Mathematical Statistics
Institute of Applied Mathematics and Mechanics
University of Warsaw
John M. Noble

Time Series
Final Assessment (Computer)
Deadline: Saturday 6th July 2024, 13.00

Exercise 1

The file `aa-rv-20m.txt` contains the realised daily volatility series of Alcoa stock returns from January 2, 2003 - May 7, 2004 based on 20-minute intradaily log returns.

1. Fit an ARIMA(0,1,1) model to the log volatility series and write down the model.

2. Estimate the local trend model, of the form:

$$\begin{cases} y_t = \mu_t + e_t & e_t \sim N(0, \sigma_e^2) \\ \mu_{t+1} = \mu_t + \eta_t & \eta_t \sim N(0, \sigma_\eta^2) \end{cases}$$

   for the log volatility series. What are the estimates of σ_e and σ_η? Obtain time plots for the filtered and smoothed state space variables with pointwise 95% confidence interval.

Exercise 2

Consider the monthly simple excess returns of Pfizer stock and the S&P 500 composite index from January 1990 to December 2003. The excess returns are in `m-pfesp-ex9003.txt` with Pfizer stock returns in the first column.

1. Fit a fixed-coefficient market model to the Pfizer stock returns. Write down the fitted model.

2. Express a time-varying CAPM (Capital Asset Pricing Model) as a state space model and fit a CAPM to the Pfizer stock return. This is a model of the form:

$$\begin{cases} r_t = \alpha_t + \beta_t r_{M,t} + e_t & e_t \sim N(0, \sigma_e^2) \\ \alpha_{t+1} = \alpha_t + \eta_t & \eta_t \sim N(0, \sigma_\eta^2) \\ \beta_{t+1} = \beta_t + \epsilon_t & \epsilon_t \sim N(0, \sigma_\epsilon^2) \end{cases}$$

   Here r_t is the excess return of an asset, while $r_{M,t}$ is the excess return of the market. The S&P returns are used for the market.

   What are the estimated standard errors of the innovations to the α_t and β_t series? Obtain time plots of the smoothed estimates of α_t and β_t.

Exercise 3

The file `m-ppiaco4709.txt` contains year, month, day, and U.S. producer price index (PPI) from January 1947 to November 2009. The index is for all commodities and not seasonally adjusted. Let $z_t = \ln(Z_t) - \ln(Z_{t-1})$, where Z_t is the observed monthly PPI. It turns out that an AR(3) model is adequate for z_t, if the minor seasonal dependence is ignored. Let y_t be the sample mean-corrected series of z_t.

1. Fit an AR(3) model to y_t and write down the fitted model.
2. Suppose that y_t has independent measurement errors so that $y_t = x_t + e_t$ where x_t is a zero-mean AR(3) process and $\text{Var}(e_t) = \sigma_e^2$. Use a state-space form to estimate the parameters, including the innovational variances to the state and σ_e^2. Write down the fitted model and obtain a time plot of the smoothed estimate of x_t. Also, show the time plot of the filtered response residuals of the fitted state space model.

Exercise 4

Consider the simple returns of IBM stock, CRSP value-weighted index, CRSP equal-weighted index, and the S&P composite index from January 1980 to December 2008 (understanding the meaning of these terms is not necessary for the exercise). The index returns include dividend distributions. The file is `d-ibm3dxwkdays8008.txt` which has 12 columns. The columns are (year, month, day, IBM, VW, EW, SP, M, T, W, R, F), where M, T, W, R, F denotes indicator variables for Monday to Friday respectively.

Use a regression model to study the effects of trading days on the equal-weighted index returns (column EW). What is the fitted model? Are the weekday effects significant in the returns at the 5% significance level? Are there correlations in the regression residuals? If so, build a suitable time series model.

Exercise 5

Consider the monthly simple returns of GE stock from January 1926 to December 2008. They are found in `m-ge2608.txt` in the course directory. Use the last three years of data for forecast evaluation.

1. Using lagged returns $r(t-1)$, $r(t-2)$, $r(t-3)$ as input, build a 3-2-1 feed forward neural network to forecast 1-step-ahead returns. Calculate the mean squared error of forecasts.
2. Again, using $r(t-1)$, $r(t-2)$, $r(t-3)$ and also their *signs*, build a 6-5-1 feed forward neural network to forecast the 1-step-ahead GE stock price *movement* (that is, we're only interested in whether the price goes up or down; not by how much), with 1 denoting upward movement. Calculate the mean squared error of the forecasts.
   If `rtn` denotes return, you can create a direction variable by:

   ```
   drtn = ifelse(rtn>0,1,0)
   ```

Exercise 6: Forecasting

The `USDEUR.csv` data set gives exchange rate data for the US dollar and Euro over a 5 year period 2008 - 2013. Consider the various prediction techniques discussed in the course, discuss their performance and select the technique that gives the greatest accuracy of prediction (as measured by sum of squares of forecast errors).

Take the data or two consecutive years (which you may select). Construct your prediction for the *second* year in your sample for each series. The first year sample serves as a training period and gives you memory for the process (so that you have sufficient data to generate forecast for the first trading day in January of the second year, for example).

To generate a forecast for day t you can only use information up to date or time $t-1$ (e.g. to create a forecast for May 6th 2012 use only information up to and including May 5th 2012. Your prediction should be recursive; at t you make a forecast for $t+1$, at $t+1$ you make a forecast for $t+2$, etc.

Please submit an R program that computes your forecast. It should take a price series (with a date) and return as output a forecast of your series (again with a date). Your prediction should cover at least the entire second year of the data.

With the R program please submit a brief explanation of your methodology.

For each forecast, compute the sum of squares of the loss:

$$\sum_t |p_t - \widehat{p}_t|^2$$

where p_t and $\widehat{p}_t$ are price at time t and its forecast, respectively.
For your forecasting model, compute the one-step forecasts on the years not used for prediction.

While working on your forecasts consider the following:

1. Do you want to work in *price* or *return* space? If you decide to forecast returns, construct log returns, and remember to translate your forecast of returns to forecasts of prices.

2. Parsimonious forecasts often work best in live applications. Keep in mind the trade-off between good performance in your sample and reliability of your model on the data you do not see.

3. You may want to use material covered in class, but do not feel obliged to do so or constrained by the material covered. Anything that gives an accurate forecasting method is encouraged.