phi_{00} = \rho_0 = 1$

$k=1 \to \phi_{11} = \rho_1$

$$k=2 \to \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix}}_{A} \begin{bmatrix} \phi_{21} \\ \phi_{22} \end{bmatrix}$$

$$k=3 \to \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{bmatrix}}_{A} \begin{bmatrix} \phi_{31} \\ \phi_{32} \\ \phi_{33} \end{bmatrix}, \text{ etc.}$$

$$k=4 \to \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \rho_3 \\ \rho_1 & 1 & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & 1 & \rho_1 \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{bmatrix} \begin{bmatrix} \phi_{41} \\ \phi_{42} \\ \phi_{43} \\ \phi_{44} \end{bmatrix}, \text{ etc.}$$

Computing 3x3 det:

$$\det\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = a\det\begin{bmatrix} e & f \\ h & i \end{bmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}.$$

4x4?

$$\begin{vmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{vmatrix} = a\begin{vmatrix} f & g & h \\ j & k & l \\ n & o & p \end{vmatrix} - b\begin{vmatrix} e & g & h \\ i & k & l \\ m & o & p \end{vmatrix} + c\begin{vmatrix} e & f & h \\ i & j & l \\ m & n & p \end{vmatrix} - d\begin{vmatrix} e & f & g \\ i & j & k \\ m & n & o \end{vmatrix}$$

---

e.g. Suppose $E(S^2) = \frac{n}{n-1}\gamma_0 - \frac{n}{n-1}Var(\bar{Z})$.

Show that $\frac{E(S^2)}{\gamma_0} = 1 - \frac{2}{n-1}\left[\sum_{k=1}^{n-1}\rho_k - \frac{1}{n}\sum_{k=1}^{n-1}k\rho_k\right]$.

So $\frac{1}{\gamma_0}E(S^2) = \frac{n}{n-1} - \frac{n}{(n-1)}\frac{Var(\bar{Z})}{\gamma_0}$.

$$Var(\bar{Z}) = \frac{1}{n^2}\sum_{t=1}^{n}\sum_{s=1}^{n} Cov(Z_t, Z_s)$$

$$= \frac{1}{n^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\gamma_{|t-s|}$$

$$\frac{Var(\bar{Z})}{\gamma_0} = \frac{1}{n^2}\sum_t\sum_s \rho_{|t-s|} = \frac{1}{n^2}\sum_{t=1}^{n}\left[\rho_{|t-1|} + \rho_{|t-2|} + \cdots + \rho_{|t-n|}\right]$$

$$\cancel{= \frac{1}{n^2}\sum_{t=1}^{n}\sum} = \frac{1}{n^2}\left[n + 2(n-1)\rho_1 + 2(n-2)\rho_2 + \cdots + 2(n-(n-1))\rho_{n-1}\right]$$

$$\cancel{= \frac{1}{n^2}\left[2\sum_{k=-(n-1)}^{n-1}(n-|k|)\rho_{|k|}\right]}$$

$$= \frac{1}{n^2}\left[n + \sum_{k=1}^{n-1}2(n-k)\rho_k\right]$$

Thus $\frac{E(S^2)}{\gamma_0} = \frac{n}{n-1} - \frac{n}{n-1}\cdot\frac{1}{n^2}\left[n + \sum_{k=1}^{n-1}2(n-k)\rho_k\right]$

$$= \frac{n}{n-1} - \frac{1}{(n-1)n}\left[n + \sum_{k=1}^{n-1}2(n-k)\rho_k\right]$$

$$= \frac{n}{n-1} - \frac{1}{n-1} - \frac{1}{n(n-1)}\sum_{k=1}^{n-1}2(n-k)\rho_k$$

$$= 1 - \frac{1}{(n-1)n}\sum_{k=1}^{n-1} 2(n-k)\rho_k$$

$$= 1 - \frac{2}{n(n-1)}\sum_{k=1}^{n-1}(n-k)\rho_k$$

$$= 1 - \frac{2}{(n-1)}\sum_{k=1}^{n-1}\frac{n-k}{n}\rho_k$$

$$= 1 - \frac{2}{n-1}\left[\sum_{k=1}^{n-1}\left(1-\frac{k}{n}\right)\rho_k\right]$$

$$= 1 - \frac{2}{n-1}\left[\sum_{k=1}^{n-1}\rho_k - \frac{1}{n}\sum_{k=1}^{n-1}k\rho_k\right].$$

Since $S^2 = \frac{1}{n-1}\sum_{t=1}^{n}(Z_t - \bar{Z})^2$ is a biased est. of TS, some suggested to use a corrective const.

$\hat{\sigma} = C_N S$, N the length of TS,

and $C_N = \left(\frac{N-1}{2}\right)^{1/2}\frac{\Gamma((N-1)/2)}{\Gamma(N/2)}$. $C_N \searrow S$.

S is a biased est. $\because$ Jensen's inequality: $\sqrt{S^2} = S$ (underest.)

As $N\uparrow$, $C_N \approx 1$.

$\widehat{ACF}$: $\hat{\gamma}_k$ and $\tilde{\gamma}_k$.

$$\hat{\gamma}_k = \frac{1}{n}\sum_{t=1}^{n-k}(Z_t - \bar{Z})(Z_{t+k} - \bar{Z})$$

$$\tilde{\gamma}_k = \frac{1}{n-k}\sum_{t=1}^{n-k}(Z_t - \bar{Z})(Z_{t+k} - \bar{Z}).$$

$$E(\hat{\gamma}_k) \approx \gamma_k - \frac{k}{n} - \frac{n-k}{n}Var(\bar{Z})$$

$$E(\tilde{\gamma}_k) = \gamma_k - Var(\bar{Z}).$$

$Y = \sum_{i=1}^{n}\alpha_i Z_{t_i}$

$Var(Y) = \sum\sum \alpha_i\alpha_j \gamma_{t_i - t_j}$

Comments ① $Bias(\hat{\gamma}_k) > Bias(\tilde{\gamma}_k)$ especially k large wrt n.

② $\hat{\gamma}_k$ is positive semidefinite but $\tilde{\gamma}_k$ isn't.

③ $Var(\hat{\gamma}_k) < Var(\tilde{\gamma}_k)$.

$$\Phi_p(B)x_t = \Theta_q(B)a_t$$

~~ARIMA(p,d,q), Unit root test~~

	ACF	PACF	Comment
AR(p)	continues	0 if k>p	Always invertible
MA(q)	0 if k>q	continues	Always stationary (causal)
ARMA(p,q)	0 if k>q-p	0 if k>p-q	

~~Model adequacy~~

— Wold decomposition

Any zero-mean process $\{X_t\}$ which is not deterministic can be expressed as a sum of $X_t = U_t + V_t$, where $\{U_t\}$ denotes an MA(∞) process and $\{V_t\}$ is a deterministic process which is uncorrelated w/ $\{U_t\}$.

$\{X_t\}$ is deterministic if the values X_{n+j}, $j \geq 1$, of the process $\{X_t \mid t = 0, \pm 1, \pm 2, \cdots\}$ were perfectly predictable in terms of $\mu_n \in \text{span}\{X_t \mid -\infty < t \leq n\}$.

— Model adequacy or diagnostic checking

Parameters are estimated by MoM, MLE, Kalman Filter, etc.

$$\text{ARMA}(p,q): \quad \Phi_p(B)X_t = \Theta_q(B)a_t.$$

$$\Leftrightarrow \quad \Phi_p(B)X_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_q a_{t-q}.$$

$$\Rightarrow \quad \hat{\Phi}_p(B)X_t - \hat{\theta}_1 a_{t-1} - \hat{\theta}_2 a_{t-2} - \cdots - \hat{\theta}_q a_{t-q} = \hat{a}_t.$$

ACF among residuals:

$$\hat{\rho}_k = \frac{\sum_{t=1}^{n-k} \hat{a}_t \hat{a}_{t+k}}{\sum_{t=1}^{n} \hat{a}_t^2} = \text{ACF of residual at lag } k.$$

$$\hat{\rho}_k = \frac{\sum_{t=1}^{n-k} \hat{a}_t \hat{a}_{t+k}}{\sum_{t=1}^{n} \hat{a}_t^2}$$

* Portmanteau test

→ the overall test that checks an entire group of residual ACFs assuming that the model is adequate.

→ Two popular tests:

$$Q_{BP} = n \cdot \sum_{k=1}^{m} \hat{\rho}_k^2 \sim \chi^2_{m-(p+q)}$$

$$Q_{LB} = \sum_{k=1}^{m} \frac{n(n+2)}{(n-k)} \hat{\rho}_k^2 \sim \chi^2_{m-(p+q)}$$

→ Pros: practical purposes; minimal req. to use the fitted model

→ Cons: lack power if comparing w/ traditional statistical tests, e.g. LRT.

Sol'ns

① Finite sample adjustments

② Complicated functional of residual ACFs

③ Monte Carlo test

④ Other applications: portmanteau tests for randomness and ARMA models w/ infinite variance innovations

— Model selection

→ Akaike Information Criterion (AIC)

→ Bayesian 〃 (BIC)

$$AIC = -2\log ML + 2k$$

$$BIC = -2\log ML + k\log(n)$$

ML denotes MLE, $\log ML$ is the value of maximized log-likelihood function for a model fitted to a given data, and k is the # of independently adjusted parameters w/i the model

<u>Remark</u> BIC puts more ~~penalty~~ penalties on the number of parameters used by fitted models, and some empirical studies indicate that the model selected by BIC performs better in the post-sample analysis, such as forecasting.

Select k w/ the lowest AIC/BIC.

* ARIMA

<u>Def'n</u> $\{X_t\}$ is said to follow an ARIMA model of order (p,d,q) if $W_t = (1-B)^d X_t$ is a stationary ARMA model.

So $(1-B)^d \Phi_p(B) X_t = \Theta_q(B) a_t$, $a_t \sim N(0,\sigma^2)$.

<u>Def'n</u> $\nabla^d = (1-B)^d$.

Suppose X_t is stationary, and $Y_t = a + bt + ct^2 + X_t$.

Then $\nabla^d Y_t = \nabla^d a + \nabla^d bt + \nabla^d ct^2 + \nabla^d X_t$. Let $d=2$.

$(1-B)^2 a = (1-2B+B^2) a = a - 2a + a = 0$.

$(1-B)^2 bt = (1-2B+B^2) bt$
$= bt - 2b(t-1) + b(t-2)$
$= 2bt - 2bt + 2b - 2b = 0$.

$(1-B)^2 ct^2 = (1-2B+B^2) ct^2$
$= ct^2 - 2c(t-1)^2 + c(t-2)^2$
$= ct^2 - 2c(t^2-2t+1) + c(t^2-2t+1)$
$= ct^2 - c(t^2-2t+1) = ct^2 - ct^2 + 2ct - c$
$= 2ct - c$

$(1-B)^2 ct^2$
$= (1-2B+B^2)ct^2$
$= ct^2 - 2c(t-1)^2 + c(t-2)^2$
$= ct^2 - 2c(t^2-2t+1) + c(t^2-4t+4)$
$= \cancel{ct^2} - \cancel{2ct^2} + \cancel{4ct} - 2c + \cancel{ct^2} - \cancel{4ct} + 4c$
$= \underline{2c}$

Is $(1-B)^2 X_t$ stationary?

Since X_t is stationary, we have
$E(X_t) = m$, $E(|X_t|^2) < \infty$,
and $\gamma_x(t,s) = \gamma_x(r+t, r+s)$ $\forall t,r,s \in T$.

$t-k \quad t$
$\gamma_x(t_1, t_2)$
$= \gamma_x(t-k, t)$
$= \gamma_x(t, t+k)$
$=: \gamma_k$.

$(1-B)^2 X_t = (1-2B+B^2)X_t = X_t - 2X_{t-1} + X_{t-2}$.

$E(X_t - 2X_{t-1} + X_{t-2})$
$= m - 2m + m = 0$. ✓

$E(|X_t - 2X_{t-1} + X_{t-2}|^2)$
$\leq E(|X_t|^2 + |-2X_{t-1}|^2 + |X_{t-2}|^2) < \infty$. ✓

$E((X_{t_1} - 2X_{t_1-1} + X_{t_1-2})(X_{t_2} - 2X_{t_2-1} + X_{t_2-2}))$ $(\because \mu = 0)$

~~$E(X_{t_1} X_{t_2}) - 2E(X_{t_1-1} X_{t_2})$~~

$Cov(X_{t_1} - 2X_{t_1-1} + X_{t_1-2}, X_{t_2} - 2X_{t_2-1} + X_{t_2-2})$
$= Cov(X_{t_1}, X_{t_2}) - 2Cov(X_{t_1-1}, X_{t_2}) + Cov(X_{t_1-2}, X_{t_2})$
$- 2Cov(X_{t_1}, X_{t_2-1}) + 4Cov(X_{t_1-1}, X_{t_2-1}) - 2Cov(X_{t_1-2}, X_{t_2-1})$
$+ Cov(X_{t_1}, X_{t_2-2}) - 2Cov(X_{t_1-1}, X_{t_2-2}) + Cov(X_{t_1-2}, X_{t_2-2})$.

So it is a function of lag τ. ✓

$\therefore$ $(1-B)^2 X_t$ is also stationary. //

Def'n
$\nabla_d = (1-B^d)$.

So $\nabla^d = (1-B)^d$, $\nabla_d = (1-B^d)$.
e.g. $\nabla_d X_t = (1-B^d)X_t = X_t - X_{t-d}$.

* $I(d)$ process and DF unit root test.

Def'n A series follows a stationary ARMA model after differencing d times is said to be integrated of order d, or $I(d)$ process.
i.e. X_t is $I(d)$ if $W_t = (1-B)^d X_t$ is stationary.

Def'n A Dickey-Fuller test tests the H_0 whether a unit root is present in an AR model. The H_1 is usually ~~stationarity~~ or trend-stationarity.

e.g. Test $I(1)$.

$X_{t-1} = \phi X_{t-2} + a_{t-1}$

$$X_t = \phi X_{t-1} + a_t, \quad a_t \sim NID(0, \sigma^2).$$
$$\nabla X_t = \phi \nabla X_{t-1} + \nabla a_t$$
$$X_t - X_{t-1} = \phi(X_{t-1} - X_{t-2}) + a_{t-1}$$
$$X_t - X_{t-1} = \phi X_{t-1} - \phi X_{t-2} + a_{t-1}$$
$$= \phi X_{t-1} - (X_{t-1} - a_t) + a_{t-1}$$
$$X_t - \cancel{X_{t-1}} = \phi X_{t-1} - \cancel{X_{t-1}} + a_t + a_{t-1}$$
$$X_t = \phi X_{t-1} + a_t + a_{t-1}$$

$$X_t = \phi X_{t-1} + a_t$$
$$X_t - X_{t-1} = \phi X_{t-1} - X_{t-1} + a_t$$
$$:= \nabla X_t = \underbrace{(\phi - 1)}_{\pi} X_{t-1} + a_t.$$

$H_0: \pi = 0$ vs. $H_1: \pi > 0$
or $X_t \sim I(1)$ — or $X_t \sim I(0)$.

- General DF-test:

$$\nabla X_t = a + \tau^T DR_t + \pi X_{t-1} + a_t,$$

$$\left(\begin{array}{l} \tau^T = (\alpha_1, \alpha_2, \cdots) \\ DR_t = (t, t^2, t^3, \cdots), \end{array} \right.$$

a = regression intercept

- Problems w/ DF test:
  - → DF test considers only a single unit root.
  - → " assumes a correct model specification, i.e. correct specification of time trend and intercept. DGP may contain both AR and MA terms, and there might be structural breaks in the data.

- Detecting multiple roots

If more than one unit root is suspected, then perform DF tests on successive differences of X_t.

e.g. Two roots suspected.

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

$$\nabla X_t = (\phi - 1) X_{t-1} + a_t$$

$$= \pi_1 X_{t-1} + a_t$$

$$\nabla^2 X_t = \pi_1 \nabla X_{t-1} + a_t$$

Use the appropriate statistic to determine whether π is sig. diff. from 0. If you cannot reject $H_0: \pi_1 = 0$, conclude that $X_t \sim I(2)$.

If $\pi_1 \approx 0$, determine if there is a single unit root:

~~$\nabla^2 X_t = \pi_1 \nabla X_{t-1} + \pi_2 \nabla X_{t-1} + a_t$~~

$$\nabla^2 X_t = \pi_1 \nabla X_{t-1} + \pi_2 X_{t-1} + a_t.$$

Single unit root $\Rightarrow$ $H_0: \pi_1 < 0$ and $\pi_2 = 0$.

We can use the DF critical values to test the null hypothesis $\pi_2 = 0$.

Reject $H_0 \Rightarrow X_t$ is stationary.

* Augmented DF (ADF) test

$$\nabla X_t = \tau^T DR_t + \pi X_{t-1} + \sum_{j=1}^{k} \gamma_j \nabla X_{t-j} + a_t, \quad k = p-1.$$

The equation uses the autoregression to take into account the presence of serial correlated errors.

Selection of p:

① Autoregression approximation

$$ARIMA(p,1,q) \approx ARIMA(n,1,0), \quad n \leq T^{1/3}.$$

② General-to-specific methodology.

Start w/ p^*. If the t-statistic of lag p^* is insignificant at some specified critical value, re-estimate the regression using p^*-1. Repeat the process until the last lag is sig. diff. from 0.

Once our tentative lag length is determined, conduct the diagnostic checking such as looking at res. autocorrelation plot or conducting portmanteau tests on regression residuals.