Chapter 2

Instructor: Pan Guang-Ming

Division of Mathematical Sciences, SPMS
Nanyang Technological University, Singapore

Outline

1. Stationary
   - *Strict stationary*
   - *Weakly stationary*
2. ARIMA(p,d,q) models
   - *MA(q)*
   - *AR(p)*
   - *ARMA*
   - *ARIMA*
3. SACF and SPACF
   - *SACF*
   - *SPACF*

Stationary

Loosely speaking, a time series $\{X_t, t = 0, \pm 1, \cdots\}$ is said to be stationary if it has statistical properties similar to those of the "time shifted" series $\{X_{t+h}, t = 0, \pm 1, \cdots\}$ for each integer h. We can make this idea precise with the following definitions.

Mean and Coariance

Definition

The expectation function of X is defined as

$$\mu(t) = E[X_t], \;\; t \in T.$$

And the covariance function of X is given by

$$\begin{aligned}\gamma(t,s) &= \operatorname{cov}(X_t, X_s) \\ &= E[(X_t - \mu_X(t))(X_s - \mu_X(s))]\end{aligned}$$

for all $t, s \in T$.
The variance function is defined by

$$\sigma_X^2(t) = \gamma(t,t) = \operatorname{var}(X_t).$$

Thus $\mu(t)$, $\operatorname{var}(X_t)$ and $\gamma(t,s)$ are just real functions of t and (t,s) respectively.

Example of Gaussian process

Example

Consider the Gaussian process $(X_t, t \in [0,1])$ of i.i.d. $N(0,1)$ random variables X_t.
Its expectation and covariance functions are given by

$$\begin{aligned} \mu_X(t) &= 0 \text{ and} \\ \gamma(t,s) &= 1 \text{ if } t = s;\ 0 \text{ otherwise.} \end{aligned}$$

Strictly stationary

Definition

A time series is said to be strictly stationary if, for any $n \in \mathbb{Z}^+$ and all integers h, $(X_1, ..., X_n)$ and $(X_{1+h}, \cdots, X_{n+h}), s \in \mathbb{Z}\}$ have the same distributions.

Such an example is i.i.d random variables.

Weakly Stationary

Definition

Denote

$$\mu_t = EX_t \quad \text{and} \quad \gamma(t,k) = \text{cov}(X_t, X_{t+k}), \quad t, k \in \mathbb{Z}.$$

A time series is said to be weakly stationary if (a) $\mu_t = \mu$ is independent of t; and (b) $\gamma(t,k)$ is independent of t for each k.

Strictly stationary against weakly stationary

- Weak stationarity is also called stationarity directly since we are mainly interested in this type of stationarity.
- Strict stationarity and finiteness of second moment ensure weak stationarity.
- However, generally speaking, weak stationarity does not imply strict stationarity.
- An exceptional case is that the strict stationarity is the same as the weak stationarity when X_t is Gaussian process.
- Note that some distributions may have infinite second moment such as Cauchy distribution.

How to verify weakly stationary ?

Usually, we first have a look at whether either

$$\mu(t) = c_0 \quad \text{or} \quad \mathrm{var}(X_t) = c_1. \tag{1}$$

If so, we need to look at whether the following holds

$$\gamma(t, s) = h(t - s). \tag{2}$$

If (1) doesn't hold, then we will conclude that X_t is not weakly stationary. The verification of (2) is then not required.

Example of linear model (I)

Example

Consider

$$X_t = \beta_0 + \beta_1 t + e_t,$$

where e_t's are i.i.d. with mean zero and variance one, and $\beta_1 \neq 0$. Is y_t stationary ?

Example of linear model (II)

Solution

It is not stationary because

$$EX_t = \beta_0 + \beta_1 t,$$

dependent on time t.

Example of random walk (I)

Example

Consider a random walk of the form

$$X_t = X_{t-1} + e_t, \;\; t = 1, 2, \ldots,$$

where e_t is a sequence of i.i.d. random variables with $E[e_t] = 0$ and $E[e_t^2] = 1$. Let $X_0 = 0$.

- Justify whether X_t is stationary.
- Is $Z_t = X_t - X_{t-1}$ stationary ?

Example of random walk(II)

Solution

- *It is not stationary. Indeed, we may rewrite*

$$X_t = e_t + e_{t-1} + \cdots + e_1$$

so that

$$\begin{aligned} Var(X_t) &= Var(e_t + e_{t-1} + \cdots + e_1) \\ &= Var(e_t) + Var(e_{t-1}) + \cdots + Var(e_1) = t, \end{aligned}$$

dependent on time t.

- *Yes, it is stationary, because*

$$Z_t = X_t - X_{t-1} = e_t.$$

Autocovariance and autocorrelation functions

In view of condition (b) in Definition 4, whenever we use the term covariance function with reference to a stationary process $\{X_t\}$, $\gamma(t, k)$ is a function of one variable, time lag k and hence denote it by γ_k.

Definition

Let $X(t)$ be a stationary time series. The autocovariance function at lag k is

$$\gamma_k = \text{cov}(X_t, X_{t+k}) = E(X_t - \mu)(X_{t+k} - \mu), \qquad k \in \mathbb{Z};$$

and its autocorrelation function (ACF) is

$$\rho_k = \frac{\text{cov}(X_t, X_{t+k})}{\sqrt{\text{var}(X_t)\text{var}(X_{t+k})}} = \gamma_k/\gamma_0 \qquad k \in \mathbb{Z}.$$

Property of ACF and ACVF

The two functions have the following properties,

- $\gamma_0 = \mathrm{var}(X_t)$; $\rho_0 = 1$.
- $\gamma_k = \gamma_{-k}$; $\rho_k = \rho_{-k}$.
  Therefore, the ACFs are often plotted only for the nonnegative lags.

Example of IID

Example

If Z_t is i.i.d with variance $Var(Z_t) = \sigma^2$ then for any s, t with $s \neq t$,

$$Cov(Z_t, Z_s) = 0$$

and $\{Z_t\}$ is stationary.

White noise

Definition

The process Z_t is said to be white noise if it is a stationary with $Cov(Z_t, Z_s) = 0$ for $s \neq t$, and $Var(Z_t) = \sigma^2$. Denote it by $z_t \sim WN(0, \sigma_z^2)$.

MA(1) model (I)

Example

We can build time series with white noise sequence. Suppose $Z_t \sim WN(0, \sigma^2)$. Let

$$X_t = Z_t + \theta Z_{t-1} \qquad (MA(1) model).$$

Purpose of defining sample ACF and ACVF

- Although we have just seen how to compute the autocorrelation function for a few simple time series model, in reality, we do not start with a model but with observations $x_1, x_2, \cdots, x_n$.
- Suppose that $\{x_t\}$ is stationary, then we have to estimate μ_x, $\gamma_x(h)$ and $\rho_x(h)$, with $h = 0, 1, 2, \cdots$.
- This estimate may suggest which of the many possible stationary models is a suitable candidate for representing the dependence in the data.
- For example, a sample ACF that is close to zero for all nonzero lags suggests that an appropriate model for the data might be i.i.d noise.

SACVF and SACF

Definition

The sample mean is

$$\hat{\mu}_x = \bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t.$$

The sample autocovariance function (SACVF) is

$$\hat{\gamma}(h) = \frac{1}{n}\sum_{t=1}^{n-h}(x_{t+h} - \bar{x})(x_t - \bar{x}).$$

The sample autocorrelation function (SACF) at lag h is

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} = \frac{\sum_{t=1}^{n-h}(x_{t+h} - \bar{x})(x_t - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}.$$

The Moving average

Definition

The process $\{X_t\}$ defined by

$$X_t = Z_t + \sum_{j=1}^{q} \theta_j Z_{t-j} = \sum_{j=0}^{q} \theta_j Z_{t-j}, \ \theta_0 = 1,$$

where $Z_t \sim WN(0, \sigma^2)$, is called a *moving average process of order* q and is denoted by MA(q). We always assume $\theta_i = 0$ for $i \leq -1$.

An example of MA(q)

Example

Let X_t be grocery sales and Z_t be an unobserved (to the analyst) coupon campaign that varies in intensity over time. Coupon will expire after some period. At time 1, we have 100 unobserved coupons and assume that the take-up rate is always 50% (it is θ_1). So 50 incremental sales will take place at that time. At time 2, we have 80 new coupons and 50 remaining ones from the last period. This gives you $40 + 25 = 0.5 \cdot 80 + 0.5^2 \cdot 100$ bonus sales. Then you will get (suppose coupon expires after period 2)

$$X_t = 0.5Z_t + 0.5^2 Z_{t-1}.$$

The above model can be rewritten as

$$X_t = \tilde{Z}_t + 0.5\tilde{Z}_{t-1},$$

with $\tilde{Z}_t = 0.5Z_t$.

The Moving average

Is it stationary ?

Expectation of MA(q)

- $E[X_t] = 0$ since $E[Z_t] = 0$, independent of time t.
- Proof:

$$EX_t = EZ_t + \sum_{j=1}^{q} \theta_j E(Z_{t-j}) = 0.$$

Variance of MA(q)

$$\begin{aligned}\mathrm{var}(X_t) &= \mathrm{var}(Z_t + \sum_{j=1}^{q} \theta_j Z_{t-j}) \\ &= \sigma^2 + \sum_{j=1}^{q} \theta_j^2 \sigma^2 \\ &= \sigma^2 \left(1 + \sum_{j=1}^{q} \theta_j^2\right),\end{aligned}$$

where $\sigma^2 = \mathrm{var}(Z_t)$, independent of time t.

Covariance functions of MA(q)

$$\gamma(\mathrm{k}) = \mathrm{cov}(X_{t-k}, X_t) = \mathrm{cov}(\sum_{j=0}^{q} \theta_j Z_{t-k-j}, \sum_{l=0}^{q} \theta_l Z_{t-l})$$
$$= (\theta_k + \theta_1\theta_{k+1} + \cdots + \theta_{q-k}\theta_q)\,\sigma^2 \quad \text{for } k \le q.$$

When $k > q$, $\mathrm{cov}(X_{t-k}, X_t) = 0$, because $k + j > q \ge l$.

Correlation functions of MA(q)

Its correlation functions are obtained by dividing $\gamma(k)$ with $\gamma(0)$:

$$\rho(k) = \begin{cases} 1, & k = 0 \\ \frac{\theta_k + \theta_1\theta_{k+1} + \cdots + \theta_{q-k}\theta_q}{1 + \sum_{j=1}^{q} \theta_j^2}, & 1 \le k \le q \\ 0, & k > q \end{cases} \tag{3}$$

Since the mean and covariance function are both independent of time t we conclude that MA(q) is stationary.

Correlation functions of MA(q)

Generally, the ACF for a MA(q) time series cut-off after lag q

Example of MA(2)

Example

(MA(2) process): If $Z_t \sim WN(0, \sigma^2)$, then $\{X_t\}$ defined by

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2}$$

is an MA(2) process. Find its ACFs,

ACFs of MA(2)

Solution

It follows that the ACFs are given by

$$\rho(1) = \frac{\theta_1 + \theta_1\theta_2}{1 + \theta_1^2 + \theta_2^2},$$

$$\rho(2) = \frac{\theta_2}{1 + \theta_1^2 + \theta_2^2},$$

$$\rho(k) = 0 \ \forall \ k \geq 3.$$

Simulated MA(1) model

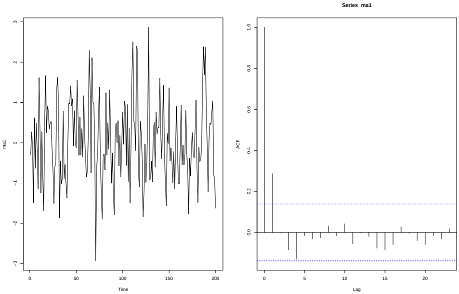


Figure: Plot of $MA(1)$ with $\theta_1 = 0.2$ and its acf

R codes for simulating MA(1) and plot ACF

```
ma1=arima.sim(list(order=c(0,0,1),ma=0.2), n = 200)
ts.plot(ma1)
acf(ma1)
```

Example of differencing of stationary process(I)

Example

Let $Z_t \sim WN(0, \sigma^2)$. Consider $\nabla^k Z_t$, which appeared when applying the differencing operator to produce stationary process. Prove that it is stationary.

Example of differencing of stationary process(II)

Solution

It is easily seen that

$$\nabla Z_t = Z_t - Z_{t-1}$$

is a MA(1) and

$$\nabla^2 Z_t = Z_t - 2Z_{t-1} + Z_{t-2}$$

is a MA(2). Similarly one may see that $\nabla^k Z_t$ *is a MA(k) model. Therefore it is stationary.*

MA(∞)

Definition

Let

$$X_t = \sum_{j=1}^{\infty} \psi_j Z_{t-j}$$

where $\psi_0 = 1$ and $Z_t \sim WN(0, \sigma^2)$ and

$$\sum_{j=0}^{\infty} |\psi_j| < \infty.$$

$\{X_t\}$ is called a **general linear process** or a **MA**(∞) process

AR(p)(I)

Definition

Assume that $\{Z_t\}$ is white noise. Then the process $\{X_t\}$ defined by

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + Z_t$$

is called an autoregressive process of order p, denoted by $AR(p)$.

Recall the following backshift operator B,

$$BX_t = X_{t-1} \text{ and } B^i X_t = X_{t-i}$$

for $i = 1, 2, \ldots$.

Real examples of AR(p)

- The current time series value X_t is a function of past time series values $X_{t-1}, \cdots, X_{t-p}$.
- Autoregressive models are heavily used in economic forecasting.
- For the GDP growth series, an autoregressive model of order one uses only the information on GDP growth observed in the last quarter to predict a future growth rate.
- The change in inflation is AR(p) model.

AR(p)(II)

Is AR(p) stationary ?

Condition to ensure stationarity

Define the p^{th} degree polynomial,

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

We now can rewrite the model of X_t by

$$\phi(B)X_t = Z_t,$$

where $\phi(B) = 1 - \sum_{j=1}^{p} \phi_j B^j$. It is called the AR *characteristic polynomial*.

Theorem

$AR(p)$ is stationary if all the roots of the polynomial $\phi(z)$ are outside the unit circle.

Convert to MA(∞)

Assume that there is a sequence of constants $\{\psi_j : j = 0, 1, \ldots\}$ such that

$$X_t = (\phi(B))^{-1} Z_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j} \tag{4}$$

(so called **Causality**). It follows that $\{X_t\}$ is stationary if and only if

$$\sum_{j=0}^{\infty} |\psi_j| < \infty.$$

Convert AR(1)to MA(∞)

Example

The characteristic polynomial for

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

is $\phi(z) = 1 - \phi z$. Evidently when $|\phi(z)| < 1$

$$\phi^{-1}(z) = \frac{1}{1 - \phi z} = \sum_{j=0}^{\infty} \phi^j z^j.$$

It follows that

$$X_t = \phi^{-1}(B) Z_t = \sum_{j=0}^{\infty} \phi^j B^j Z_t = \sum_{j=0}^{\infty} \phi^j Z_{t-j}.$$

It is consistent with the previous example.

The expectation of stationary $AR(p)$(I)

$EX_t = 0$ because taking expectation on both sides of $AR(p)$ yields

$$(1 - \phi_1 - \cdots - \phi_p)EX_t = 0.$$

Why

$$(1 - \phi_1 - \cdots - \phi_p) \neq 0?$$

The expectation of stationary $AR(p)$(II)

Prove by contradiction: If

$$(1 - \phi_1 - \cdots - \phi_p) = 0,$$

then 1 is a root of the polynomial $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$, contradict to stationarity of $AR(p)$.

An important property of AR(p)

$E(Z_t X_{t-k}) = 0$ for any integer $k > 0$, because via (4)

$$E(Z_t X_{t-k}) = E\left(Z_t \sum_{j=0}^{\infty} \psi_j Z_{t-k-j}\right) = \sum_{j=0}^{\infty} \psi_j E\left(Z_t Z_{t-k-j}\right) = 0. \qquad (5)$$

Roughly speaking, this means that Z_t is uncorrelated with any value of X before time t.

The Yule-Walker equations

Theorem

The Yule-Walker equations are

$$\rho(k) = \phi_1\rho(k-1) + \phi_2\rho(k-2) + \cdots + \phi_p\rho(k-p). \; .$$

For $k \geq 1$, $\gamma(k) = \phi_1\gamma(k-1) + \phi_2\gamma(k-2) + \cdots + \phi_p\gamma(k-p)$.

Proof of The Yule-Walker equations

Multiplying X_{t-k} on both sides of AR(p) and then taking expectation we have

$$E(X_tX_{t-k}) = \phi_1 E(X_{t-1}X_{t-k}) + \phi_2 E(X_{t-2}X_{t-k}) + \cdots + \phi_p E(X_{t-p}X_{t-k})$$
$$+E(Z_tX_{t-k})$$

By (5)

$$\gamma(k) = cov(X_t, X_{t-k}) = E(X_tX_{t-k}) - EX_tEX_{t-k} = E(X_tX_{t-k}).$$

When $k \geq 1$ the above implies that

$$\gamma(k) = \phi_1\gamma(k-1) + \phi_2\gamma(k) + \cdots + \phi_p\gamma(p-k).$$

Dividing the above further by $\gamma(0)$ yields the Yule-Walker equations.

AR(1)(I)

Example

(AR(1) process continued): Consider an AR(1) model of the form

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

- (i) Show that X_t is stationary if $|\phi_1| < 1$.
- (ii) Find the ACFs.

AR(1) (II)

Solution

(i). If $|\phi_1| < 1$*, then the root to* $1 - \phi_1 z = 0$ *is* $z = 1/\phi_1$ *which lies outside the unit circle. Hence* $\{X_t\}$ *is stationary.*

AR(1) (III)

Solution

(ii) By the Yule-Walker equations

$$\rho_k = \phi_1 \rho_{k-1}.$$

This yields

$$\rho_k = \phi_1^k \rho_0 = \phi_1^k.$$

Hence

$$|\rho(k)| = |\phi_1|^k = e^{-k \log(|\phi_1|^{-1})} \to 0$$

exponentially. This fact is very useful in identifying an AR(1) model. Note that the exponential behaviour of the ACF of AR(1).

ARMA(p, q)

Definition

(The ARMA(p,q) Process). The process $\{X_t\}$ is said to be an auto-regressive moving average or ARMA(p,q) process if $\{X_t\}$ is stationary and for every t,

$$\begin{aligned} & X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} \\ & = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}, \end{aligned} \tag{6}$$

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

ARMA(p, q) with nonzero mean

We say that $\{X_t\}$ is an ARMA(p, q) process with mean μ if $\{X_t - \mu\}$ is an ARMA(p, q) process.

Equivalent expression of ARMA(p, q)

Model (6) can be written symbolically in the more compact form

$$\phi(B)X_t = \theta(B)Z_t, \ \ t = 0, \pm 1, \pm 2, \ldots,$$

where ϕ and θ are the p^{th} and q^{th} degree polynomials

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

and

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

and B is the backward shift operator defined by

$$B^j X_t = X_{t-j}, \ \ j = 0, \pm 1, \pm 2, \ldots.$$

AR(p) as a special ARMA(p,q)

Example

(The AR(p) Process). If $\theta(z) \equiv 1$ then

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

and the process is said to be an autoregressive process of order p (or AR(p)). This model can also be written as

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

MA(q) as a special ARMA(p,q)

Example

(The MA(q) Process). If $\phi(z) \equiv 1$ then

$$X_t = \theta(B)Z_t = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$$

and the process is said to be a moving-averaging process of order q (or MA(q)).

Condition to ensure stationarity of ARMA(p,q)

Theorem

The ARMA process is stationary if all roots of the polynomial of $\phi(z)$ are outside the unit root circle.

An application of Theorem

Example

Consider the model

$$X_t = X_{t-1} + Z_t.$$

Note that $\phi(z) = 1 - z$. Its root is one which does not lie outside the unit circle. It follows that it is not stationary.

Why PACF ?

The partial autocorrelation coefficient function, like the autocorrelation function, conveys vital information regarding the dependence structure of a stationary process.

Motivation of PACF(I)

To motivate the idea, consider an $AR(1)$ model,

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

Then

$$\begin{aligned}\gamma(2) &= cov(X_t, X_{t-2}) = cov(\phi X_{t-1} + Z_t, X_{t-2}) \\ &= cov(\phi^2 X_{t-2} + \phi Z_{t-1} + Z_t, X_{t-2}) = \phi^2 \gamma(0),\end{aligned}$$

because Z_t and Z_{t-1} are both uncorrelated with X_{t-2} (so called causal property). The correlation between X_t and X_{t-2} is not equal to zero because X_t is dependent on X_{t-2} through X_{t-1}.

Motivation of PACF(II)

Suppose we break this chain of dependence by removing X_{t-1}. That is, consider the correlation between $X_t - \phi X_{t-1}$ and $X_{t-2} - \phi X_{t-1}$ because the correlation between X_t and X_{t-2} with the linear dependence of each on X_{t-1} removed. In this way, we have broken the dependence chain between X_t and X_{t-2}. Indeed,

$$cov(X_t - \phi X_{t-1}, X_{t-2} - \phi X_{t-1})$$

$$= cov(Z_t, X_{t-2} - \phi X_{t-1}) = 0.$$

The aim of introducing PACF

The aim of introducing PACF is to detect the relation between x_t and x_{t-h} with the effect of $x_{t-1}, \cdots, x_{t-h+1}$ removed.

$$x_t, \underbrace{x_{t-1}, \cdots, x_{t-h+1}}, x_{t-h}.$$

Definition of PACF

Definition

For a stationary process $\{X_t\}$, the partial autocorrelation function (PACF) is defined as

$$\begin{aligned} \phi_{11} &= corr(X_1, X_0) = \rho(1) \\ \phi_{kk} &= corr(X_k - f_{1,k-1}, X_0 - f_{2,k-1}),\ \ k \geq 2 \end{aligned}$$

where $f_{1,k-1} = f_1(X_{k-1}, \cdots, X_1)$ and $f_{2,k-1} = f_2(X_{k-1}, \cdots, X_1)$ minimize, respectively, the mean square linear prediction error

$$E(X_k - f_{1,k-1})^2, \quad E(X_0 - f_{2,k-1})^2$$

Usually, $f_1(X_{k-1}, \cdots, X_1) = \beta_1 X_1 + \cdots + \beta_{k-1} X_{k-1}$.

Explanation of PACF

For a stationary process $\{X_t\}$, the partial autocorrelation function (PACF) at lag k can be interpreted as the correlation function between X_0 and X_k after removing their linear dependency on $X_1, ..., X_{k-1}$.

PACFs of AR(p)(I)

Example

If $x_t \sim AR(p)$

$$x_t = \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + a_t,$$

then its PACFs of are

$$\begin{aligned}
\phi_{1,1} &= \rho(1), \\
\phi_{p,p} &= \phi_p, \\
\phi_{k,k} &= 0 \text{ for } \; k > p.
\end{aligned}$$

PACFs of AR(p)(II)

When $k > p$, the linear combination minimizing the mean square linear prediction error is

$$f_{1,p} = \phi_1 x_{k-1} + \cdots + \phi_p x_{k-p}$$

(obtaining such an expression is complicated and beyond the scope of the course). Then PACF at lag $k > p$ is Then PACF at lag $k > p$ is

$$\phi_{k,k} = corr(X_k - f_{1,p}, X_0 - f_{2,p}) = corr(Z_k, X_0 - f_{2,p}) = 0,$$

because Z_k is uncorrelated with the past values of X_{k-j} with $j > 0$. The calculation of $\phi_{p,p} = \phi_p$ is also beyond the scope of the course.

PACFs of AR(p)(III)

Theorem

Generally, the PACF for an AR(p) time series cut-off after lag p.

Equivalent definition of PACF

The general formula of ϕ_{kk} of an ARMA process is the last component of

$$\phi_h = \Gamma_h^{-1} g_h$$

where the matrix $\Gamma_h = [\gamma(i-j)]_{i,j=1}^{h}$ and $g_h = [\gamma(1), \gamma(2), \cdots, \gamma(h)]^T$ (this formula is for your reference only).

Summary of ACF and PACF of ARMA

Identification of ARMA model

model	ACF	PACF
AR(p)	dies down	cut-off after lag p
MA(q)	cut-off after lag q	dies down
ARMA(p,q)	dies down	dies down

Model for nonstationary time series

We have already discussed the class of ARMA models for representing stationary series. A generalization of this class, which incorporates a wide range of nonstationary series, is provided by the ARIMA processes, which, after differencing finitely many times, reduces to ARMA models.

ARIMA Process

Definition

(Auto–regressive Integrated Moving Average (ARIMA) Process)
A process $\{X_t\}$ is ARIMA(p, d, q) iff $\{\bigtriangledown^d X_t\}$ is a stationary ARMA(p, q) process.

ARIMA processes are **not stationary** when $d > 0$.

An example of nonstationary process

Example

Suppose that $\{X_t\}$ is a stationary process. Let $Y_t = X_t + c_1 + c_2 t + c_3 t^2$.
Is $\{Y_t\}$ stationary? $(1-B)Y_t$? $(1-B)^2 Y_t$?
Answer: $(1-B)^2 Y_t$ is stationary.
Can it be classified an ARIMA if X_t is white noise ?

ARIMA(0,1,0)

Example

ARIMA(0,1,0) Consider

$$(1-B)X_t = Z_t, \text{ or } X_t = X_0 + \sum_{i=1}^{t} Z_i,$$

where Z_t is white noise. The above special case is also called random walk. It can be classified as ARIMA(0,1,0).

ARIMA(0,1,1)

Example

(ARIMA$(0,1,1)$) Consider a time series model of the form

$$X_t = X_{t-1} + Z_t + \theta Z_{t-1}. \tag{7}$$

where $\{Z_t\}$ is a white noise.
Model (7) is an ARIMA$(0,1,1)$ since

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

defines $\{\bigtriangledown X_t\}$ an ARMA$(0,1)$=MA(1).

Attention: ARIMA

If p, d or q are zero, names can be abbreviated in the obvious way. e.g., ARMA(1,0) $\equiv$ AR(1); ARIMA(0,0,3) $\equiv$ MA(3).

Plot of $ARIMA(0,1,1)$

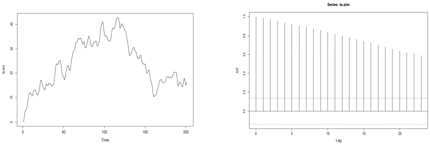
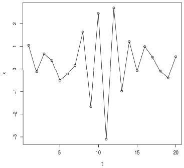


Figure: Plot of $ARIMA(0,1,1)$ with $\theta_1 = 0.8$ and its acf

A real data (I)

Example

Consider the following data set of x_t :
1.0445, -0.1338, 0.6706, 0.3755, -0.5110, -0.2352, 0.1595, 1.6258, -1.6739, 2.4478, -3.1019, 2.6860, -0.9905, 1.2113, -0.0929, 0.9905, 0.5213, -0.1139, -0.4062, 0.5438



Why ARMA models (I)?

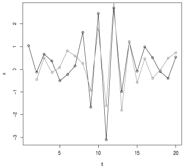
We cannot make any "good" prediction using trend and seasonality model.
Suppose we try the follow model:

$$x_t = 0.40 - 0.82 x_{t-1}$$

This is a AR(1) model (we will explain how we get 0.4 and -0.82 in the subsequent lectures).

A real data (II)

The fitted values are (time: prediction) 2: -0.4569, 3: 0.5130, 4: -0.1491, 5: 0.0938, 6: 0.8235, 7: 0.5965, 8: 0.2716, 9: -0.9354, 10: 1.7809, 11: -1.6121, 12: 2.9564, 13: -1.8082, 14: 1.2183, 15: -0.5942, 16: 0.47939, 17: -0.4124, 18: -0.0262, 19: 0.4966, 20: 0.73730



which can make reasonable prediction. AR model is USEFUL in prediction

Why ARMA models (II)

For the reason above (and more), we need to investigate models: AR(p), MA(q) .... More generally, stationary time series. To this end, we need answer the following questions.

- How to check stationarity ?
- ACF; PACF?
- Properties of AR, MA, ARMA models.
- How to calculate ACF and PACF for a given ARMA model.

Selecting a model based on SACF and SPACF

What will the SACF look like if the time series is not stationary ?

SACF of Nonstationary data

Here is an example.

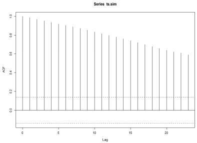


Figure: Plot of $ARIMA(0,1,1)$ with $\theta_1 = 0.8$ and its acf

R codes for simulating ARIMA models The R codes for generating the above the plots are as follows.

```
ts.sim<- arima.sim(list(order = c(0,1,1), ma = 0.7), n = 20
ts.plot(ts.sim)
acf(ts.sim)
```

General principles

- If the SAC dies down very quickly or even cuts off, then the time series should be considered stationary. If the SAC dies down extremely slowly, the time series values should be considered nonstationary.
- If the time series values (time plot) seem to fluctuate with constant variation around a constant mean μ, then it is reasonable to believe that the time series is stationary.

SACF

Given observations of a time series, one approach to the fitting of a model to the data is to match the sample ACF and sample PACF of the data with the ACF and PACF of the model, respectively. Suppose that $z_1, \cdots, z_n$ is a realization (sample, observations) from a stationary time series.
The sample ACF (SACF) at lag h is

$$r_h = \frac{\sum_{t=1}^{n-h}(z_t - \bar{z})(z_{t+h} - \bar{z})}{\sum_{t=1}^{n}(z_t - \bar{z})^2}.$$

SPACF

The sample PACF (SPACF) is

$$
\begin{aligned}
r_{1,1} &= r_1 \\
r_{k,k} &= \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j}, \text{ for } \ k = 2, 3, \cdots \\
r_{k,j} &= r_{k-1,j} - r_{k,k} r_{k-1,k-j}, \\
&\qquad \text{for } \ j = 1, 2, \cdots, k-1
\end{aligned}
$$

where $\bar{z} = \sum_{i=1}^{n} z_t / n$.

Standard error of r_h and r_{hh}

The *standard error of r_h* is

$$s_{r_h} = \left(\frac{1 + 2\sum_{j=1}^{h-1} r_j^2}{n} \right)^{1/2}.$$

and the *standard error of r_{hh}* is

$$s_{r_{hh}} = \left(\frac{1}{n} \right)^{1/2}.$$

Recall of summary of ACF and PACF of ARMA

Recall that

model	ACF	PACF
AR(p)	dies down quickly	cut-off after lag p
MA(q)	cut-off after lag q	dies down quickly
ARMA(p,q)	dies down quickly	dies down quickly
Nonstationary	dies down slowly	—

How to select MA(q) ?

Theorem

If the SACF r_h is significantly different from zero for $0 \leq h \leq q$ and negligible for $h > q$, then an $MA(q)$ model might be suitable for the data.

Meaning of "negligible" (I)

How do we classify ACF as negligible ? If (with probability approximately 0.95)

$$|r_h| \leq 1.96 s_{r_h},$$

then SACFs are considered as negligible. In practice, we frequently use the more stringent values $1.96/\sqrt{n}$ as the bound.

How to select AR(q) ?

Theorem

If the SPACF r_{hh} is significantly different from zero for $0 \leq h \leq p$ and negligible for $h > p$, then an $AR(p)$ model might be suitable for the data.

Meaning of "negligible" (I)

How do we classify PACF as negligible ? If (with probability approximately 0.95)

$$|r_{hh}| \leq \frac{1.96}{\sqrt{n}},$$

then SPACFs are considered as negligible.

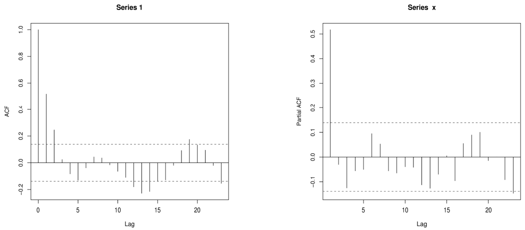
Exception in practice

- There is no theoretical model for which the ACF has nonzero values at lags $1, \cdots, q$ and cuts off after lag q, and the PACF has nonzero values at lags $1, \cdots, p$ and cuts off at all lags after lag p.
- However, in practice, sometimes for the time series values SACF has spikes at lags $1, \cdots, q$ and cuts off after lag q, and SPACF has spikes at lags $1, \cdots, p$ and cuts off after lag p.
- If this occurs, experience indicates that we should attempt to determine which of the SACF or SPACF is cutting off more abruptly. If the SACF cuts off more abruptly, then an MA(q) might be appropriate.

Can you identify them (I)?

Can you identify appropriate models for the following time series ?

Time series 1



Can you identify them (II)?

Time series 2