Time Series Analysis

Stationary ARMA Process

Binzhi Chen

Department of Economics
University of Birmingham

March 4, 2020

White noise

The basic building block for all the processes considered in this handout is a sequence $\{\epsilon_t\}_{t=-\infty}^{\infty}$ whose elements have mean zero and variance σ^2

$$E(\epsilon_t) = 0$$
$$E(\epsilon_t^2) = \sigma^2$$
$$E(\epsilon_t \epsilon_s) = 0$$

Then we have the *Gaussian white noise process*

Expectations

Suppose we have a sample size T of a random variable Y_t

$$\{y_1, y_2, \ldots, y_T\}$$

with a *Gaussian white noise* process

$$\{\epsilon_1, \epsilon_2, \ldots, \epsilon_T\} \qquad \epsilon_t \sim N(0, \sigma^2)$$

The *expectation* of the observation of a time series refers to the mean of the probability distribution:

$$E(Y_t) = \int_{-\infty}^{\infty} y_t f_{Y_t}(y_t) dy_t$$

Variance

The variance of the random variable Y_t is defined as:

$$Var(Y_t) = E[Y_t - E(Y_t)]^2 = \int_{-\infty}^{\infty} [Y_t - E(Y_t)]^2 f_{Y_t}(y_t) dy_t$$

Consider Y_t is a time trend plus a Gaussian white noise

$$Y_t = \beta t + \epsilon_t \tag{1}$$

The expected value of Y_t is βt and the variance of Y_t is σ^2

Autocovariance

The j th *autocovariance* of Y_t is:

$$\gamma_{jt} = \int\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} [Y_t - E(Y_t)][Y_{t-j} - E(Y_{t-j})]$$
$$\times f_{Y_t, Y_{t-1}, \ldots, Y_{t-j}}(y_t, y_{t-1}, \ldots, y_{t-j}) dy_t dy_{t-1} \ldots dy_{t-j} \tag{2}$$
$$= E[Y_t - E(Y_t)][Y_{t-j} - E(Y_{t-j})]$$

The autocovariance of equation (1) is:

$$\gamma_{jt} = E[Y_t - E(Y_t)][Y_{t-j} - E(Y_{t-j})] = E(\epsilon_t \epsilon_{t-j}) = 0$$

Autocorrelation Function and Coefficient Significant Test

Thus, autocorrelation can be calculated as:

$$\rho_j = \frac{\gamma_j}{\gamma_0}$$

Where $j = 1, 2, \ldots$

Single autocorrelation coefficient is significant and we will reject the null hypothesis if the coefficient lies inside the confidence interval:

$$(-1.96 \times \frac{1}{\sqrt{T}}, 1.96 \times \frac{1}{\sqrt{T}})$$

Where T is the sample size.

Autocorrelation Coefficient Significant Test

We can also test the joint hypothesis that all m of the ρ_k correlation coefficients are simultaneously equal to zero using the Q-statistic developed by Box and Pierce:

$$Q = T \sum_{k=1}^{m} \hat{\rho}_k{}^2$$

Where T is the sample size and m is the maximum lag length.

Another joint hypothesis test is Ljung-Box statistic:

$$Q^* = T(T+2) \sum_{k=1}^{m} \frac{\hat{\rho}_k{}^2}{T-K}$$

Both Q-statistic and Ljung-Box statistic are distributed as χ^2_m, but Ljung-Box statistic is more general in terms of the number of samples.

Autocorrelation Coefficient Significant Test

Suppose that a researcher had estimated the first 5 autocorrelation coefficients using a series of length 100 observations, and found them to be (from 1 to 5): 0.207, -0.013, 0.086, 0.005, -0.022.

Test each of the individual coefficient for significance.

Use both the Box-Pierce and Ljung-Box tests to establish whether the autocorrelation coefficients from previous slide are jointly significant. Compared with a tabulated $\chi^2_{(5)} = 11.1$ at the 5% level.

Stationary Conditions

If neither the mean μ nor the autocovariance γ_j depend on the date t then the process for Y_t is *stationary*

$$E(Y_t) = \mu \qquad \text{for all t} \tag{3}$$

$$E[Y_t - \mu]^2 = \sigma^2 \qquad \text{for all t} \tag{4}$$

$$E[Y_t - \mu][Y_{t-j} - \mu] = \gamma_j \qquad \text{for all t and j} \tag{5}$$

For example, equation (1) is not stationary because it is a function of time

MA Process

Consider a q th-order moving average process:

$$Y_t = \mu + \epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \cdots + \theta_q\epsilon_{t-q} \tag{6}$$

Where $\{\epsilon_t\}$ is a white noise process

Then the expectation is:

$$E(Y_t) = \mu + E(\epsilon_t) + E(\theta_1\epsilon_{t-1}) + E(\theta_2\epsilon_{t-2}) + \cdots + E(\theta_q\epsilon_{t-q}) = \mu$$

The variance is:

$$\begin{aligned} Var(Y_t) &= E[Y_t - \mu]^2 = E(\epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \cdots + \theta_q\epsilon_{t-q})^2 \\ &= (1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)\sigma^2 \end{aligned}$$

The j th covariance is:

$$\begin{aligned} \gamma_j &= [E(\epsilon_t) + E(\theta_1\epsilon_{t-1}) + E(\theta_2\epsilon_{t-2}) + \cdots + E(\theta_q\epsilon_{t-q}) \\ &\times (\epsilon_{t-j}) + E(\theta_1\epsilon_{t-j-1}) + E(\theta_2\epsilon_{t-j-2}) + \cdots + E(\theta_q\epsilon_{t-j-q})] \\ &= [\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \cdots + \theta_q\theta_{q-j}]\sigma^2 \end{aligned}$$

MA Process

Consider a MA(1) process:

$$Y_t - \mu = (1 + \theta L)\epsilon_t$$

If $|\theta| < 1$, multiply $(1 + \theta L)^{-1}$ to obtain:

$$(1 - \theta L + \theta^2 L^2 - \theta^3 L^3 + \dots)(Y_t - \mu) = \epsilon_t \tag{7}$$

which could be viewed as an AR(∞) representation.

If a moving average representation can be rewritten as an AR(∞) representation simply by inverting the moving average lag operator, then the moving average representation is said to be *invertible*

MA Process

Consider a MA(q) process:

$$Y_t - \mu = (1 + \theta_1 L + \theta_2 L^2 + \theta_3 L^3 + \cdots + \theta_q L^q)\epsilon_t$$

If the roots of

$$(1 + \theta_1 z + \theta_2 z^2 + \theta_3 z^3 + \cdots + \theta_q z^q) = 0 \tag{8}$$

lie outside the unit circle, then the process can be viewed as a AR(∞) representation:

$$(1 + \eta_1 L + \eta_2 L^2 + \eta_3 L^3 + \dots)(Y_t - \mu) = \epsilon_t \tag{9}$$

Where

$$(1 + \eta_1 L + \eta_2 L^2 + \eta_3 L^3 + \dots) = (1 + \theta_1 z + \theta_2 z^2 + \theta_3 z^3 + \cdots + \theta_q z^q)^{-1}$$

MA Process

Consider a MA(2) process:

$$Y_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2}$$

where ϵ_t is white noise process.

1. Calculate the mean and variance of Y_t
2. Derive the autocorrelation function for this process
3. If $\theta_1 = 0.25$ and $\theta_2 = 0.75$, is it invertible?

AR Process

Consider a p th-order autoregressive process:

$$Y_t = \mu + \phi Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \epsilon_t \tag{10}$$

Where $\{\epsilon_t\}$ is a white noise process

We can rewrite as:

$$(1 - \phi L - \phi_2 L^2 - \cdots - \phi_p L^p) Y_t = c + \epsilon_t$$

If the roots of:

$$(1 - \phi_1 z - \phi_2 z^2 - \phi_3 z^3 - \cdots - \phi_q z^q) = 0$$

lie outside the unit circle, the series is weakly stationary.

AR Process

The reason behind the absolute value of the coefficient in AR(1) less than unity, denoted $|\phi| < 1$, is the root of z.
Consider a first-order autoregressive process, denoted AR(1):

$$Y_t = c + \phi Y_{t-1} + \epsilon_t \tag{11}$$

Or:

$$(1 - \phi_1 L) Y_t = c + \epsilon_t$$

If AR(1) is stationary, the process must satisfy that the root of:

$$(1 - \phi z) = 0$$

lies outside the unit circle. Thus,

$$z = \frac{1}{\phi} > \pm 1$$

AR Process

Transforming equation(11) by recursive substitution approach we can get:

$$Y_t = \frac{c}{1-\phi} + \epsilon_t + \phi\epsilon_{t-1} + \phi^2\epsilon_{t-2} + \cdots + \phi^T Y_0 \tag{12}$$

If $\mu = \frac{c}{1-\phi}$, we transform a AR(1) process into MA(∞) process.

AR Process

Another reason behind the absolute value of the coefficient in AR(1) less than unity, denoted $|\phi| < 1$, is MA(∞) process.

1.If $|\phi| < 1$, $\phi^T \to 0$ as $T \to \infty$
So the shocks to the system gradually die away.

2.If $|\phi| = 1$, $\phi^T = 1$ for all T
So shocks persist in the system and never die away. The current value of y is just an infinite sum of past shocks plus some starting value of Y_0. We will recall this scenario later in Unit Root.

3.If $|\phi| > 1$
It has an intuitively unappealing property: shocks to the system are not only persistent through time, they are propagated so that a given shock will have an increasingly large influence. It is also not reasonable in many data series in economics and finance.

AR Process

Taking expectations in equation(12) we can get:

$$E(Y_t) = \mu + 0 + 0 + \cdots + 0 = \mu$$

The variance is:

$$\begin{aligned} Var(Y_t) &= E[Y_t - E(Y_t)]^2 \\ &= E(\epsilon_t + \phi\epsilon_{t-1} + \phi^2\epsilon_{t-2} + \dots)^2 \\ &= (1 + \phi^2 + \phi^4 + \cdots +) \cdot \sigma^2 \\ &= \frac{\sigma^2}{1 - \phi^2} \end{aligned}$$

AR Process

The j th autocovariance is:

$$
\begin{aligned}
\gamma_j &= E[Y_t - E(Y_t)]E[Y_{t-j} - E(Y_{t-j})] \\
&= E[\epsilon_t + \phi\epsilon_{t-1} + \phi^2\epsilon_{t-2} + \ldots] \times [\epsilon_{t-j} + \phi\epsilon_{t-j-1} + \phi^2\epsilon_{t-j-2}] \\
&= [\phi^j + \phi^{j+2} + \phi^{j+4} + \ldots] \cdot \sigma^2 \\
&= \frac{\phi^j}{1-\phi^2} \cdot \sigma^2
\end{aligned}
$$

What is the autocovariance for AR(1) process?

ARMA Process

Now we can turn to ARMA process:

$$\begin{aligned} Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} \\ + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} \end{aligned} \tag{13}$$

Or:

$$Y_t = c + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{i=1}^{q} \theta_q \epsilon_{t-q}$$

Or:

$$\begin{aligned} &(1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p) Y_t \\ &\qquad = c + (1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q) \epsilon_t \end{aligned} \tag{14}$$

Or:

$$\phi(L) Y_t = c + \theta(L) \epsilon_t$$

ARMA Process

Both sides can be divided by $(1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p)$ to obtain:

$$Y_t = \mu + \psi(L)\epsilon_t$$

Where

$$\psi(L) = \frac{(1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q)}{(1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p)}$$

and

$$\mu = c/(1 - \phi_1 - \phi_2 - \cdots - \phi_p)$$

Overparameterization Problem

Consider a simple white noise process:

$$Y_t = \epsilon_t \tag{15}$$

Multiple both sides by $(1 - \psi L)$:

$$(1 - \psi L)Y_t = (1 - \psi L)\epsilon_t \tag{16}$$

Clearly, if equation (15) a valid representation, then so is equation (16) for any value of ψ. Thus, equation (16) might be described as an ARMA(1, 1) process, with $\phi_1 = \psi$ and $\theta_1 = -\psi$. It is important to avoid such a parameterization. Since any value of ψ in equation (16) describes the data equally well, we will obviously get into trouble trying to estimate the parameter ψ in equation (16) by maximum likelihood. Moreover, theoretical manipulations based on a representation such as equation (16) may overlook key cancellations. If we are using an ARMA(1, 1) model in which ϕ_1 is close to $-\theta_1$ then the data might better be modeled as simple white noise.

Overparameterization Problem

A related overparameterization can arise with an ARMA(p, q) model. Consider factoring the lag polynomial operators:

$$\begin{aligned}(1-\lambda_1 L)(1-\lambda_2 L)\dots(1-\lambda_p L)Y_t \\ (1-\eta_1 L)(1-\eta_2 L)\dots(1-\eta_p L)\epsilon_t\end{aligned} \tag{17}$$

We assume that $|\lambda_i| < 1$, o that the process is stationary. If the autoregressive operator and the moving average operator have any roots in common, such as $\lambda_i = \eta_j$, then both sides can be divided by $(1-\lambda_i L)$:

$$\prod_{k=1, k\neq i}^{p}(1-\lambda_k L)Y_t = \prod_{k=1, k\neq j}^{q}(1-\eta_k L)\epsilon_t \tag{18}$$

Overparameterization Problem

Or:

$$(1-\phi_1^*L-\phi_2^*L^2-\cdots-\phi_{p-1}^*L^{p-1})Y_t = (1+\theta_1^*L+\theta_2^*L^2+\cdots+\theta_{q-1}^*L^{q-1})\epsilon_t$$

Where

$$(1-\phi_1^*L-\phi_2^*L^2-\cdots-\phi_{p-1}^*L^{p-1}) = (1-\lambda_1L)(1-\lambda_2L)\dots(1-\lambda_{i-1}L)(1-\lambda_{i+1}L)\dots(1-\lambda_pL)$$

$$(1+\theta_1^*L+\theta_2^*L^2+\cdots+\theta_{q-1}^*L^{q-1}) = (1-\eta_1L)(1-\eta_2L)\dots(1-\eta_{j-1}L)(1-\eta_{j+1}L)\dots(1-\eta_qL)$$

The stationary ARMA(p, q) process satisfying equation (14) is clearly identical to the stationary ARMA($p-1$, $q-1$) process satisfying.