STAT 626: Outline of Lectures 5-6

Stationary Time Series (§2.2)

1. Autocovariance of a Time Series (TS),

2. Autocorrelation Function (ACF) of a Stationary TS and Correlogram:

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}, \quad h = 0, 1, \ldots.$$

3. Important Example of Stationary Time Series: Moving Average (MA), Autoregressive (AR) Models,...

4. Wold Decomposition: A good stationary TS is linear in the WN.

5. Bivariate TS and Stationarity.

Topics From Chapter 1

6. White Noise: The Building Blocks

7. Autoregression: The Birth of Modern Time Series Analysis

8. Random Walks: The Engine of Financial Engineering

9. Signal + Noise: For Other Engineering

DEFINITIONS

1. A Time Series $\{x_t\}$ is **stationary** if

   (a) the mean function $E(x_t)$ does not depend on the time t,

   (b) the covariance function $\text{cov}(x_s, x_t)$ depends on the times s, t only through the distance $|s-t|$.

2. **Autocovariance Function** of a Stationary Time Series:

$$\gamma(h) = \text{cov}(x_{t+h}, x_t), \quad h = 0, 1, \ldots.$$

   NOTE:

$$\gamma(0) = \text{cov}(x_t, x_t) = \text{var}(x_t).$$

3. **The Autocorrelation Function (ACF)**

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}, \quad h = 0, 1, \ldots,$$

   is a **symmetric** function of the lag h. **Correlogram** is the plot of $\rho(h)$ vs h.

4. **Multivariate Time Series:**

   Cross-Covariance Function (CCF) of Two Time Series:

$$\gamma_{xy}(h) = \text{Cov}(x_{t+h}, y_t), h = 0, \pm 1, \pm 2, \ldots$$

   Why ACF is symmetric and CCF is not?

   Proof without words!

$$\gamma_{xx}(h) = \text{Cov}(x_{t+h}, x_t) = \text{Cov}(x_t, x_{t+h}) = \gamma_{xx}(-h), \quad h = 1, 2, \ldots.$$

2.2 Stationarity

Although we have previously not made any special assumptions about the behavior of the time series, many of the examples we have seen hinted that a sort of regularity may exist over time in the behavior of a time series. Stationarity requires regularity in the mean and autocorrelation functions so that these quantities (at least) may be estimated by averaging.

Definition 2.13. *A* **stationary** *time series is a finite variance process where*

(i) the mean value function, μ_t, defined in (2.1) is constant and does not depend on time t, and

(ii) the autocovariance function, $\gamma(s,t)$, defined in (2.2) depends on times s and t only through their time difference.

As an example, for a stationary hourly time series, the correlation between what happens at 1AM and 3AM is the same as between what happens at 9PM and 11PM because they are both two hours apart.

Example 2.14. A Random Walk is Not Stationary
A random walk is not stationary because its autocovariance function, $\gamma(s,t) = \min\{s,t\}\sigma_w^2$, depends on time; see Example 2.9 and Problem 2.5. Also, the random walk with drift violates both conditions of Definition 2.13 because the mean function, $\mu_{xt} = \delta t$, depends on time t as shown in Example 2.3

Because the mean function, $E(x_t) = \mu_t$, of a stationary time series is independent of time t, we will write

$$\mu_t = \mu. \tag{2.10}$$

Also, because the autocovariance function, $\gamma(s,t)$, of a stationary time series, x_t, depends on s and t only through time difference, we may simplify the notation. Let $s = t + h$, where h represents the time shift or *lag*. Then

$$\gamma(t+h,t) = \text{cov}(x_{t+h}, x_t) = \text{cov}(x_h, x_0) = \gamma(h,0)$$

because the time difference between $t+h$ and t is the same as the time difference between h and 0. Thus, the autocovariance function of a stationary time series does not depend on the time argument t. Henceforth, for convenience, we will drop the second argument of $\gamma(h,0)$.

Definition 2.15. *The* **autocovariance function of a stationary time series** *will be written as*

$$\gamma(h) = \text{cov}(x_{t+h}, x_t) = E[(x_{t+h} - \mu)(x_t - \mu)]. \tag{2.11}$$

Definition 2.16. *The* **autocorrelation function (ACF) of a stationary time series** *will be written using (2.7) as*

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}. \tag{2.12}$$

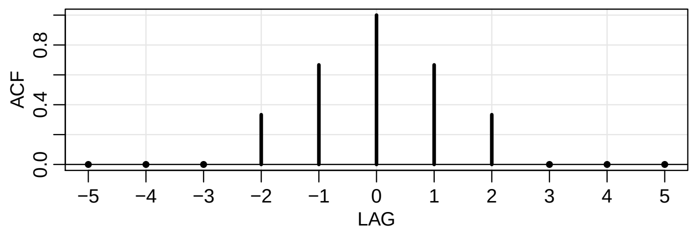

Figure 2.1: *Autocorrelation function of a three-point moving average.*

Because it is a correlation, we have $-1 \leq \rho(h) \leq 1$ for all h, enabling one to assess the relative importance of a given autocorrelation value by comparing with the extreme values -1 and 1.

Example 2.17. Stationarity of White Noise

The mean and autocovariance functions of the white noise series discussed in Example 1.7 and Example 2.6 are easily evaluated as $\mu_{wt} = 0$ and

$$\gamma_w(h) = \text{cov}(w_{t+h}, w_t) = \begin{cases} \sigma_w^2 & h = 0, \\ 0 & h \neq 0. \end{cases}$$

Thus, white noise satistfies Definition 2.13 and is stationary. $\lozenge$

Example 2.18. Stationarity of a Moving Average

The three-point moving average process of Example 1.8 is stationary because, from Example 2.2 and Example 2.8, the mean and autocovariance functions $\mu_{vt} = 0$, and

$$\gamma_v(h) = \begin{cases} \frac{3}{9}\sigma_w^2 & h = 0, \\ \frac{2}{9}\sigma_w^2 & h = \pm 1, \\ \frac{1}{9}\sigma_w^2 & h = \pm 2, \\ 0 & |h| > 2 \end{cases}$$

are independent of time t, satisfying the conditions of Definition 2.13. Note that the ACF, $\rho(h) = \gamma(h)/\gamma(0)$, is given by

$$\rho_v(h) = \begin{cases} 1 & h = 0, \\ 2/3 & h = \pm 1, \\ 1/3 & h = \pm 2, \\ 0 & |h| > 2 \end{cases}.$$

Figure 2.1 shows a plot of the autocorrelation as a function of lag h. Note that the autocorrelation function is symmetric about lag zero.

```
ACF = c(0,0,0,1,2,3,2,1,0,0,0)/3
LAG = -5:5
tsplot(LAG, ACF, type="h", lwd=3, xlab="LAG")
```

```
abline(h=0)
points(LAG[-(4:8)], ACF[-(4:8)], pch=20)
axis(1, at=seq(-5, 5, by=2))
```

Example 2.19. Trend Stationarity
A time series can have stationary behavior around a trend. For example, if

$$x_t = \beta t + y_t \,,$$

where y_t is stationary with mean and autocovariance functions μ_y and $\gamma_y(h)$, respectively. Then the mean function of x_t is

$$\mu_{x,t} = E(x_t) = \beta t + \mu_y \,,$$

which is not independent of time. Therefore, the process is not stationary. The autocovariance function, however, is independent of time, because

$$\begin{aligned}\gamma_x(h) = \text{cov}(x_{t+h}, x_t) &= E[(x_{t+h} - \mu_{x,t+h})(x_t - \mu_{x,t})] \\ &= E[(y_{t+h} - \mu_y)(y_t - \mu_y)] = \gamma_y(h).\end{aligned}$$

This behavior is sometimes called *trend stationarity*. An example of such a process is the export price of salmon series displayed in Figure 3.1. ◊

The autocovariance function of a stationary process has several useful properties. First, the value at $h = 0$ is the variance of the series,

$$\gamma(0) = E[(x_t - \mu)^2] = \text{var}(x_t)\,. \tag{2.13}$$

Another useful property is that the autocovariance function of a stationary series is symmetric around the origin,

$$\gamma(h) = \gamma(-h) \tag{2.14}$$

for all h. This property follows because

$$\begin{aligned}\gamma(h) &= \gamma((t+h) - t) = E[(x_{t+h} - \mu)(x_t - \mu)] \\ &= E[(x_t - \mu)(x_{t+h} - \mu)] = \gamma(t - (t+h)) = \gamma(-h),\end{aligned}$$

which shows how to use the notation as well as proving the result.

Example 2.20. Autoregressive Models
The stationarity of AR models is a little more complex and is dealt with in Chapter 4. We'll use an AR(1) to examine some aspects of the model,

$$x_t = \phi x_{t-1} + w_t.$$

Since the mean must be constant, if x_t is stationary the mean function $\mu_t = E(x_t) = \mu$ is constant so

$$E(x_t) = \phi E(x_{t-1}) + E(w_t)$$

implies $\mu = \phi\mu + 0$; thus $\mu = 0$. In addition, assuming x_{t-1} and w_t are uncorrelated,

$$\begin{aligned} \text{var}(x_t) &= \text{var}(\phi x_{t-1} + w_t) \\ &= \text{var}(\phi x_{t-1}) + \text{var}(w_t) + 2\text{cov}(\phi x_{t-1}, w_t) \\ &= \phi^2 \text{var}(x_{t-1}) + \text{var}(w_t)\,. \end{aligned}$$

If x_t is stationary, the variance, $\text{var}(x_t) = \gamma_x(0)$, is constant, so

$$\gamma_x(0) = \phi^2 \gamma_x(0) + \sigma_w^2\,.$$

Thus

$$\gamma_x(0) = \sigma_w^2 \frac{1}{(1-\phi^2)}\,.$$

Note that for the process to have a positive, finite variance, we should require $|\phi| < 1$. Similarly,

$$\begin{aligned} \gamma_x(1) &= \text{cov}(x_t, x_{t-1}) = \text{cov}(\phi x_{t-1} + w_t, x_{t-1}) \\ &= \text{cov}(\phi x_{t-1}, x_{t-1}) = \phi \gamma_x(0). \end{aligned}$$

Thus,

$$\rho_x(1) = \frac{\gamma_x(1)}{\gamma_x(0)} = \phi\,,$$

and we see that ϕ is in fact a correlation, $\phi = \text{corr}(x_t, x_{t-1})$.

It should be evident that we have to be careful when working with AR models. It should also be evident that, as in Example 1.9, simply setting the initial conditions equal to zero does not meet the stationary criteria because x_0 is not a constant, but a random variable with mean μ and variance $\sigma_w^2/(1-\phi^2)$. ◊

In Section 1.3, we discussed the notion that it is possible to generate realistic time series models by filtering white noise. In fact, there is a result by Wold (1954) that states that any (non-deterministic[1]) stationary time series is in fact a filter of white noise.

Property 2.21 (Wold Decomposition). *Any stationary time series, x_t, can be written as linear combination (filter) of white noise terms; that is,*

$$x_t = \mu + \sum_{j=0}^{\infty} \psi_j w_{t-j}, \tag{2.15}$$

where the ψs are numbers satisfying $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ and $\psi_0 = 1$. We call these **linear processes**.

[1]This means that no part of the series is deterministic, meaning one where the future is perfectly predictable from the past; e.g., model (1.6).

Remark. Property 2.21 is important in the following ways:

- As previously suggested, stationary time series can be thought of as filters of white noise. It may not always be the best model, but models of this form are viable in many situations.
- Any stationary time series can be represented as a model that does not depend on the future. That is, x_t in (2.15) depends only on the present w_t and the past $w_{t-1}, w_{t-2}, \ldots$.
- Because the coefficients satisfy $\psi_j^2 \to 0$ as $j \to \infty$, the dependence on the distant past is negligible. Many of the models we will encounter satisfy the much stronger condition $\sum_{j=0}^{\infty} |\psi_j| < \infty$.

The models we will encounter in Chapter 4 are linear processes. For the linear process, we may show that the mean function is $E(x_t) = \mu$, and the autocovariance function is given by

$$\gamma(h) = \sigma_w^2 \sum_{j=0}^{\infty} \psi_{j+h}\psi_j \tag{2.16}$$

for $h \geq 0$; recall that $\gamma(-h) = \gamma(h)$. To see (2.16), note that

$$\begin{aligned}
\gamma(h) = \operatorname{cov}(x_{t+h}, x_t) &= \operatorname{cov}\Big(\sum_{j=0}^{\infty} \psi_j w_{t+h-j},\ \sum_{k=0}^{\infty} \psi_k w_{t-k}\Big) \\
&= \operatorname{cov}\big[w_{t+h} + \cdots + \psi_h w_t + \psi_{h+1} w_{t-1} + \cdots,\ \psi_0 w_t + \psi_1 w_{t-1} + \cdots\big] \\
&= \sigma_w^2 \sum_{j=0}^{\infty} \psi_{h+j}\psi_j\,.
\end{aligned}$$

The moving average model is already in the form of a linear process. The autoregressive model such as the one in Example 1.9 can also be put in this form as we suggested in that example.

When several series are available, a notion of stationarity still applies with additional conditions.

Definition 2.22. *Two time series, say, x_t and y_t, are* **jointly stationary** *if they are each stationary, and the* cross-covariance function

$$\gamma_{xy}(h) = \operatorname{cov}(x_{t+h}, y_t) = E[(x_{t+h} - \mu_x)(y_t - \mu_y)] \tag{2.17}$$

is a function only of lag h.

Definition 2.23. *The* **cross-correlation function (CCF)** *of jointly stationary time series x_t and y_t is defined as*

$$\rho_{xy}(h) = \frac{\gamma_{xy}(h)}{\sqrt{\gamma_x(0)\gamma_y(0)}}. \tag{2.18}$$

As usual, we have the result $-1 \leq \rho_{xy}(h) \leq 1$ which enables comparison with the extreme values -1 and 1 when looking at the relation between x_{t+h} and y_t. The cross-correlation function is *not* generally symmetric about zero because when $h > 0$, y_t happens before x_{t+h} whereas when $h < 0$, y_t happens after x_{t+h}.

Example 2.24. Joint Stationarity
Consider the two series, x_t and y_t, formed from the sum and difference of two successive values of a white noise process, say,

$$x_t = w_t + w_{t-1} \quad \text{and} \quad y_t = w_t - w_{t-1},$$

where w_t is white noise with variance σ_w^2. It is easy to show that $\gamma_x(0) = \gamma_y(0) = 2\sigma_w^2$ because the w_ts are uncorrelated. In addition,

$$\gamma_x(1) = \text{cov}(x_{t+1},\, x_t) = \text{cov}(w_{t+1} + w_t,\, w_t + w_{t-1}) = \sigma_w^2$$

and $\gamma_x(-1) = \gamma_x(1)$; similarly $\gamma_y(1) = \gamma_y(-1) = -\sigma_w^2$. Also,

$$\begin{aligned}
\gamma_{xy}(0) &= \text{cov}(x_t, y_t) = \text{cov}(w_{t+1} + w_t, w_{t+1} - w_t) = \sigma_w^2 - \sigma_w^2 = 0\,;\\
\gamma_{xy}(1) &= \text{cov}(x_{t+1}, y_t) = \text{cov}(w_{t+1} + w_t, w_t - w_{t-1}) = \sigma_w^2\,;\\
\gamma_{xy}(-1) &= \text{cov}(x_{t-1}, y_t) = \text{cov}(w_{t-1} + w_{t-2}, w_t - w_{t-1}) = -\sigma_w^2\,.
\end{aligned}$$

Noting that $\text{cov}(x_{t+h}, y_t) = 0$ for $|h| > 2$, using (2.18) we have,

$$\rho_{xy}(h) = \begin{cases} 0 & h = 0, \\ \frac{1}{2} & h = 1, \\ -\frac{1}{2} & h = -1, \\ 0 & |h| \geq 2. \end{cases}$$

Clearly, the autocovariance and cross-covariance functions depend only on the lag separation, h, so the series are jointly stationary. ◊

Example 2.25. Prediction via Cross-Correlation
Consider the problem of determining leading or lagging relations between two stationary series x_t and y_t. If for some unknown integer ℓ, the model

$$y_t = Ax_{t-\ell} + w_t$$

holds, the series x_t is said to **lead** y_t for $\ell > 0$ and is said to **lag** y_t for $\ell < 0$. Estimating the lead or lag relations might be important in predicting the value of y_t from x_t. Assuming that the noise w_t is uncorrelated with the x_t series, the cross-covariance function can be computed as

$$\begin{aligned}
\gamma_{yx}(h) &= \text{cov}(y_{t+h}, x_t) = \text{cov}(Ax_{t+h-\ell} + w_{t+h},\, x_t)\\
&= \text{cov}(Ax_{t+h-\ell},\, x_t) = A\gamma_x(h-\ell)\,.
\end{aligned}$$

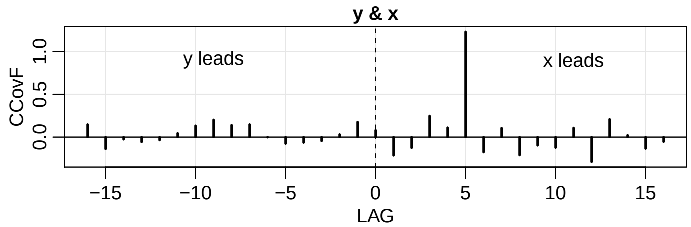

Figure 2.2 *Demonstration of the results of Example 2.25 when $\ell = 5$. The title indicates which series is leading.*

Since the largest value of $|\gamma_x(h-\ell)|$ is $\gamma_x(0)$, i.e., when $h = \ell$, the cross-covariance function will look like the autocovariance of the input series x_t, and it will have an extremum on the positive side if x_t leads y_t and an extremum on the negative side if x_t lags y_t. Below is the R code of an example with a delay of $\ell = 5$ and $\hat{\gamma}_{yx}(h)$, which is defined in Definition 2.30, shown in Figure 2.2.

```
x = rnorm(100)
y = lag(x,-5) + rnorm(100)
ccf(y, x, ylab="CCovF", type="covariance", panel.first=grid())
```

◊

2.3 Estimation of Correlation

For data analysis, only the sample values, $x_1, x_2, \ldots, x_n$, are available for estimating the mean, autocovariance, and autocorrelation functions. *In this case, the assumption of stationarity becomes critical and allows the use of averaging to estimate the population mean and covariance functions.*

Accordingly, if a time series is stationary, the mean function (2.10) $\mu_t = \mu$ is constant so we can estimate it by the *sample mean*,

$$\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t. \tag{2.19}$$

The estimate is unbiased, $E(\bar{x}) = \mu$, and its standard error is the square root of $\mathrm{var}(\bar{x})$, which can be computed using first principles (Property 2.7), and is given by

$$\mathrm{var}(\bar{x}) = \frac{1}{n}\sum_{h=-n}^{n}\Big(1 - \frac{|h|}{n}\Big)\gamma_x(h)\,. \tag{2.20}$$

If the process is white noise, (2.20) reduces to the familiar σ_x^2/n recalling that $\gamma_x(0) = \sigma_x^2$. Note that in the case of dependence, the standard error of $\bar{x}$ may be smaller or larger than the white noise case depending on the nature of the correlation structure (see Problem 2.10).

The theoretical autocorrelation function, (2.12), is estimated by the sample ACF as follows.