

Lecture 1: Stationary Time Series*

1 Introduction

If a random variable X is indexed to time, usually denoted by t , the observations $\{X_t, t \in \mathbb{T}\}$ is called a time series, where \mathbb{T} is a time index set (for example, $\mathbb{T} = \mathbb{Z}$, the integer set).

Time series data are very common in empirical economic studies. Figure 1 plots some frequently used variables. The upper left figure plots the quarterly GDP from 1947 to 2001; the upper right figure plots the the residuals after linear-detrending the logarithm of GDP; the lower left figure plots the monthly S&P 500 index data from 1990 to 2001; and the lower right figure plots the log difference of the monthly S&P. As you could see, these four series display quite different patterns over time. Investigating and modeling these different patterns is an important part of this course.

In this course, you will find that many of the techniques (estimation methods, inference procedures, etc) you have learned in your general econometrics course are still applicable in time series analysis. However, there are something special of time series data compared to cross sectional data. For example, when working with cross-sectional data, it usually makes sense to assume that the observations are independent from each other, however, time series data are very likely to display some degree of dependence over time. More importantly, for time series data, we could observe only one history of the realizations of this variable. For example, suppose you obtain a series of US weekly stock index data for the last 50 years. This sample can be said to be large in terms of sample size, however, it is still one data point, as it is only one of the many possible realizations.

2 Autocovariance Functions

In modeling finite number of random variables, a covariance matrix is usually computed to summarize the dependence between these variables. For a time series $\{X_t\}_{t=-\infty}^{\infty}$, we need to model the dependence over infinite number of random variables. The autocovariance and autocorrelation functions provide us a tool for this purpose.

Definition 1 (*Autocovariance function*). *The autocovariance function of a time series $\{X_t\}$ with $Var(X_t) < \infty$ is defined by*

$$\gamma_X(s, t) = Cov(X_s, X_t) = E[(X_s - EX_s)(X_t - EX_t)].$$

Example 1 (Moving average process) Let $\epsilon_t \sim i.i.d.(0, 1)$, and

$$X_t = \epsilon_t + 0.5\epsilon_{t-1}$$

*Copyright 2002-2006 by Ling Hu.

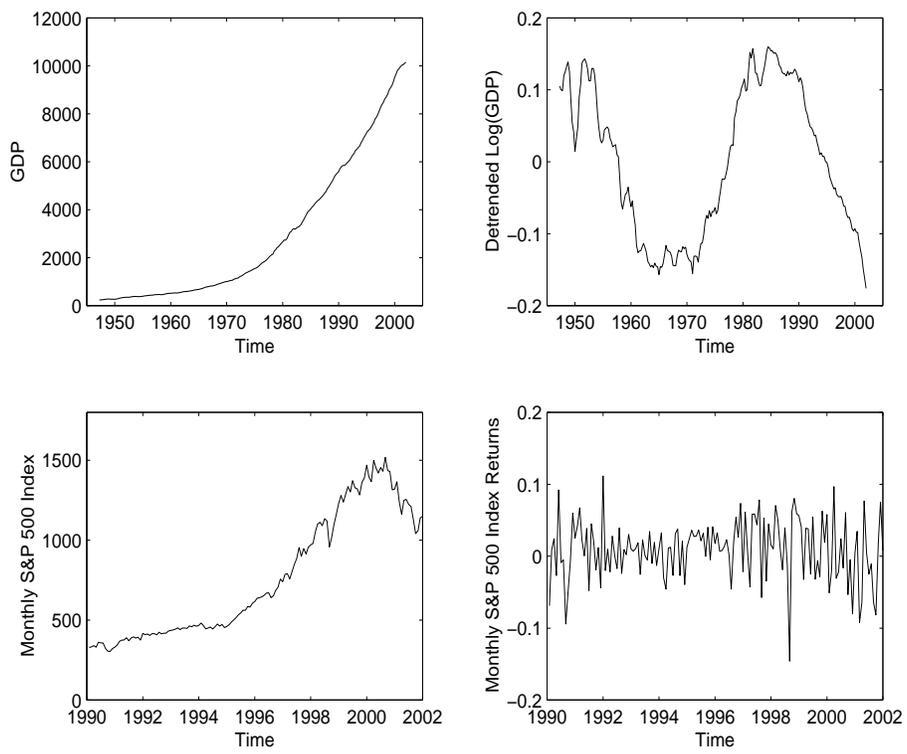


Figure 1: Plots of some economic variables

then $E(X_t) = 0$ and $\gamma_X(s, t) = E(X_s X_t)$. Let $s \leq t$. When $s = t$,

$$\gamma_X(t, t) = E(X_t^2) = 1.25,$$

when $t = s + 1$,

$$\gamma_X(t, t + 1) = E[(\epsilon_t + 0.5\epsilon_{t-1})(\epsilon_{t+1} + 0.5\epsilon_t)] = 0.5,$$

when $t - s > 1$, $\gamma_X(s, t) = 0$.

3 Stationarity and Strict Stationarity

With autocovariance functions, we can define the covariance stationarity, or weak stationarity. In the literature, usually stationarity means weak stationarity, unless otherwise specified.

Definition 2 (*Stationarity or weak stationarity*) The time series $\{X_t, t \in \mathbb{Z}\}$ (where \mathbb{Z} is the integer set) is said to be stationary if

(I) $E(X_t^2) < \infty \forall t \in \mathbb{Z}$.

(II) $EX_t = \mu \forall t \in \mathbb{Z}$.

(III) $\gamma_X(s, t) = \gamma_X(s + h, t + h) \forall s, t, h \in \mathbb{Z}$.

In other words, a stationary time series $\{X_t\}$ must have three features: finite variation, constant first moment, and that the second moment $\gamma_X(s, t)$ only depends on $(t - s)$ and not depends on s or t . In light of the last point, we can rewrite the autocovariance function of a stationary process as

$$\gamma_X(h) = Cov(X_t, X_{t+h}) \quad \text{for } t, h \in \mathbb{Z}.$$

Also, when X_t is stationary, we must have

$$\gamma_X(h) = \gamma_X(-h).$$

When $h = 0$, $\gamma_X(0) = Cov(X_t, X_t)$ is the variance of X_t , so the autocorrelation function for a stationary time series $\{X_t\}$ is defined to be

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)}.$$

Example 1 (continued): In example 1, we see that $E(X_t) = 0$, $E(X_t^2) = 1.25$, and the autocovariance functions does not depend on s or t . Actually we have $\gamma_X(0) = 1.25$, $\gamma_X(1) = 0.5$, and $\gamma_X(h) = 0$ for $h > 1$. Therefore, $\{X_t\}$ is a stationary process.

Example 2 (Random walk) Let S_t be a random walk $S_t = \sum_{s=0}^t X_s$ with $S_0 = 0$ and X_t is independent and identically distributed with mean zero and variance σ^2 . Then for $h > 0$,

$$\begin{aligned} \gamma_S(t, t+h) &= Cov(S_t, S_{t+h}) \\ &= Cov\left(\sum_{i=1}^t X_i, \sum_{j=1}^{t+h} X_j\right) \\ &= Var\left(\sum_{i=1}^t X_i\right) \quad \text{since } Cov(X_i, X_j) = 0 \quad \text{for } i \neq j \\ &= t\sigma^2 \end{aligned}$$

In this case, the autocovariance function depends on time t , therefore the random walk process S_t is not stationary.

Example 3 (Process with linear trend): Let $\epsilon_t \sim iid(0, \sigma^2)$ and

$$X_t = \delta t + \epsilon_t.$$

Then $E(X_t) = \delta t$, which depends on t , therefore a process with linear trend is not stationary.

Among stationary processes, there is simple type of process that is widely used in constructing more complicated processes.

Example 4 (White noise): The time series ϵ_t is said to be a white noise with mean zero and variance σ_ϵ^2 , written as

$$\epsilon \sim WN(0, \sigma_\epsilon^2)$$

if and only if ϵ_t has zero mean and covariance function as

$$\gamma_\epsilon(h) = \begin{cases} \sigma_\epsilon^2 & \text{if } h = 0 \\ 0 & \text{if } h \neq 0 \end{cases}$$

It is clear that a white noise process is stationary. Note that white noise assumption is weaker than identically independent distributed assumption.

To tell if a process is covariance stationary, we compute the unconditional first two moments, therefore, processes with conditional heteroskedasticity may still be stationary.

Example 5 (ARCH model) Let $X_t = \epsilon_t$ with $E(\epsilon_t) = 0$, $E(\epsilon_t^2) = \sigma^2 > 0$, and $E(\epsilon_t \epsilon_s) = 0$ for $t \neq s$. Assume the following process for ϵ_t^2 ,

$$\epsilon_t^2 = c + \rho \epsilon_{t-1}^2 + u_t$$

where $0 < \rho < 1$ and $u_t \sim WN(0, 1)$.

In this example, the conditional variance of X_t is time varying, as

$$E_{t-1}(X_t^2) = E_{t-1}(\epsilon_t^2) = E_{t-1}(c + \rho \epsilon_{t-1}^2 + u_t) = c + \rho \epsilon_{t-1}^2.$$

However, the unconditional variance of X_t is constant, which is $\sigma^2 = c/(1 - \rho)$. Therefore, this process is still stationary.

Definition 3 (*Strict stationarity*) The time series $\{X_t, t \in \mathbb{Z}\}$ is said to be strict stationary if the joint distribution of $(X_{t_1}, X_{t_2}, \dots, X_{t_k})$ is the same as that of $(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_k+h})$.

In other words, strict stationarity means that the joint distribution only depends on the ‘difference’ h , not the time (t_1, \dots, t_k) .

Remarks: First note that finite variance is not assumed in the definition of strong stationarity, therefore, strict stationarity does not necessarily imply weak stationarity. For example, processes like i.i.d. Cauchy is strictly stationary but not weak stationary. Second, a nonlinear function of a strict stationary variable is still strictly stationary, but this is not true for weak stationary. For example, the square of a covariance stationary process may not have finite variance. Finally, weak

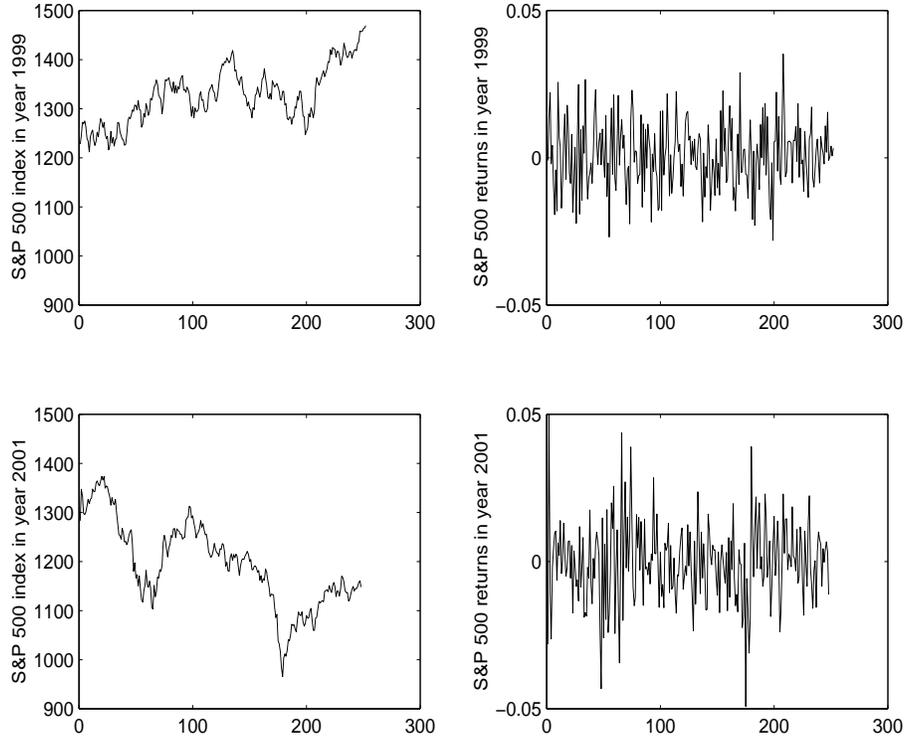


Figure 2: Plots of S&P index and returns in year 1999 and 2001

stationarity usually does not imply strict stationarity as higher moments of the process may depend on time t . However, if process $\{X_t\}$ is a Gaussian time series, which means that the distribution functions of $\{X_t\}$ are all multivariate Gaussian, i.e. the joint density of

$$f_{X_t, X_{t+j_1}, \dots, X_{t+j_k}}(x_t, x_{t+j_1}, \dots, x_{t+j_k})$$

is Gaussian for any j_1, j_2, \dots, j_k , weak stationarity also implies strict stationarity. This is because a multivariate Gaussian distribution is fully characterized by its first two moments.

For example, a white noise is stationary but may not be strict stationary, but a Gaussian white noise is strict stationary. Also, general white noise only implies uncorrelation while Gaussian white noise also implies independence. Because if a process is Gaussian, uncorrelation implies independence. Therefore, a Gaussian white noise is just $i.i.d.N(0, \sigma^2)$.

Stationary and nonstationary processes are very different in their properties, and they require different inference procedures. We will discuss this in much details through this course. At this point, note that a simple and useful method to tell if a process is stationary in empirical studies is to plot the data. Loosely speaking, if a series does not seem to have a constant mean or variance, then very likely, it is not stationary. For example, Figure 2 plots the daily S&P 500 index in year 1999 and 2001. The upper left figure plots the index in 1999, upper right figure plots the returns in 1999, lower left figure plots the index in 2001, and lower right figure plots the returns in 2001.

Note that the index level are very different in 1999 and 2001. In year 1999, it is wandering at a higher level and the market rises. In year 2001, the level is much lower and the market drops.

In comparison, we did not see much difference in the returns in year 1999 and 2001 (although the returns in 2001 seem to have thicker tails). Actually, only judging from the return data, it is very hard to tell which figure plots the market in booms, and which figure plots the market in crashes. Therefore, people usually treat stock price data as nonstationary and stock return data as stationary.

4 Ergodicity

Recall that Kolmogorov's law of large number (LLN) tells that if $X_i \sim i.i.d.(\mu, \sigma^2)$ for $i = 1, \dots, n$, then we have the following limit for the ensemble average

$$\bar{X}_n = n^{-1} \sum_{i=1}^n X_i \rightarrow \mu.$$

In time series, we have time series average, not ensemble average. To explain the differences between ensemble average and time series average, consider the following experiment. Suppose we want to track the movements of some particles and draw inference about their expected position (suppose that these particles move on the real line). If we have a group of particles (group size n), then we could track down the position of each particle and plot a distribution of their positions. The mean of this sample is called ensemble average. If all these particles are i.i.d., LLN tells that this average converges to its expectation as $n \rightarrow \infty$. However, as we remarked earlier, with time series observations, we only have one history. That means, in this experiment, we only have one particle. Then instead of collecting n particles, we can only track this single particle and record its position, say x_t , for $t = 1, 2, \dots, T$. The mean we computed by averaging over time, $T^{-1} \sum_{t=1}^T x_t$ is called time series average.

Does the time series average converges to the same limit as the ensemble average? The answer is yes if X_t is stationary and ergodic. If X_t is stationary and ergodic with $E(X_t) = \mu$, then the time series average has the same limit as ensemble average,

$$\bar{X}_T = T^{-1} \sum_{t=1}^T X_t \rightarrow \mu.$$

This result is given as ergodic theorem, and we will discuss it later in our lecture 4 on asymptotic theory. Note that this result require both stationarity and ergodicity. We have explained stationarity and we see that stationarity allows time series dependence. Ergodicity requires 'average asymptotic independence'. Note that stationarity itself does not guarantee ergodicity (page 47 in Hamilton and lecture 4).

Readings:

Hamilton, Ch. 3.1

Brockwell and Davis, Page 1-29

Hayashi, Page 97-102