

Statistics 153 (Time Series) : Lecture Five

Aditya Guntuboyina

31 January 2012

1 Contents

1. Definitions of strong and weak stationary sequences.
2. Examples of stationary sequences.

2 Stationarity

The concept of a stationary time series is the most important thing in this course.

Stationary time series are typically used for the residuals after trend and seasonality have been removed.

Stationarity allows a systematic study of time series forecasting.

In order to avoid confusion, we shall, from now on, make a distinction between the observed time series data $x_1, \dots, x_n$ and the sequence of random variables $X_1, \dots, X_n$. The observed data $x_1, \dots, x_n$ is a realization of $X_1, \dots, X_n$. It is convenient to think of the random variables $\{X_t\}$ as forming a doubly infinite sequence:

$$\dots, X_{-2}, X_{-1}, X_0, X_1, X_2, \dots$$

The notion of stationarity will apply to the doubly infinite sequence of random variables $\{X_t\}$. Strictly speaking, stationarity does not apply to the data $x_1, \dots, x_n$. However, one frequently abuses terminology and uses stationary data to mean that the random variables of which the observed data are a realization have a stationary distribution.

Definition 2.1 (Strict or Strong Stationarity). A doubly infinite sequence of random variables $\{X_t\}$ is **strictly stationary** if the joint distribution of $(X_{t_1}, X_{t_2}, \dots, X_{t_k})$ is the same as the joint distribution of $(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_k+h})$ for every choice of times $t_1, \dots, t_k$ and h .

Roughly speaking, stationarity means that the joint distribution of the random variables remains constant over time. For example, under stationarity, the joint distribution of today's and tomorrow's random variables is the same as the joint distribution of the variables from *any* two successive days (past or future).

Note how stationarity makes the problem of forecasting or prediction feasible. From the data, we can study how a particular day's observation depends on the those of the previous days and because under stationarity, such a dependence is assumed to be constant over time, one can hope to use it to predict future observations from the current data.

Many of the things that we shall be doing with stationarity actually go through even with the following notion that is weaker than strong stationarity.

Definition 2.2 (Second-Order or Weakly or Wide-Sense Stationarity). A doubly infinite sequence of random variables $\{X_t\}$ is **second-order stationary** if

1. The mean of the random variable X_t , denoted by $\mathbb{E}X_t$, is the same for all times t .
2. The covariance between X_{t_1} and X_{t_2} is the same as the covariance between X_{t_1+h} and X_{t_2+h} for every choice of times t_1, t_2 and h .

Second stationarity means that the *second order properties* (means and covariances) of the random variables remain constant over time. It should be noted that, unlike strong stationarity, the joint distribution of the random variables may well change over time.

Another way of phrasing the condition

$$\text{cov}(X_{t_1}, X_{t_2}) = \text{cov}(X_{t_1+h}, X_{t_2+h}) \quad \text{for all } t_1, t_2 \text{ and } h$$

is that the covariance between two random variables only depends on the *time lag* between them. In other words, the covariance between X_{t_1} and X_{t_2} only depends on the time lag $|t_1 - t_2|$ between them. Thus, if we define $\gamma(h)$ to be the covariance between the random variables corresponding to any two times with a lag of h , we have

$$\text{cov}(X_{t_1}, X_{t_2}) = \gamma(|t_1 - t_2|) \quad \text{for all times } t_1 \text{ and } t_2$$

as a consequence of second-order stationarity.

The function $\gamma(h)$ is called the **Autocovariance Function** of the stationary sequence $\{X_t\}$. The notion of autocovariance function only applies to a stationary sequence of random variables but not to data. The Autocovariance function is abbreviated to *acvf*.

The variance of each random variable X_t is given by $\gamma(0)$. By stationarity, all random variables have the same variance.

Let $\rho(h)$ denote the correlation between two random variables in the stationary sequence $\{X_t\}$ that are separated by a time lag of h . Because $\gamma(h)$ denotes the corresponding covariance, it follows that

$$\rho(h) := \frac{\gamma(h)}{\gamma(0)}.$$

The function $\rho(h)$ is called the **Autocorrelation Function** of the stationary sequence $\{X_t\}$. Once again, this is a notion that only applies to random variables but not to data. On the contrary, the Sample Autocorrelation Function that we looked at before only applies to data. The Autocorrelation function is abbreviated to *acf*.

Because $\rho(h)$ is a correlation, it follows that $|\rho(h)| \leq 1$. Also $\rho(0)$ equals 1.

Definition 2.3 (Gaussian Process). *The sequence $\{X_t\}$ is said to be **gaussian** if the joint distribution of $(X_{t_1}, \dots, X_{t_k})$ is multivariate normal for every choice of times $t_1, \dots, t_k$.*

Important Note: $(X_{t_1}, \dots, X_{t_k})$ is multivariate normal means that every **linear combination** of $(X_{t_1}, \dots, X_{t_k})$ is univariate normal. In particular, it is much stronger than saying that each of $X_{t_1}, \dots, X_{t_k}$ has a univariate normal distribution. If $X_{t_1}, \dots, X_{t_k}$ are **independent** normal random variables, then their joint distribution is an example of a multivariate normal distribution.

The following statement is a consequence of the fact that multivariate normal distributions are determined means and covariances:

second-order stationarity + Gaussian Process $\implies$ Strong Stationarity.

In the rest of this course, stationarity would always stand for second-order stationarity. Unless explicitly mentioned, do not assume that a stationary series is strongly stationary.

3 Examples

3.1 White Noise Series

The random variables $\dots, X_{-2}, X_{-1}, X_0, X_1, X_2, \dots$ are said to form a *white noise* series if they have mean zero and the following covariance:

$$\begin{aligned}\text{cov}(X_{t_1}, X_{t_2}) &= \sigma^2 && \text{if } t_1 = t_2 \\ &= 0 && \text{if } t_1 \neq t_2.\end{aligned}\tag{1}$$

In other words, the random variables in a white noise series are uncorrelated, have mean zero and a constant variance.

This is clearly a stationary series. What is its acf?

The random variables $\{X_t\}$ are said to be *purely random* if they are independent and identically distributed with a finite constant mean and variance. Unless explicitly specified, do not assume that the common distribution is normal or that the common mean is zero or that the common variance is one. A purely random series $\{X_t\}$ with mean zero is strongly stationary.

The white noise series is only a very special example of stationarity. Stationarity allows for considerable dependence between successive random variables in the series. The only requirement is that the dependence should be constant over time.

3.2 Moving Average

Given a white noise series Z_t with variance σ^2 and a number θ , set

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

This is called a *moving average* of order 1. The series is stationary with mean zero and acvf:

$$\begin{aligned}\gamma_X(h) &= \sigma^2(1 + \theta^2) && \text{if } h = 0 \\ &= \theta\sigma^2 && \text{if } h = 1 \\ &= 0 && \text{otherwise.}\end{aligned}\tag{2}$$

As a consequence, X_{t_1} and X_{t_2} are uncorrelated whenever t_1 and t_2 are two or more time points apart. This time series has *short memory*.

More generally, one can consider a Moving Average of order q defined by

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \dots + \theta_q Z_{t-q}\tag{3}$$

for some numbers $\theta_1, \theta_2, \dots, \theta_q$.

This is also stationary with mean zero and acvf (below: θ_0 equals 1):

$$\begin{aligned}\gamma_X(h) &= \sigma^2 \sum_{i=0}^{q-h} \theta_i \theta_{i+h} & \text{if } h = 0, 1, \dots, q \\ &= 0 & \text{if } h > q.\end{aligned}\tag{4}$$

The acvf now *cuts off* at lag q .

3.3 Infinite Order Moving Average

We can extend the definition of moving average (5) to infinite order by:

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \dots + \theta_q Z_{t-q} + \theta_{q+1} Z_{t-q-1} + \dots \tag{5}$$

Here, once again, $\{Z_t\}$ represents a white noise sequence with mean zero and variance σ_Z^2 .

The right hand above is an infinite sum and hence we need to address convergence issues. One particular choice of the weights $\{\theta_i\}$ where one does not encounter problems with convergence is:

$$\theta_i = \phi^i \quad \text{for some } \phi \text{ with } |\phi| < 1.$$

The resulting process can be written succinctly as $X_t = \sum_{i=0}^{\infty} \phi^i Z_{t-i}$. It can be shown, using the fact that $|\phi| < 1$, that the sequence $\sum_{i=0}^n \phi^i Z_{t-i}$ converges as $n \rightarrow \infty$ to a random variable with finite variance in both the L^2 sense (mean square sense) and the almost sure sense. X_t is defined to be that limit.

For this choice of weights, is $\{X_t\}$ a stationary sequence. What is its acvf:

$$\gamma(h) := \frac{\phi^{|h|} \sigma^2}{1 - \phi^2}.$$

Here is an important property of this process X_t :

$$\begin{aligned}X_t &= Z_t + \phi Z_{t-1} + \phi^2 Z_{t-2} + \dots = Z_t + \phi (Z_{t-1} + \phi Z_{t-2} + \phi^2 Z_{t-3} + \dots) \\ &= Z_t + \phi X_{t-1} \quad \text{for every } t = \dots, -1, 0, 1, \dots\end{aligned}$$

Thus X_t satisfies the following first order *difference equation*:

$$X_t = \phi X_{t-1} + Z_t. \tag{6}$$

For this reason, $X_t = \sum_{i=0}^{\infty} \phi^i Z_{t-i}$ is called the *Stationary Autoregressive Process of order one*. Note here that $|\phi| < 1$.

This raises an interesting question. Is $X_t = \sum_{i=0}^{\infty} \phi^i Z_{t-i}$ the only solution to the difference equation (6) for $|\phi| < 1$?

The answer is no. Indeed, define X_0 to be an arbitrary random variable that is uncorrelated with the white noise series $\{Z_t\}$ and define $X_1, X_2, \dots$ as well as $X_{-1}, X_{-2}, \dots$ using the difference equation (6). The resulting sequence surely satisfies (6).

However, an **important fact** is that $X_t = \sum_{i=0}^{\infty} \phi^i Z_{t-i}$ is the only **stationary** solution to the difference equation (6) for $|\phi| < 1$. We shall prove this in the next class.