STAT 580 : Time Series Analysis

Take Home Final (Due: 5PM on the scheduled final exam day)

Note: You are not allowed to work with other students–I will be very strict on this. Your answer must be concise and esay to follow. Your answer must be well organized in separated papers, other than computer outputs. You may cut and paste the plots inside your answers or may attach those as references at the end. Hand-written exam or late exam will not be accepted

1. **Analysis of Influenza Deaths.** The file *infl.asd* shows the number of monthly influenza deaths (n=132 months) in the U.S. for the ten year period 1968-1978. Perform your analyses on the *mean-adjusted* monthly influenza.

   (a) Examine the autocorrelation and partial autocorrelation functions and give one possible suggested ARIMA model that is sugeested by the PACF.

   (b) An analyst decided to investigate ARIMA models of the form $ARIMA(p, d, q) \times (P, D, Q)_{12}$ for this data. Using the ACF and PACF carefully identify possible min and max of each order, $p, d, q, P, D, and\ Q$. Perform the ARIMA search and identify the best model using AIC_C. Be sure to confirm that your choice is reasonable by using residual analysis and checking the statistical significance of the parameters. Write the final model. Your expression shouldn't involve the backshift operator B.

   (c) Compute the 12 month forecast and prediction limits under the best model obtained in (b).

   (d) Another analyst claims that the model $ARIMA(4, 0, 0) \times (0, 0, 0)_{12}$ would be a better choice than the model obtained in (b). What do you think of this assertion? How can you support your assertion?

2. **Revisiting the Oil and Gas Series.** In Homework Problem 1.3, you looked at monthly oil and gas prices in the file *oil.asd*, say O_t and G_t, measured over the period beginning in July, 1973 and ending in December, 1987. In that problem, it seemed to be reasonable to consider the transformed variables

   $$x_t = \log \frac{O_t}{O_{t-1}}$$

   and

   $$y_t = \log \frac{G_t}{G_{t-1}}.$$

   We would like to look at (1) predicting y_t from its own past using an ARIMA models as compared with (2) predicting y_t from its own past as well as from x_t and its past using transfer function methodology.
   Answer the questions below with this general objective in mind.

   (a) Provide the CCF between pre-whitten x_t and y_t and interpret.

   (b) Fit an ARIMA model to the transformed oil series x_t and save the residuals and estimated model parameters. Apply the ARIMA transformation from this model to y_t and compute the cross correlation function between the two residuals. Argue that this suggests the model

   $$y_t = \beta_1 x_t + e_t$$

   Compare with the model that might be suggested by the impulse response function relating x_t to y_t.

   (c) Fit the regression model suggested in (c) and verify the structure of the residuals. What does this imply about e_t? Use the $ARIMA(1, 0, 0)$ for x_t combined with the regression model to write the forecast function for y_t. Predict the next 12 percentage changes using this model.

   (d) Fit an ARIMA model to the gas price series y_t and predict the next 12 percentage changes in gas price.

   (e) What would be a reasonable method for comparing the two models. Justify your answer and use ASTSA to make the comparison.

3. Consider the SOI series we discussed in class. From the spectrum of the series, we noted that essentially two components have power, the El Nino frequency at about 0.02 cycles per month and a yearly frequency of about 0.08 cycles per month.

   (a) Assume that we wish to preserve the lower frequency as signal and to eliminate the higher frequencies. Design a filter for this purpose and apply the filter to the series. Give the plot of the filtered series and its spectrum. Discuss your findings.

   (b) Now, smooth the original series using the symmetric 12-month moving average. What is the effect of the smoothing filter. Compare the smoothed series and the signal extracted in (a) in various ways.

4. Consider the *varve* data.

   (a) Fit an ARIMA(0,1,1) model to the logarithmically transformed series y_t. Give 10 year forecasts with 95% forecast interval in original scale.

   (b) As an alternative to the model in (a), consider fitting the state-space model

   $$y_t = x_t + v_t$$

   where x_t is an unoberved *climate signal* satisfying

   $$x_t = x_{t-1} + w_t$$

   where the parameters to be estimated are $var v_t = \sigma_v^2$, $var w_t = \sigma_w^2$ and $\mu = Ex_o$, the mean value of the initial state. Assume that $var x_0 = 1$. Give 10 year forecasts with 95% forecast interval in original scale. Compare with the results obtained in (a).