MAE 5870 – Análise de Séries temporais
Lista #2
Data de entrega: 20/04/2017

2.1 For the Johnson & Johnson data, say y_t, shown in Figure 1.1, let $x_t = \log(y_t)$.

(a) Fit the regression model

$$x_t = \beta t + \alpha_1 Q_1(t) + \alpha_2 Q_2(t) + \alpha_3 Q_3(t) + \alpha_4 Q_4(t) + w_t$$

where $Q_i(t) = 1$ if time t corresponds to quarter $i = 1, 2, 3, 4$, and zero otherwise. The $Q_i(t)$'s are called indicator variables. We will assume for now that w_t is a Gaussian white noise sequence. What is the interpretation of the parameters β, α_1, α_2, α_3, and α_4? (*Detailed code is given in Appendix R on page 574.*)

(b) What happens if you include an intercept term in the model in (a)?

(c) Graph the data, x_t, and superimpose the fitted values, say $\widehat{x}_t$, on the graph. Examine the residuals, $x_t - \widehat{x}_t$, and state your conclusions. Does it appear that the model fits the data well (do the residuals look white)?

2.2 For the mortality data examined in Example 2.2:

(a) Add another component to the regression in (2.25) that accounts for the particulate count four weeks prior; that is, add P_{t-4} to the regression in (2.25). State your conclusion.

(b) Draw a scatterplot matrix of M_t, T_t, P_t and P_{t-4} and then calculate the pairwise correlations between the series. Compare the relationship between M_t and P_t versus M_t and P_{t-4}.

2.6 Consider a process consisting of a linear trend with an additive noise term consisting of independent random variables w_t with zero means and variances σ_w^2, that is,

$$x_t = \beta_0 + \beta_1 t + w_t,$$

where β_0, β_1 are fixed constants.

(a) Prove x_t is nonstationary.

(b) Prove that the first difference series $\nabla x_t = x_t - x_{t-1}$ is stationary by finding its mean and autocovariance function.

(c) Repeat part (b) if w_t is replaced by a general stationary process, say y_t, with mean function μ_y and autocovariance function $\gamma_y(h)$.

2.8 The glacial varve record plotted in Figure 2.6 exhibits some nonstationarity that can be improved by transforming to logarithms and some additional nonstationarity that can be corrected by differencing the logarithms.

(a) Argue that the glacial varves series, say x_t, exhibits heteroscedasticity by computing the sample variance over the first half and the second half of the data. Argue that the transformation $y_t = \log x_t$ stabilizes the variance over the series. Plot the histograms of x_t and y_t to see whether the approximation to normality is improved by transforming the data.
(b) Plot the series y_t. Do any time intervals, of the order 100 years, exist where one can observe behavior comparable to that observed in the global temperature records in Figure 1.2?
(c) Examine the sample ACF of y_t and comment.
(d) Compute the difference $u_t = y_t - y_{t-1}$, examine its time plot and sample ACF, and argue that differencing the logged varve data produces a reasonably stationary series. Can you think of a practical interpretation for u_t? Hint: For $|p|$ close to zero, $\log(1+p) \approx p$; let $p = (y_t - y_{t-1})/y_{t-1}$.
(e) Based on the sample ACF of the differenced transformed series computed in (c), argue that a generalization of the model given by Example 1.23 might be reasonable. Assume

$$u_t = \mu + w_t - \theta w_{t-1}$$

is stationary when the inputs w_t are assumed independent with mean 0 and variance σ_w^2. Show that

$$\gamma_u(h) = \begin{cases} \sigma_w^2(1+\theta^2) & \text{if } h = 0, \\ -\theta\, \sigma_w^2 & \text{if } h = \pm 1, \\ 0 & \text{if } |h| > 1. \end{cases}$$

(f) Based on part (e), use $\widehat{\rho}_u(1)$ and the estimate of the variance of u_t, $\widehat{\gamma}_u(0)$, to derive estimates of θ and σ_w^2. This is an application of the method of moments from classical statistics, where estimators of the parameters are derived by equating sample moments to theoretical moments.

2.9 In this problem, we will explore the periodic nature of S_t, the SOI series displayed in Figure 1.5.

(a) Detrend the series by fitting a regression of S_t on time t. Is there a significant trend in the sea surface temperature? Comment.
(b) Calculate the periodogram for the detrended series obtained in part (a). Identify the frequencies of the two main peaks (with an obvious one at the frequency of one cycle every 12 months). What is the probable El Niño cycle indicated by the minor peak?

2.11 Consider the two weekly time series `oil` and `gas`. The oil series is in dollars per barrel, while the gas series is in cents per gallon; see Appendix R for details.

(a) Plot the data on the same graph. Which of the simulated series displayed in §1.3 do these series most resemble? Do you believe the series are stationary (explain your answer)?

(b) In economics, it is often the percentage change in price (termed *growth rate* or *return*), rather than the absolute price change, that is important. Argue that a transformation of the form $y_t = \nabla \log x_t$ might be applied to the data, where x_t is the oil or gas price series [see the hint in Problem 2.8(d)].

(c) Transform the data as described in part (b), plot the data on the same graph, look at the sample ACFs of the transformed data, and comment. [Hint: `poil = diff(log(oil))` and `pgas = diff(log(gas))`.]

(d) Plot the CCF of the transformed data and comment The small, but significant values when `gas` leads `oil` might be considered as feedback. [Hint: `ccf(poil, pgas)` will have `poil` leading for negative lag values.]

2.12 Use two different smoothing techniques described in §2.4 to estimate the trend in the global temperature series displayed in Figure 1.2. Comment.