Specialist Courses

Daniele Zago

November 22, 2021

Contents

Part I

Time Series

This is a short course aimed at giving an introduction to time series models and their application to modelling and prediction of data collected sequentially in time. The aim is to provide specific techniques for handling data and at the same time to provide some understanding of the theoretical basis for the techniques. Topics covered will include univariate linear and non linear models (both in mean and variance) and some basics of spectral analysis.

Course references

Brockwell and Davis (2016)	Introduction to Time Series and Forecasting
Fan and Yao (2005)	Nonlinear Time Series: Nonparametric and Parametric Methods
Shumway and Stoffer (2017)	Time Series Analysis and Its Applications: With R Examples
Tsay (2013)	Multivariate Time Series Analysis: With R and Financial Applications
Wei (2019)	Multivariate Time Series Analysis and Applications
Brockwell (2009)	Time Series: Theory and Methods
Douc et al. (2014)	Nonlinear Time Series: Theory, Methods and Applications with R Examples

Course contents

- Introduction to linear time series models
- Introduction to spectral analysis
- Transfer function models
- Introduction to nonlinear time series models
  - Threshold AR models
  - Markov Switching models
  - Bilinear models
  - ARCH type models
- Long memory models
- INAR models

Lecture 1: Introduction to time series

2021-11-12

In general, in time series we are interested in *a)* understanding the stochastic mechanism that gives rise to an observed series and *b)* to forecast future values of a series based on the observed history. As this course is introductory, we will restrict our analysis to *univariate* time series.

Assumption We assume that future behaviour is equal to previous behaviour, i.e. we are able to forecast the future based on the observed past data.

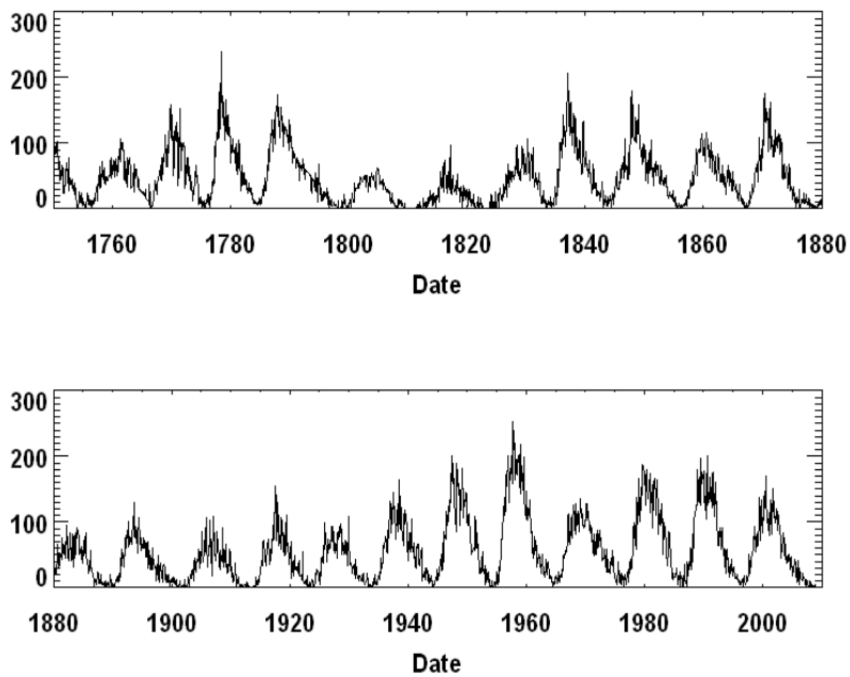

Figure 1: Linear time series models are not able to explain this behaviour, since cyclic components are unequal over time.

There are several approaches in modern time series, namely

› Classical approach

› Modern approach of Box and Jenkins.

› State-space approach of Durbin and Koopman (2012).

1.1 Classical approach

We assume a data-generating process given by the sum of different components

$$Y_t = f(t) + \varepsilon_t, \quad \varepsilon_t \sim \mathrm{WN}(0, \sigma_\varepsilon^2),$$

such that $\mathbb{E}[\varepsilon_t] = 0$, $\mathbb{V}[\varepsilon_t] = \sigma_\varepsilon^2$, $\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0$ for $i \neq j$. Assuming different shapes of $f(t)$ lets us obtain different types of time series, i.e.

› Additive TREND + SEASONAL + CYCLES: $f(t) = T_t + S_t + C_t$

› Multiplicative TREND · SEASONAL · CYCLES: $f(t) = T_t \cdot S_t \cdot C_t$

The classical approach establishes that trend, seasonal, and cyclic components should be estimated separately with simple models and then combined. For example, we can use a linear model for the trend such as

$$T_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \ldots + \alpha_g t^g.$$

In order to model S_t we could use a dummy variables with sine/cosine transform to include cyclic behaviour.

Problem Empirical time series contain deterministic component and stochastic trend components, which cannot be modeled by stationary processes.

1.2 Modern approach

We can consider a DGP as a stochastic process, with time series seen as a realization from a stochastic process. To perform statistical inference, we need to assume that at least some features of the underlying probability law are *stationary* over the time period of interest.

Def. (Stochastic process)
A collection of random variables $X = (X_t)_t$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called a ***stochastic process***

Remark A stochastic process is therefore a function of two arguments $X : \mathcal{T} \times \Omega \to X$, $(t, \omega) \mapsto X_t(\omega)$ and for a fixed value of ω we obtain a *path* from the stochatstic process.

Sample We only observe a portion of the infinite path of the stochastic process,

$$\ldots, X_{-t}, X_{-t-1}, \ldots, X_0, \underbrace{X_1, X_2, \ldots, X_t}_{x_1, x_2, \ldots, x_t}, \ldots,$$

therefore if we want to make inference over the DGP we must make some strong assumptions.

Def. (Mean function)
For a stochastic process X_t, the ***mean function*** is

$$\mu_t = \mathbb{E}[X_t] \quad \text{for } t \in \mathcal{T}.$$

Def. (Autocovariance function)
For a stochastic process X_t, the ***autocovariance function*** is

$$\gamma_{t,s} = \text{Cov}(X_t, X_s) = \mathbb{E}[(X_t - \mu_t)(X_s - \mu_s)], \quad \text{for } t, s = 0, \pm 1, \pm 2, \ldots$$

The autocorrelation function is then defined as

$$\mathrm{Cor}(X_t, X_s) = \frac{\mathrm{Cov}(X_t, X_s)}{\sqrt{\mathbb{V}[X_t]}\sqrt{\mathbb{V}[X_s]}}.$$

We need to make some strong assumptions on the structure of the process in order to make inference possible.

Def. (Strong stationarity)
A process $(X_t)_t$ is ***strictly stationary*** if it is invariant under time shifts, i.e. if

$$(X_{t_1}, \dots, X_{t_n}) \stackrel{\mathrm{d}}{=} (X_{t_1+k}, \dots, X_{t_n+k})$$

for any $n \geq 1$, any choice of $t_1, \dots, t_n$ and al time shifts $k \in \mathbb{Z}$.

Marginals Choosing for instance $n = 1$ means that the marginal distribution of X_t the same as that of X_{t-k} for all t and k.

Def. (Weak stationarity)
A process $(X_t)_t$ is ***weakly stationary*** if

1. $\mathbb{E}[X_t] = \mu < \infty$ for all t.
2. $\mathbb{V}[X_t] = \sigma^2 < \infty$ for all t.
3. $\mathrm{Cov}(X_t, X_{t-k}) = \gamma(k)$ is independent of t for each k.

Weaker Rather than imposing conditions on all possible distributions, we impose conditions only on the first two moments of the series.

Implications

- Strong stationarity + $\mathbb{E}[X_t]^2 < \infty \implies$ weak stationarity.
- Weak stationarity $\not\Longrightarrow$ strong stationarity.
- Weak stationarity + Gaussian $\implies$ Strong stationarity.

Example (Random walk)
For a random walk $Y_t = Y_{t-1} + \varepsilon_t$, we have that $\mathbb{V}[Y_t] = t\sigma^2$ and the process is therefore non-stationary.

Since our objective is to find a model which is able to take into account the ***linear*** dependence between the observations, two very important functions are the autocorrelation and autocovariance functions.

Since for a stationary time series X_t we have $\mathrm{Cov}(X_t, X_{t-k}) = \gamma(k)$ for all k, we can therefore define the ACF as

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)}, \quad k = 0, \pm 1, \pm 2, \dots$$

from which we can see that γ and ρ are even functions, namely

$$\gamma(-k) = \gamma(k), \quad \rho(-k) = \rho(k).$$

Def. (Sample autocorrelation function)
We define the ***sample autocorrelation function*** (ACF) as

$$\widehat{\rho}(k) = \frac{\widehat{\gamma}(k)}{\widehat{\gamma}(0)},$$

where

$$\widehat{\gamma}(k) = \frac{1}{n}\sum_{t=1}^{n-|k|}(X_t - \overline{X})(X_{t+|k|} - \overline{X}).$$

Bias Even if this estimator is biased in finite samples, this is preferred to the unbiased estimator since when dividing by n we have a nonnegative-definite estimator.

Autocorrelogram picture

When fitting a model that relies on correlation between X_t and X_{t-k}, conditioned on immediate variables. We will only include a further lagged variable X_{t-k} in the model for X_t if X_{t-k} makes a genuine and additional contribution to X_t in addition to those from $X_{t-1}, \dots, X_{t-k+1}$. We measure this dependence using the *partial autocorrelation.*

Def. (Partial autocorrelation)
The ***partial autocorrelation*** at lag k is the autocorrelation between z_t and z_{t+k} with the linear dependence of z_t on $z_{t+1}, \dots, z_{t+k-1}$ removed. Namely,

$$\begin{cases} \alpha(1) = \mathrm{Cor}(z_{t+1}, z_t) \\ \alpha(k) = \mathrm{Cor}\left(z_{t+k} - \pi_{t,k}(z_{t+k}), z_t - \pi_{t,k}(z_t)\right) \quad \text{if } k \geq 2 \end{cases}$$

where $\pi_{t,k}(x)$ is the orthogonal projection (regression) of x onto $z_{t+1}, \dots, z_{t+k-1}$.

The ACF and PACF are the main instruments that we use for choosing the most appropriate model for the DGP.

References

Brockwell, P. J. (2009). *Time Series: Theory and Methods.* 2° edizione. New York, N.Y: Springer.

Brockwell, P. J. and Davis, R. A. (2016). *Introduction to Time Series and Forecasting.* Third. Springer Texts in Statistics. New York: Springer-Verlag.

Douc, R. et al. (2014). *Nonlinear Time Series: Theory, Methods and Applications with R Examples.* 1° edizione. Boca Raton: Chapman and Hall/CRC.

Durbin, T. l. J. and Koopman, S. J. (2012). *Time Series Analysis by State Space Methods.* 2 edizione. Oxford: OUP Oxford.

Fan, J. and Yao, Q. (2005). *Nonlinear Time Series: Nonparametric and Parametric Methods.* Springer.

Shumway, R. H. and Stoffer, D. S. (2017). *Time Series Analysis and Its Applications: With R Examples.* 4th ed. New York, NY: Springer.

Tsay, R. S. (2013). *Multivariate Time Series Analysis: With R and Financial Applications.* 1. edizione. Hoboken, New Jersey: John Wiley & Sons Inc.

Wei, W. W. S. (2019). *Multivariate Time Series Analysis and Applications.* 1. edizione. Hoboken, NJ: John Wiley & Sons Inc.