STAT 153 - Introduction to Time Series Lecture One

Fall 2022, UC Berkeley

Aditya Guntuboyina

August 29, 2022

1 Time Series Models

The primary objective of time series analysis is to develop mathematical models that provide plausible descriptions for sample time series data.

We assume that the observed data $y_0, \dots, y_T$ is a realization of a sequence of random variables $Y_0, \dots, Y_T$. The basic strategy of modelling is always to start simple and to build up. We shall study the following models in this class:

1. (Gaussian) White Noise
2. Parametric Trend models (parametric function + white noise)
3. Nonparametric Trend models
4. ARIMA models
5. Linear Gaussian State Space Models
6. Nonlinear State Space Models

We shall focus on the Bayesian approach while learning these models.

2 Simplest Model: Gaussian White Noise

The simplest model is

$$Y_0, \dots, Y_T \overset{\text{i.i.d}}{\sim} N(0, \sigma^2).$$

We shall refer to this as Gaussian White Noise or simply White Noise. The variance σ^2 is a parameter in this model. This simple model will never be a good description for any observed time series data. Observed time series often displays various kinds of trends and seasonal patters which cannot be explained by this very simple model.

Time Series often also display "autocorrelation" and this is also not explained by this model. This means the following. Take the observed time series $y_0, \dots, y_T$ and then pair each successive pair of observations to create the bivariate dataset $(y_0, y_1), (y_1, y_2), \dots, (y_{T-1}, y_T)$.

It is often the case that both the individual variables in this bivariate dataset show considerable linear dependence. In other words, the correlation coefficient

$$r := \frac{\sum_{t=0}^{T-1} (y_t - \bar{y}_{(1)})(y_{t+1} - \bar{y}_{(2)})}{\sqrt{\sum_{t=0}^{T-1} (y_t - \bar{y}_{(1)})^2} \sqrt{\sum_{t=0}^{T-1} (y_{t+1} - \bar{y}_{(2)})^2}}, \tag{1}$$

where

$$\bar{y}_{(1)} := \frac{\sum_{t=0}^{T-1} y_t}{T} \quad \text{and} \quad \bar{y}_{(2)} := \frac{\sum_{t=0}^{T-1} y_{t+1}}{T},$$

is often far from zero. On other hand, for the White Noise model, the theoretical correlation between successive observations Y_t and Y_{t+1} is zero. The expression (1) is oftten simplified as

$$r_1 := \frac{\sum_{t=0}^{T-1} (y_t - \bar{y})(y_{t+1} - \bar{y})}{\sum_{t=0}^{T} (y_t - \bar{y})^2} \tag{2}$$

where $\bar{y}$ is the overall mean:

$$\bar{y} := \frac{y_0 + \cdots + y_T}{T+1}.$$

The quantity r_1 is known as the (Sample) Autocorrelation Coefficient at Lag One for the time series data $y_0, \ldots, y_T$. Lag One here refers to the fact that the correlation is between y_t and y_{t+1}. More generally, one can define the Autocorrelation Coefficient at lag k as:

$$r_k := \frac{\sum_{t=0}^{T-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=0}^{T} (y_t - \bar{y})^2}$$

for $k = 1, 2, \ldots$. When $k = 0$, r_0 is by definition equal to 1.

One commonly plots the Autocorrelation Coefficient r_k as a function of k. Generally one or more of these autocorrelations will be quite different from zero. On the other hand, if we generate data from the white noise model, these r_k's will be quite close to zero (especially when T is large). In this sense, the white noise model does not explain autocorrelations present in observed time series. More precisely, it can be shown that if the data $y_0, \ldots, y_T$ are generated according to the white noise model, then the autocorrelations $r_1, r_2, \ldots$ will be **independently** distributed according to the normal distribution with mean zero and variance $1/T$. This variance decreases to zero as T increases and the mean is zero. Thus for large T, the sample autocorrelations should be very close to zero. Also note that the sample autocorrelations for different lags are independent.

A plot of the autocorrelations with lag is known as the "Correlogram". Use the function **acf** in **R** to get the correlogram. The blue bands in the correlogram correspond to levels of $\pm 1.96 T^{-1/2}$ and these can be used as evidence against the suitability of the Gaussian White Noise model for the particular dataset.