Chapter2

January 22, 2019

1 Chapter 2: Linear Time Series Analysis and Its Applications

Finance 5330: Financial Econometrics Tyler J. Brough First Date: January 19, 2019 Last Date: January 21, 2019

1.1 Introduction

These notes are based on Chapter 2 of the book *Analysis of Financial Time Series 3rd Ed* by Ruey Tsay.

Understanding the simple time series models introduced here will go a long way to better appreciate the more sophisticated financial econometric models of later chapters.

Treating an asset return (e.g. log return r_t of a stock) as a collection of random variables over time., we have a time series $\{r_t\}$. The Linear time series models of this chapter are a natural first attemp at modeling such dynamic behavior.

The theories of linear time series discussed include:

- stationarity
- dynamic dependence
- autocorrelation function
- modeling
- forecasting

The econometric models introduced include:

- (a) simple autoregressive (AR) models
- (b) simple moving-average (MA) models
- (c) mixed autoregressive moving-average (ARMA) models
- (d) unit-root nonstationarity
- (e) regression models with times series errors
- (f) fractionally differenced models for long-range dependence

1.2 Section 2.1 Stationarity

The foundation of time series analysis is stationarity. A time series $\{r_t\}$ is said to be *strictly stationary* if the joint distribution of $(r_{t_1+t}, \ldots, r_{t_k+t})$ for all t, where k is an arbitrary positive integer and $(t_1, \ldots, t_k)$ is a collection of k positive integers.

Strict stationarity requires that the joint distribution of $(r_{t_1+t}, \ldots, r_{t_k+t})$ is invariant under time shift. This is a very strong requirement that is challenging to verify empirically. For this reason, we often employ a simpler form of stationarity.

A time series is $\{r_t\}$ *weakly stationary* if both the mean of r_t and the covariance between r_t and r_{t-l} are time invariant, where l is an arbitrary integer.

More specifically, $\{r_t\}$ is weakly stationary if:

- (a) $E(r_t) = \mu$, which is constant
- (b) $Cov(r_t, r_{t-l}) = \gamma_l$, which only depends on l

In practice, suppose that we have observed T data points $\{r_t|1, \ldots, T\}$. Weak stationarity implies that a time plot of the data would show that the T values fluctuate with constant variation around a fixed level. In application, weak stationarity enables one to make inference concerning future observations (e.g. prediction).

Implicitly, in the condition of weak stationarity, we assume that the first two moments of r_t are finite. From the definitions, if r_t is strictly stationary and its first two moments are finite, then r_t is also weakly stationary. The converse is not true in general.

If the time series r_t is normally distributed, then weak stationarity is equivalent to strict stationarity.

We will be mainly concerned with weakly stationary time series.

The covariance $\gamma_l = Cov(r_t, r_{t-1})$ is called the lag-l autocovariance of r_t. It has two important properties:

- (a) $\gamma_0 = Var(r_t)$
- (b) $\gamma_{-l} = \gamma_l$

The second property holds because $Cov(r_t, r_{t-(-l)}) = Cov(r_{t-(-l)}, r_t) = Cov(r_{t+l}, r_t) = Cov(r_{t_1}, r_{t_1-l})$, where $t_1 = t + l$.

In the finance literature, is common to assume that an asset return series is weakly stationary. We can check this empirically given a sufficient number of historical returns observations. In particular, we can divide the historical returns into subsamples and check the consistency of the results obtained across subsamples.

1.3 Section 2.2 Correlation and Autocorrelation Function

Recall that the correlation between two random variables X and Y can be defined as:

$$\rho_{x,y} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{E[(X-\mu_x)(Y-\mu_y)]}{\sqrt{E[(X-\mu_x)^2]E[(Y-\mu_y)^2]}}$$

This coefficient measures the strength between X and Y, and can be shown that $-1 \leq \rho_{x,y} \leq +1$, and that $\rho_{x,y} = \rho y, x$. The two random variables are uncorrelated if $\rho_{x,y} = 0$. In addition, if both X and Y are normally distributed random variables then the condition that $\rho_{x,y} = 0$ also indicates that they are independent.

When the sample $\{(x_t, y_t)\}_{t=1}^T$ then the population parameter can be estimated by its sample counterpart:

$$\hat{\rho}_{x,y} = \frac{\sum_{t=1}^T (x_t - \tilde{x})(y_t - \tilde{y})}{\sqrt{\sum_{t=1}^T (x_t - \tilde{x})^2) \sum_{t=1}^T (y_t - \tilde{y})^2}}$$

where $\tilde{x} = \frac{1}{T}\sum_{t=1}^T x_t$ and $\tilde{y} = \frac{1}{T}\sum_{t=1}^T y_t$ are the sample mean of X and Y, respectively.

Simulating Correlated Data We can simulate correlated data with the following algorithm:

1. Draw $z_1 \sim N(0,1)$
2. Draw $z_2 \sim N(0,1)$
3. Set $\epsilon_1 = z_1$
4. Set $\epsilon_2 = \rho z_1 + \sqrt{1-\rho^2} z_2$, where ρ is value of the correlation coefficient desired.

We can do this in Python as follows:

```
In [4]: import numpy as np

        M = 10000
        z1 = np.random.normal(size=M)
        z2 = np.random.normal(size=M)
        rho = 0.5
        e1 = z1
        e2 = rho * z1 + np.sqrt(1 - rho**2) * z2

        np.corrcoef(e1,e2)
```

```
Out[4]: array([[1.        , 0.50372534],
               [0.50372534, 1.        ]])
```

Autocorrelation Function (ACF) Consider a weakly stationary return series r_t. When the linear dependence between r_t and its past values r_{t-i} is of interest, the concept of correlation is generalized to autocorrelation.

The correlation coefficient between r_t and r_{t-l} is called the lag-l *autocorrelation* of r_t and is commonly denoted by ρ_l.

Specifically, we define

$$\rho_l = \frac{Cov(r_t, r_{t-l})}{\sqrt{Var(r_t)Var(r_{t-l})}} = \frac{Cov(r_t, r_{t-l})}{Var(r_t)} = \frac{\gamma_l}{\gamma_0}$$

For a given sample of returns $\{r_t\}_{t=1}^T$, let $\tilde{r}$ be the sample mean (i.e. $\tilde{r} = \frac{1}{T}\sum_{t=1}^T r_t$). Then the lag-1 sample autocorrelation of r_t is

$$\hat{\rho}_1 = \frac{\sum_{t=2}^T (r_t - \tilde{r})(r_{t-1} - \tilde{r})}{\sum_{t=1}^T (r_t - \tilde{r})^2}$$

Under general conditions, $\hat{\rho}_1$ is a consistent estimate of ρ_1. For example:

- If $\{r_t\}_{t=1}^T$ is an independent and identically distributed (iid) sequence
- And $E(r_t^2) < \infty$
- Then $\hat{\rho}_1$ is asymptotically normal with mean 0 and variance $1/T$

We can use this to test the hypothesis $H_0 : \rho_1 = 0$ against the alternative hypothesis $H_a : \rho_1 \neq 0$. The test statistic is the usual t ratio, which is $\sqrt{T}\hat{\rho}_1$ and follows asymptotically the standard normal distribution. The null hypothesis H_0 is rejected if the t ratio is large in magnitude, or if the p-value of the t ratio is small, say less than 0.05.

In general, the lag-l sample autocorrelation of r_t is defined as

$$\hat{\rho}_l = \frac{\sum_{t=l+1}^T (r_t - \tilde{r})(r_{t-l} - \tilde{r})}{\sum_{t=1}^T (r_t - \tilde{r})^2}, \qquad 0 \leq l < T - 1.$$

If $\{r_t\}$ is an iid sequence satisfying $E(r_t^2) < \infty$, then $\hat{\rho}_l$ is asymptotically normal with mean zero and variance $1/T$ for any fixed positive integer l.

More generally, if r_t is a weakly stationary time series satisfying $r_t = \mu + \sum_{i=0}^q \psi_i a_{t-i}$, where $\psi_0 = 1$ and $\{a_j\}$ is a sequence of iid random variables with mean zero, then $\hat{\rho}_l$ is asymptotically normal with mean zero and variance $(1 + 2\sum_{i=1}^q \hat{\rho}_i^2)/T$ for $l > q$. This is known as Bartlett's formula.

Testing Individual ACF For a given positive interger l, the previous result can be used to test $H_0 : \rho_l = 0$ vs $H_a : \rho_l \neq 0$. The test statistic is

$$t\text{ratio} = \frac{\hat{\rho}_l}{\sqrt{(1 + 2\sum_{i=1}^{l-1} \hat{\rho}_i^2)/T}}$$

If $\{r_t\}$ is a stationary Gaussian series satisfying $\rho_j = 0$ for $j > l$, the t ratio is asymptotically distributed as a standard normal random variable.

The decision rule is to reject H_0 if $|t\text{-ratio}| > z_{\alpha/2}$, where $z_{\alpha/2}$ is the $100(1 - \alpha/2)th$ percentile of the standard normal distribution.

1.4 Section 2.3 White Noise and Linear Time Series

White Noise A time series r_t is called a white noise if $\{r_t\}$ is a sequence of independent and identically distributed random variables with finite mean and variance. If r_t is normally distributed with mean zero and variance σ_2, the series called a Gaussian white noise.

For a white noise series, all of the ACFs are zero. In practice, as long as the ACFs are close to zero the process is treated as a white noise.

Linear Time Series A time series r_t is called linear if it can be written as

$$r_t = \mu + \sum_{i=0}^{\infty} \psi_i a_{t-i}$$

where μ is the mean of r_t, $\psi_0 = 1$, and $\{a_i\}$ is a sequence of iid random variables with mean zero and a well-defined distribution (i.e. $\{a_i\}$ is a white noise series).

It will be seen that $\{a_i\}$ denotes that new information at time t of the time series and is often referred to as the *innovation* or *shock* at time t.

If r_t is weakly stationary, we can obtain its mean and variance easily by using the independence of $\{a_t\}$ as

$$E(r_t) = \mu, \qquad Var(r_t) = \sigma_a^2 \sum_{i=0}^{\infty} \psi_i^2$$

where σ_a^2 is the variance of a_t. Because $Var(r_t) < \infty$, $\{\psi_i^2\}$ must be a convergent sequence, that is, $\psi_i^2 \to 0$ as $i \to \infty$. Consequently, for a stationary time series, impact of the remote shock a_{t-i} on the return r_t vanishes as i increases.

The lag-l autocovariance of r_t is

$$\begin{aligned}
\gamma_l = Cov(r_t, r_{t-l}) &= E\left[\left(\sum_{i=0}^{\infty} \psi_i a_{t-i}\right)\left(\sum_{j=0}^{\infty} \psi_j a_{t-l-j}\right)\right] \\
&= E\left(\sum_{i,j=0}^{\infty} \psi_i \psi_j a_{t-i} a_{t-l-j}\right) \\
&= \sum_{j=0}^{\infty} \psi_{j+l}\psi_j E(a_{t-l-j}^2) = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+l}
\end{aligned}$$

Consequently, the ψ weights are related to the autocorrelations of r_t as follows:

$$\rho_l = \frac{\gamma_l}{\gamma_0} = \frac{\sum_{i=0}^{\infty} \psi_i \psi_{i+l}}{1 + \sum_{i=1}^{\infty} \psi_i^2}, \qquad l \geq 0,$$

where $\psi_0 = 1$. Linear time series models are econometric and statistical models used to describe the pattern of the ψ weights of r_t. For a weakly stationary time series, $\psi_i \to 0$ as $i \to \infty$ and, hence, ρ_l converges to zero as l increases.

For asset returns, this means that, as expected, the linear dependence of current return r_t on the remote past return r_{t-l} diminishes for large l.

1.5 Section 2.4 Simple AR Models

A simple model that makes use of predictive power is

$$r_t = \phi_0 + \phi_1 r_{t-1} + a_t$$

where $\{a_t\}$ is assumed to be a white noise series with mean zero and variance σ_a^2. This model is in the same form as the well-known simple linear regression model in which r_t is the dependent variable and r_{t-1} is the explanatory variable. In the time series literature, this model is referred to as an autoregressive (AR) model of order 1 or simply an AR(1) model. This model is also used in stochastic volatility modeling when r_t is replaced by its log volatility.

The AR(1) model has several properties similar to those of the simple linear regression model. But there are also some significant differences between the two models, which we'll discuss later. For now, we note that an AR(1) model implies that, conditional on the past return r_{t-1}, we have

$$E(r_t | r_{t-1}) = \phi_0 + \phi_1 r_{t-1}, \qquad Var(r_t | r_{t-1}) = Var(a_t) = \sigma_a^2$$

That is, given the past return r_{t-1}, the current return is centered around $\phi_0 + \phi_1 r_{t-1}$ with standard deviation σ_a.