Unit 5: Exploratory Data Analysis I

Taylor R. Brown PhD

Department of Statistics, University of Virginia

Spring 2020

Readings for Unit 5

Textbook chapter 2.2 (pages 54 to 62).

Last Unit

1. Linear Regression Overview

This Unit

1. Detrending with Linear Regression
2. Differencing for Stationarity
3. Backshift Operator

Motivation

With time series data, we need to account for the dependence between the values in the series. Stationarity for a time series enables us to measure the dependence, since the dependence structure is regular and does not change over time. When we do not have stationarity, we need methods to reduce the effects of nonstationarity so that stationary properties can be obtained.

Detrending with Linear Regression

Recall that in time series analysis, we prefer to work with stationary processes, to better estimate autocorrelation and other quantities of interest. If our process has a linear trend, we could use linear regression to remove the trend (**"detrend"**).

Detrending with Linear Regression

Consider the following model:

$$x_t = \mu_t + y_t \tag{1}$$

where y_t is a **zero mean stationary process**, e.g. MA(2), AR(1), white noise, etc., and μ_t is a **deterministic trend**, e.g. $\mu_t = \beta_0 + \beta_1 t$.

We can view x_t as having stationary behavior around a trend. A strong trend, μ_t, can obscure the behavior of the stationary process, y_t. Hence, we may want to remove the trend as a first step in exploratory analysis of such time series, especially if the goal is to understand the behavior of y_t.

Detrending with Linear Regression

The steps involved are

- Obtain an **estimate** of the trend component, $\hat{\mu}_t$, e.g. via OLS.
- Work with the residuals $e_t = x_t - \hat{\mu}_t$.

Back to Marriages in Church of England

Recall in this earlier example that a regression of marriages in the Church of England against time should include a quadratic term, so the trend is

$$\mu_t = \beta_1 + \beta_2 t + \beta_3 t^2.$$

Using OLS regression, we obtained

$$\hat{\mu}_t = -15130 + 16.43t - 0.00444t^2.$$

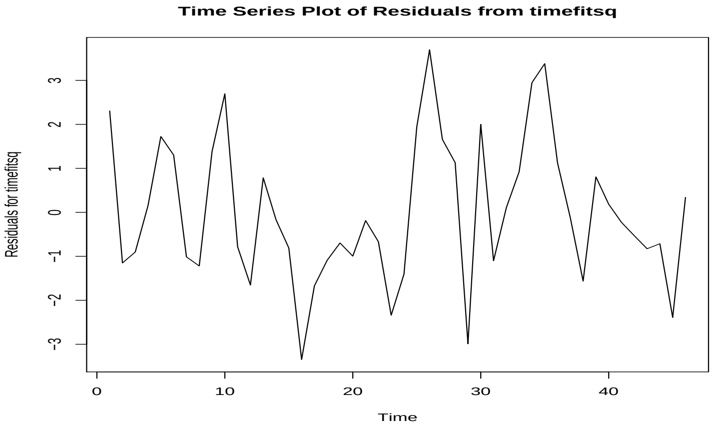
Back to Marriages in Church of England

The detrended series is obtained by subtracting $\hat{\mu}_t$ from the observed values x_t

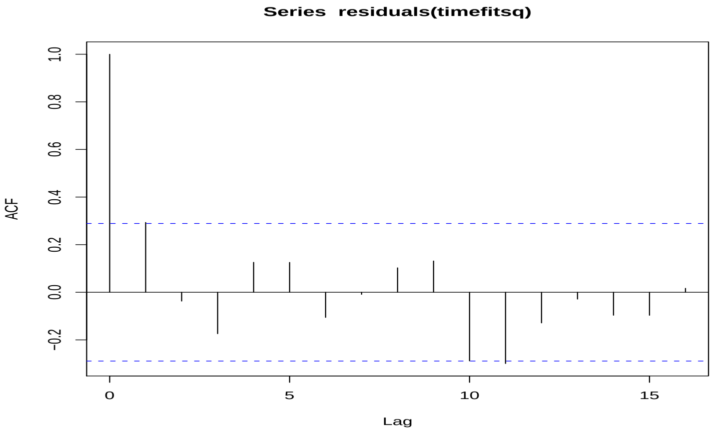
$$e_t = x_t + 15130 - 16.43t + 0.00444t^2.$$

Next, we observe the plot of e_t and its ACF.

Back to Marriages in Church of England



Back to Marriages in Church of England



Summary of Detrending

We find a good point estimate of μ_t and then look at the residuals

$$e_t = x_t - \hat{\mu}_t.$$

One thing we could do is to "detrend" with linear regression, e.g. use OLS regression to obtain $\hat{\mu}_t = \hat{\beta}_1 + \hat{\beta}_2 t$. With the residuals, based on the structure of their ACF, we have an idea on the model that may reasonably describe the stationary process y_t.

Summary of Detrending

Question: What is an implication if the residuals after detrending have a pattern / trend?

Differencing

Let ∇ denote the differencing operation. Applying ∇ to x_t results in

$$\nabla x_t = x_t - x_{t-1}, \tag{2}$$

which is also called the **first difference**.

Differencing

Suppose we assume the trend component μ_t follows a simple linear regression, then

$$\begin{aligned} \nabla x_t &= \nabla(\beta_0 + \beta_1 t + y_t) \\ &= \\ &= \end{aligned}$$

Notice that the detrending may give us a more accurate representation of y_t, whereas differencing completely **removes** β_0 and turns β_1 into the **mean** of the series $\{\nabla x_t\}$.

Differencing

However, differencing changes y_t and often introduces additional dependency. Consider regression with MA(1).

$$x_t = \beta_0 + \beta_1 t + w_t + \theta_1 w_{t-1}$$

Differencing the above expression, we have

$$\begin{aligned} \nabla x_t &= \nabla(\beta_0) + \nabla(\beta_1 t) + \nabla w_t + \theta_1 \nabla w_{t-1} \\ &= \\ &= \end{aligned}$$

which becomes a mean value β_1 plus MA(2). $\{\nabla x_t\}$ is stationary.

Differencing

Question: Suppose instead of a trend, μ_t follows a random walk with drift process, derive ∇x_t.

Differencing Vs Detrending

- An advantage of differencing over detrending is that fewer parameters are estimated after the differencing operation.
- A disadvantage of differencing is that it often makes an estimate of the stationary process y_t more difficult.

Differencing Vs Detrending

- If estimating y_t is the goal, then **detrending** may be more appropriate.
- If the goal is to make the data stationary without thinking about what μ_t is, then **differencing** may be more appropriate.

However, it's dangerous to generalize. You should consider several models, and derive the implications of each tool.

Backshift Operator

The **backshift operator** is defined as

$$Bx_t = x_{t-1},$$

which can be extended to powers

$$B^k x_t = x_{t-k}. \tag{3}$$

Backshift Operator

Thus, the first difference (2) can be written as

Backshift Operator

Question: Denote the second difference, $\nabla^2 x_t$ in terms of the backshift operator.

Backshift Operator

In general, for all positive integer d,

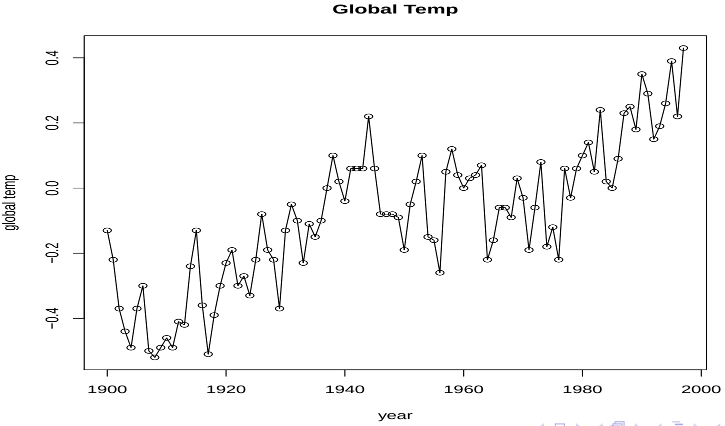
$$\nabla^d = (1 - B)^d. \tag{4}$$

Backshift Operator

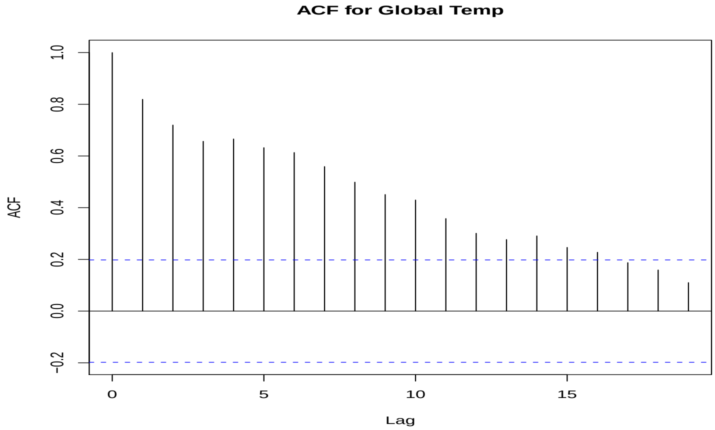
We showed earlier that the first difference eliminates a linear trend. This is an example of a **linear filter** applied to eliminate a linear trend. Next, we show that the second difference eliminates a quadratic trend.

Example: Global Temperature

Let's look at average global temperatures from 1900 to 1997. The data are measured in deviation in degrees centigrade from the 1961-1990 average. You can find the data (global.txt) in Collab.



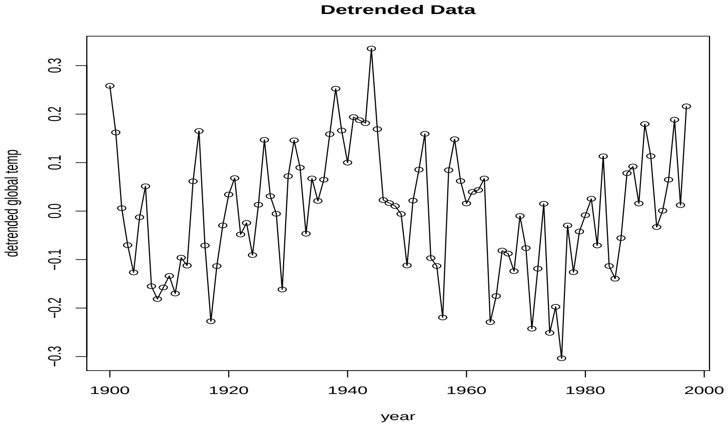
Example: Global Temperature



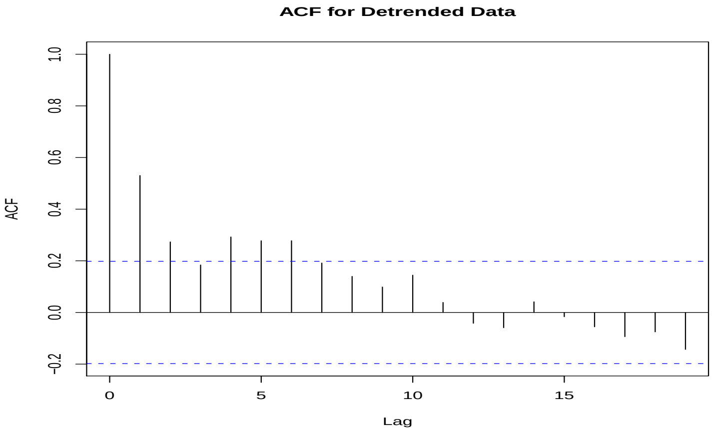
Question: What does this indicate about stationarity?

Example: Global Temperature

Next, we look at detrended data. These are the residuals obtained from regressing x_t on t

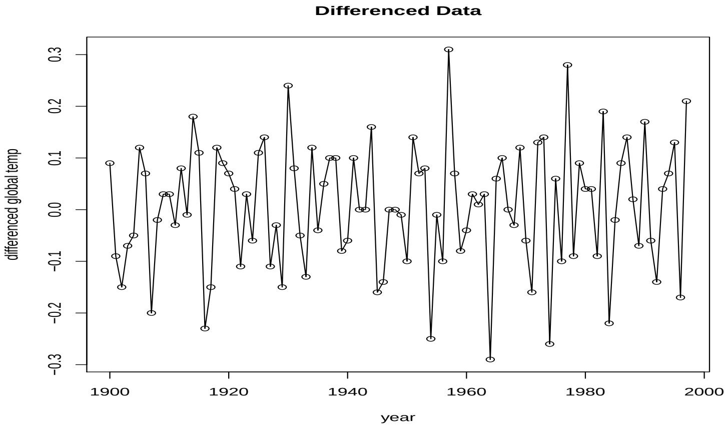


Example: Global Temperature

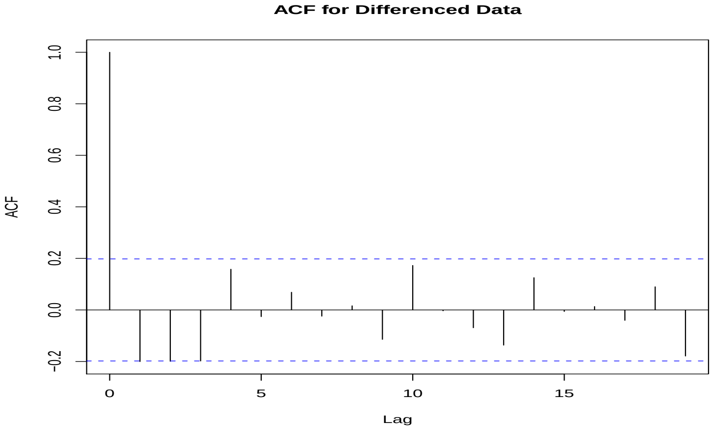


Example: Global Temperature

Finally, we look at differenced data.



Example: Global Temperature



Example: Global Temperature

Question: Based on the plots and ACFs, what can we say about the behavior of global temperature?

Example: Global Temperature

Question: How would you estimate the parameters of the process you are considering for the behavior of global temperature?