Time Series Project

Due Thursday, July 27

Your paper (not including appendices) should be no more than 15 pages, and written as an essay. Look at a few applied time series journal articles to see how this is done. You can put any extra information (for example computer programs and output) in an appendix at the end of the paper, but the first 15 pages should be self-contained. I will only look in the appendix in exceptional circumstances so that all the relevant information should be in the main body of the paper.

Good writing is important. I can't give you a good grade if I don't understand what you write. If you are a poor writer, have a friend help you, or pay somebody to help you with your writing.

When presenting your results you need to convince me that you know what you are doing, and that you have done everything correctly. Simply writing down numbers without any explanation does not tell me how you arrived at these numbers, and it is how you arrived at your numbers that I am interested in. Deciding what to include and what not to include is always difficult. But to help you to decide, imagine I am thinking you cheated and you simply made up your results, or that you don't know anything and did everything wrong. What would you need to write to convince me otherwise?

Here then is what I want for the project.

1. Let W_{1t} be the raw seasonally adjusted Canadian GDP series in the attached file "GDP_CONS_CANADA." Let W_{2t} be the raw time series that you worked with in Assignment 1. Plot both raw time series.

2. Define $X_t = \ln(W_{1t})$. Describe your data series. Construct a measure of the business cycle Y_t using both the trend stationary (TS) and the difference stationary (DS) approaches. Plot Y_t for both cases.

3. Fit an $AR(p)$ process to both the TS and DS Y_t using the Bayesian Information Criterion to decide on the appropriate value of p. If $p < 2$ set $p = 2$ for both the TS series or the DS series as required. (This is to force you to do some work, AR(1) or white noise is too easy!) From the estimated TS model for Y_t, calculate $\gamma(0)^{1/2}$, the infinite moving average weights ψ_k, and the autocorrelation functions $\rho(k)$ for $k = 0, 1, 2, ..., 6$.

4. Using both the DS and TS Y_t (with $p \geq 2$), forecast the growth rate as $E_T[\triangle X_{T+k}]$ for $k = 0, 1, 2, ..., 8$, where T is the last observation in your sample. Provide separate graphs of your TS and DS forecasts, and in the graphs plot 95% confidence intervals for $\triangle X_{T+k}$.

5. Do an augmented Dickey-Fuller test on X_t, augmenting the regression with 5 lags (or if you wish use $AIC(k)$ or $BIC(k)$). Based on the p-value, decide if X_t is Difference or Trend stationary.

6. Use Box-Jenkins identification to identify and estimate an $ARMA(p, q)$ model for both the DS and TS Y_t. Plot the standardized residuals $z_t = \frac{\varepsilon_t}{\sigma}$ for both models and comment about whether the normal distribution appears to be appropriate. Perform the following diagnostic tests on both $ARMA(p, q)$ models: 1)Box-Pierce, 2) Overfitting with $r = 4$, 3) Jarque-Bera, 4) a test for $ARCH(6)$.

7. Program the log-likelihood for the $AR(1)$
$$Y_t = \phi Y_{t-1} + \sigma z_t,$$
where z_t has a student's t distribution with 5 degrees of freedom. Using either TS or DS series, find the maximum likelihood estimates of ϕ and σ. Now, let z_t have a student's t distribution with r degrees of freedom and estimate r. Estimate the regime switching model
$$Y_t = \delta_t \phi^E Y_{t-1} + (1 - \delta_t)\,\phi^R Y_{t-1} + a_t, \quad a_t \sim i.i.N\left[0, \sigma^2\right],$$
where $\delta_t = 1$ if $Y_{t-1} \geq c$ and $\delta_t = 0$ if $Y_{t-1} < c$, and where $c = 0$. Now, estimate c. For extra marks, find another type of time series model we have not covered in class, and estimate this model.

8. Find a monthly (or higher frequency) financial time series P_t that is either an exchange rate or a stock index. Using the estimated autocorrelation function $\widehat{\rho}_a(k)$ for a_t, determine if this series follows a random walk as
$$\ln(P_t) = \delta + \ln(P_{t-1}) + a_t \text{ with } a_t \sim i.i.N\left[0, \sigma^2\right].$$

Test whether a_t is normally distributed. Estimate the autocorrelation function for a_t^2. Estimate a $GARCH(1,1)$ model for a_t. For extra marks estimate another type of model for $\ln(P_t)$ not covered in class that displays conditional heteroskedasticity.

9. Consider W_{1t} and W_{2t} in part (1). Define

$$Y_{1t} \equiv \triangle X_{1t} - \mu_1, \quad Y_{2t} \equiv \triangle X_{2t} - \mu_2,$$

where $X_{1t} \equiv \ln(W_{1t})$ and $X_{2t} \equiv \ln(W_{2t})$. Construct the 2×1 vector time series

$$Y_t \equiv \begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix}.$$

Estimate a $VAR(p)$ for Y_t using $BIC(k)$ to determine p. Test

$$H_0 : Y_t \sim VAR(p) \text{ versus } H_0 : Y_t \sim VAR(p+2)$$

using a likelihood ratio test. Test for causality in both directions between Y_{1t} and Y_{2t}, including instantaneous causality using $p \geq 2$. What do you conclude about the direction of causality based on these tests? Calculate $E_t\left[c^{'}Y_{t+2}\right]$ and a 95% confidence interval for $c^{'}Y_{t+2}$, where

$$c = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Test whether X_{1t} and X_{2t} are cointegrated by regressing

$$X_{1t} = c + \beta X_{2t} + Z_t$$

and performing an augmented Dickey-Fuller test on Z_t with $p = 5$. Estimate the *Error Correction Model*

$$\triangle X_t = \alpha + dZ_{t-1} + \phi_1 \triangle X_{t-1} + \phi_2 \triangle X_{t-2} + a_t,$$

where α and d are 2×1 vectors.