FINANCIAL ECONOMETRICS

ASSIGNMENT 3

INSTRUCTIONS

In time series analysis, distinguishing conditional from unconditional moments becomes paramount. For this assignment, Ruey Tsay's *Analysis on Financial Time Series* should be a useful guide if you can find a copy.

DETAILS ON SUBMISSION

Like Assignment 2, the solution to this assignment needs to be written in Word (or Latex). Do not include the code in the write-up: only report what has been asked and what matters for interpreting the results. In addition to the write-up, send your code that generates the results to the TA so that they can check your work.

FINITE SAMPLE AR(1) BIAS

Consider the dynamics of an AR(1) model:

$$x_t = \phi x_{t-1} + \epsilon_t, \quad \epsilon_t \overset{iid}{\sim} \mathcal{N}(0,1).$$

In small samples, the OLS estimate of the first-order autoregression coefficient is biased. In this exercise we are going to run a simple Monte Carlo experiment to evaluate the magnitude of the bias.

- Consider the same set $\phi = \{0.5, 0.9, 0.98\}$ for different sample sizes $T = \{25, 100, 1000\}$, for a total of nine (ϕ, T) pairs
- For each (ϕ, T) pair do the following:
    1. Simulate a sample of observations $\{x_t\}_{t=1}^{T}$. Initialize x_0 at the unconditional mean of zero, so that $x_1 = \epsilon_1$
    2. Using the simulated sample, compute the OLS estimate of $\widehat{\phi}$ and save it
    3. Repeat 1. and 2. 500 times
- Now for each (ϕ,T) pair, compute the mean of the 500 estimated autoregression coefficients $\widehat{\phi}$ and the frequency of $\mathbf{1}\{\widehat{\phi} < \phi\}$. Report the mean and frequency results in two 3 x 3 tables (ϕ x T tables). Discuss your results

GDP FORECASTING

Go to FRED and download quarterly data on real gross domestic product at

`https://fred.stlouisfed.org`

The series ID is *GDPC1* and you should be able to directly type this into FRED's search engine and retrieve it. The data will start in 1947Q1 and end at the most recent data point.

1. Convert the level data into an annual growth rate (year-on-year percentage change) and plot the data along with the NBER recession dates: `http://www.nber.org/cycles.html`

   Note: after conversion, you should have quarterly data from 1948Q1 until now, where each quarterly data point is an annual growth rate; for example, the 1948Q1 data point is the annual growth at 1948Q1 relative to 1947Q1; 1948Q2 data point is annual growth at 1948Q2 relative to 1947Q2; and so on

   Tip: if you can figure it out, you might be able to get FRED to plot the NBER dates for you

2. We're going to try to forecast the severity of the 2008 recession using data only up until 2007Q4. Restrict the data sample to this period and do all of the following analysis on this restricted sample period

3. Compute the ACF for 12 lags and report the plot as a function of lag j. Use the ACF function we programmed together in class. Does it oscillate and what does that tell you about what order AR you would need in order to accurately capture the series' dynamics? If you thought the process were a pure MA(q), how many MA lags would you want to consider?

4. Now look at the PACF for 12 lags and also report its plot as a function of lag j. Similarly to our ACF program, write a PACF program that takes data and a lag (J) input and returns the first 12 PACF lags (as with the ACF, the PACF at lag 0 is equal to one). If you thought the process were a pure AR(p), how many AR lags would you want to consider?

5. Let's assume that the process is a pure AR, so that we can use our analysis in the previous question (4.) to determine the correct AR model. For this assignment, assume that the first PACF lag that is not different from zero is one lag too many for the candidate AR model; that is, choose only the first p lags that are significant.

   Estimate the AR(p) you have chosen and report all coefficients, including the estimate of error variance ($\widehat{\sigma}^2$):

   $$x_{t+1} = a + \phi_1 x_t + \phi_2 x_{t-1} + \cdots + \phi_p x_{t-p+1} + \epsilon_{t+1}, \quad \epsilon_{t+1} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$$

   What is the unconditional mean of the AR(p) model? What is the time series mean of the data? Are they close?

6. With this model chosen, let's now look at the residuals of the fitted AR(p) model. Do they look like white noise? Because the true data-generating process could potentially be a ARMA(p,q) order and we're only considering AR(p) model they might not. That is fine for this assignment. Show the plot and describe it in words. And we'll proceed anyway

7. Now let's forecast out-of-sample using our data up until 2007Q4. Forecast the next eight quarters using the model. You do not need to compute the variance of the forecast error. Simply compute the conditional expectation of real GDP growth eight quarters head: $\mathbb{E}_t[x_{t+h}]$ for $h = 1, 2, \ldots, 8$. For part of the forecast I recommend doing a loop. Plot the model's fitted values and its 2007Q4 forecast over the entire true data. Does the model do a good job of forecasting the recession? Why or why not? To help you answer this question, plot the unconditional mean of the model over the figure