Final project STK4060/STK9060-sp14 - Time Series

This is the problem set for the project part of the finals in STK4060-sp14. The reports shall be individually written. You may discuss the solutions with your fellow students, but the intention is that the final formulations shall be done individually.

The deadline for turning in the reports is

Monday May 12'st at 4 pm.

Two copies marked with your candidate number shall be placed in Anders Rygh Swensen's post box at room B700 at the seventh floor in N. H. Abel's house. Handwritten reports are acceptable. Enclose the parts of the computer outputs which are necessary for the answering the questions. The other parts can be collected in appendices. When you refer to material in these, be careful to indicate explicitly where.

Anders Rygh Swensen

Problem 1

Consider the Norwegian quarterly time series for total private consumption and export of traditional goods, i.e except oil and natural gas, for the period 1970:1 to 2013:4, which can be found on the course web page as `consumption` and `export`.

The amounts are in million kr with basis year 2011, so the total consumption in the fourth quarter 2013 was around 300 billions kr.

a) Use the techniques for ARIMA and seasonal ARIMA modeling to determine suitable models for the two series. Estimate the unknown coefficients. Remember that a log transform is often applied for this kind of series.

b) Plot the residuals, and describe how you evaluate the fitted model from the properties of the residuals.

The purpose with this problem is not that you find an waterproof model, but rather that you explain the choices you do in the modeling process. Relate your model search to the strategy described on pages 159-160 in Shumway and Stoffer: Time Series Analysis and its Applications. Two pages with explanations for each series are sufficient in addition to the plots and R-output.

Problem 2

Consider the weakly stationary ARMA(2,1) time series defined by

$$x_t - 0.1x_{t-1} - 0.06x_{t-2} = w_t - 0.5w_{t-1}$$

where $\{w_t\}$ is white noise with mean zero and variance σ_w^2.

a) Explain why the series is causal and find the coefficients in the representation $x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j},\ \psi_0 = 1, \sum_{j=0}^{\infty} |\psi_j| < \infty$.

b) Explain why the series is invertible and find the coefficients in the representation $w_t = \sum_{j=0}^{\infty} \pi_j x_{t-j},\ \pi_0 = 1, \sum_{j=0}^{\infty} |\pi_j| < \infty$.

c) Find the expectation of

$$(x_t - 0.1x_{t-1} - 0.06x_{t-2})x_{t-h}, h = 0, 1, \dots$$

and

$$(w_t - 0.5w_{t-1})x_{t-h}, h = 0, 1, \dots$$

and show how this can be used to formulate a difference equation to determine the covariance $\gamma(h) = E[x_{t+h}x_t]$.

d) Solve the difference equation from part c) and obtain an expression for $\gamma(h) = E[x_{t+h}x_t]$.

Problem 3

In this exercise we shall consider some aspect of an alternative estimator of the spectral density to smoothing of the periodogram which is discussed i Shumway and Stoffer: Time Series Analysis and its Applications on pages 209-212.

Let $x_1, \cdots, x_n$ be realizations from a stationary time series with spectral density $f(\omega)$. Consider estimators of the following form

$$\tilde{f}(\omega) = \sum_{|h| \le r} w(\frac{h}{r}) \widehat{\gamma}(h) \exp^{-2\pi i \omega h}$$

where $\widehat{\gamma}(h)$ is the covariance function and w is an even function satisfying $w(0) = 1$, $|w(x)| \le 1$ all x and $w(x) = 0, |x| > 1$. The function $w(.)$ is called the *lag window.*

Define

$$\tilde{I}_n(\omega) = \sum_{|h|<n} \widehat{\gamma}(h) \exp^{-2\pi i \omega h},\ 1/2 \le \omega \le 1/2.$$

which coincides with the periodogram at the Fourier frequencies $0, 1/n, \dots, (n-1)/n$, and also the so-called *smoothing window*

$$W(\omega) = \sum_{|h| \le r} w(\frac{h}{r}) \exp^{-2\pi i \omega h}.$$

a) Show that $\widehat{\gamma}(h) = \int_{-1/2}^{1/2} \tilde{I}_n(\omega) \exp^{2\pi i \omega h} \, d\omega$.

b) Then show that

$$\tilde{f}(\omega) = \int_{-1/2}^{1/2} W(\lambda)\tilde{I}_n(\omega + \lambda) d\lambda$$

and explain why $\tilde{f}(\omega_k)$ can be approximated by

$$\sum_{|j| \leq [n/2]} (W(\omega_j)/n) I_n(\omega_k + \omega_j),$$

where I_n is the periodogram. Thus $\tilde{f}(\omega_k)$ is approximated by a weighted sum of the periodogram ordinates with weights $W_n(j) = W(\omega_j)/n, |j| \leq [n/2]$. Here $[n/2]$ is the integer value of $n/2$, which is the largest integer smaller or equal to $n/2$.

Problem 4

In this problem we will consider the forecasting method exponentially weighted moving average (EWMA) which is described in Example 3.28 in the textbook by Shumway and Stoffer.

Consider an ARIMA(0,1,1) model of the form

$$x_t = x_{t-1} + w_t - \lambda w_{t-1} \text{ where } x_0 = 0, \ |\lambda| < 1.$$

where $w_t \sim wn(0, \sigma_w^2)$. Let $\tilde{x}_{n+1}$ denote the conditional expectations $\tilde{x}_{n+1} = E[x_{n+1} | x_n, x_{n-1}, \ldots, x_1, x_0, x_{-1}, \ldots]$.

a) Explain why

$$x_t = \sum_{j=1}^{\infty} (1-\lambda)\lambda^{j-1} x_{t-j} + w_t \tag{1}$$

is an approximation to x_t for large t.

b) Use the result from part a) to argue why

$$\tilde{x}_{n+1} = (1-\lambda)x_n + \lambda \tilde{x}_n$$

and that the truncated predictor $\tilde{x}_{n+1}^n$, obtained by setting $x_0 = x_{-1} = \cdots = 0$, satisfies

$$\tilde{x}_{n+1}^n = (1-\lambda)x_n + \lambda \tilde{x}_n^{n-1}, \ n = 1, \ldots$$

Would you chose $\tilde{x}_1^0 = 0$ or $\tilde{x}_1^0 = x_1$ as initial value?

Thus the forecast for the next period is a linear combination of the forecast for the present period and the observed value for this period.

We shall now see that a similar scheme arise in forecasting using a simple state space model described as a random walk plus a noise term. Consider the local level model

$$\begin{aligned} x_t &= x_{t-1} + w_t,\ t = 1, \dots, n \\ y_t &= x_t + v_t,\ t = 1, \dots, n \end{aligned} \tag{2}$$

with the initial value x_0 and where $w_t \sim i.i.d.N(0, q\sigma^2)$ and $v_t \sim i.i.d.N(0, \sigma^2)$ are independent.

c) Formulate this as a state space model and explain what the observation and the expectation equation are.

d) If $x_t^{t-1} = E[x_t|y_{t-1}, \dots, y_1]$ and $P_t^{t-1} = var(x_t|y_{t-1}, \dots, y_1)$ with initial values $x_1^0 = x_0$ and $P_1^0 = \kappa$, explain why the Kalman filter can be written

$$\begin{aligned} x_{t+1}^t &= (1 - K_t)x_t^{t-1} + K_t y_t \\ K_t &= P_t^{t-1}/(1 + P_t^{t-1}) \\ P_{t+1}^t &= P_t^{t-1} - [(P_t^{t-1})^2/(1 + P_t^{t-1})] + q \end{aligned}$$

e) Let $P_1^0 = \kappa$. Show that when $\kappa \to \infty$, $x_2^1 \to y_1$ and $P_2^1 \to 1 + q$.

f) Show that the sequence P_{t+1}^t converges as t increases.

g) Then explain why $P_{t+1}^t \to p$ where $p = (q + \sqrt{q^2 + 4q})/2$.

h) Explain why the Kalman filter is for large t approximated by the recursion $\tilde{x}_{t+1} = (1 - k)\tilde{x}_t + k y_t$. What is k as a function of q? Verify that $|1 - k| < 1$.

i) Discuss the effect of increasing values of the signal-noise ratio q.

Problem 5

Consider the ratio of quarterly private consumption and quarterly gross national product for mainland Norway for the period 1970:1 to 2013:4, i.e $x_t = \text{consump}_t/\text{gnpmnlnd}_t$. The series can be found on the course web page as $\texttt{consumption}_t$ and $\texttt{gnpmnlnd}_t$.

a) Compute the autocorrelation function, ACF, and the partial autocorrelation function, PACF, to the ratio. Discuss the main features, especially the periodicity.

b) Graph the periodogram , and show the effect of estimating the spectrum by smoothing the perodogram using various smoothing methods.

c) Now consider the fourth difference $x_t - x_{t-4}$. Perform the same operations as in part b). How do you explain the plots now?