14.384: Time Series Analysis.

Bank of sample problems for 14.384 Time series

Disclaimer. The problems below do not constitute the full set of problems given as homework assignments for the course. Some of the problems are well-known folklore, some were inspired by the problem sets given at different times at Harvard, Upenn and Duke. I am thankful to Jim Stock, Frank Schorfheide and Barbara Rossi for giving me access to their course materials.

1. *Transforming AR(p) to MA.* If a p -order autoregressive process $\phi(L)y_t = \varepsilon_t$ is stationary, with moving average representation $y_t = \psi(L)\varepsilon_t$, show that

$$0 = \sum_{j=0}^{p} \phi_j \psi_{k-j} = \phi(L)\psi_k, \quad k = p, p+1, \ldots$$

i.e., show that the moving average coefficients satisfy the autoregressive difference equation.

2. *Sims' formula for spectrum.* Assume that we have a sample $\{y_t, x_t\}_{t=1}^{T}$ from infinite distributed lag model $y_t = B(L)x_t + e_t$, $B(L) = \sum_{j=1}^{\infty} b_j L^j$ with absolutely summable coefficients $\sum |b_j| < \infty$ (here e_t is a white-noise, x_t is stationary and weakly exogenous). Assume that one estimates (misspecified) model with q lags, that is, he regresses y_t on to $x_{t-1}, ..., x_{t-q-1}$ and obtains $\widehat{a}_1, ..., \widehat{a}_q$. As the sample size increases to infinity (but q is kept constant), the estimated coefficients converge to some non-random limits: $\widehat{a}_j \xrightarrow{p} a_j$. Let $A(L) = a_1 L + ... + a_p L^p$. Show that $A(\cdot)$ is a solution to the following problem:

$$\min_{a_1,\ldots,a_q} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(A(e^{-i\omega}) - B(e^{-i\omega})\right) S_X(\omega) \left(A(e^{i\omega}) - B(e^{i\omega})\right),$$

where $S_X(\cdot)$ is the spectrum of the process x_t. That is, one minimizes the quadratic form in the differences between true and estimated polynomial, assigning the greatest weights to the frequencies for which spectral density is the greatest.

3. *Spectrum and filters.* This is your first empirical problem. Choose a software package you feel comfortable using (I would recommend MatLab).You may use any users-written codes you find on Internet. Always make sure that the code is doing what you think it is doing. Please, do not forget to cite whatever you are using.

   (i) Download quarterly values of Real GDP for the US from Mark W. Watson personal web-site (*you may use any other aggregate macro time series from any other source if you wish. Economic Database (FRED II) maintained by the Federal Reserve Bank of St. Louis is a fantastic source*).

   (ii) Define the growth rate for real GDP. Estimate and plot spectrum for the growth rate. Discuss the graph. Find which peak in the spectrum corresponds to business cycles.

   (iii) Use the following three cycle removing devices: a) run the OLS to detrend the series ; b) use Prescott-Hodrick filter; c) apply Baxter-King filter.

   (iv) Re-estimate spectrum for all series after applying each of the three procedures. Draw spectrum functions. Discuss the differences.

   Note. As in real life empirical research, you will need to make a lot of choices while performing the task, such as choosing lag length, kernel function, and so on. Try to be reasonable, always check whether you results are sensitive to these choices. Also check original papers for suggestions.

4. *Factor model and Principle components.* Let $X = (X_{it})_{i=1,...,N,t=1,...,T}$ is $T \times N$ matrix of observations. The matrices $\Lambda(N \times k)$ and $F(T \times k)$ are both unknown.

   Factor model could be written as an N-dimension time series with T observations:

$$X_t = \Lambda F_t + e_t,$$

   where $X_t = (X_{1t}, X_{2t}, ..., X_{Nt})'$, $\Lambda = (\lambda_1, ..., \lambda_N)'$ and $e_t = (e_{1t}, e_{2t}, ..., e_{Nt})'$. Alternatively, it can be written as a T-dimension system with N observations:

$$X_i = F\lambda_i + e_i,$$

   where $X_i = (X_{i1}, ..., X_{iT})'$, $F = (F_1, ..., F_T)'$, and $e_i = (e_{i1}, ..., e_{iT})'$.

   The method of principle components minimizes

$$V(k) = \min_{\Lambda,F} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (X_{it} - \lambda_i' F_t)^2$$

   (a) Write down the first order condition for minimization over Λ. Concentrate out Λ.

(b) Assume the normalization $F'F/T = I_k$. Show that minimization problem is equivalent to maximizing trace of $F'(XX')F$.

(c) Argue that F consists of the linear subspace containing k eigenvector corresponding to the k largest eigenvalues.

(d) What are the estimates of factor loadings and common component?

5. *Subsampling with of nearly unit root process.* Assume that you have a sample $\{x_1, ..., x_T\}$ of size T from an AR(1) process with the autoregressive coefficient $0 < \rho \leq 1$. The goal is to construct an asymptotic confidence set for ρ. Subsampling (Romano and Wolf, *Econometrica* 2001) is the following procedure.

Step 1. Regress x_t on its lag and calculate the OLS estimate of $\widehat{\rho}$ and variance $\widehat{\sigma}^2_{\widehat{\rho}}$.

Step 2 Choose a subsample size $b_T < T$ and let b be an index changing between 1 and $T - b_T$. For each value of b consider a subsample $Z_b = \{x_b, x_{b+1}, ..., x_{b+b_T}\}$ of the size b_T from the initial sample. For each block Z_b run OLS regression to get the t-statistics, $t_b = \frac{\widehat{\rho}_b - \widehat{\rho}}{\widehat{\sigma}_{\widehat{\rho}_b}}$.

Step 3 Order statistics $\{t_b\}_{b=1}^{T-b_T}$ in ascending order and get $\alpha/2$ and $1-\alpha/2$ quantiles (q_1 and q_2) of this distribution. The confidence set for ρ is $C(x) = [\widehat{\rho} - q_2\widehat{\sigma}_{\widehat{\rho}}, \widehat{\rho} - q_1\widehat{\sigma}_{\widehat{\rho}}]$.

The purpose of this problem is to understand the asymptotic coverage of the described procedure. Let $t(T, \rho_0) = \frac{\widehat{\rho} - \rho_0}{\widehat{\sigma}_{\widehat{\rho}}}$ be the t-statistics for testing $H_0 : \rho = \rho_0$ with the full sample. The described procedure uses an approximation of unknown distribution of $t(T, \rho_0)$ by the distribution of t-statistic in subsample $t_b(b_T) = \frac{\widehat{\rho}_b - \widehat{\rho}}{\widehat{\sigma}_{\widehat{\rho}_b}}$, here $\widehat{\rho}_b$ is calculated for a subsample of size b_T, and $\widehat{\rho}$ - for the whole sample. The distribution of $t_b(b_T)$ could be simulated (it is done in Step 3). Assume known that simulated in Step 3 distribution (quantiles) of $t_b(b_T)$ are uniformly close(converge) to the theoretical distribution(quantiles) of $t_b(b_T)$. Assume that $b_T \to \infty$ and $b_T/T \to 0$ as $T \to \infty$.

(a) Let the true value $0 < \rho_0 < 1$ be fixed while $T \to \infty$. What is the limiting distribution of $t(T, \rho_0)$? What is the limiting distribution of $t_b(b_T)$? Calculate the limiting coverage $\lim_{T\to\infty} P_{\rho_0}\{\rho_0 \in C(x)\}$.

(b) Now assume that we have a unit root, that is, $\rho_0 = 1$. What is the limiting distribution of $t(T, \rho_0 = 1)$? What is the limiting distribution of $t_b(b_T)$? Calculate the limiting coverage $\lim_{T\to\infty} P_{\rho_0=1}\{1 \in C(x)\}$.

For the next steps use the following statement. Assume that $\rho_T = 1 + c_T/T$.

- If $c_T \to 0$ as $T \to \infty$, then $t(T, \rho_T) \Rightarrow \frac{\int_0^1 w(t)dw(t)}{\sqrt{\int_0^1 w^2(t)dt}}$.
- If $c_T \to -\infty$ as $T \to \infty$, then $t(T, \rho_T) \Rightarrow N(0, 1)$.

(c) Let $\rho_T = 1 + c/T, c < 0$. What is the limiting distribution of $t(T, \rho_T)$? What is the limiting distribution of $t_b(b_T)$? What can you say about the limiting coverage $\lim_{T\to\infty} P_{\rho_T}\{\rho_T \in C(x)\}$? What is the intuition of your result?

(d) Let $\rho_T = 1 + c/b_T, c < 0$. What is the limiting distribution of $t(T, \rho_T)$? What is the limiting distribution of $t_b(b_T)$? What can you say about the limiting coverage $\lim_{T\to\infty} P_{\rho_T}\{\rho_T \in C(x)\}$? What is the intuition of your result?

(e) Is the subsampling interval uniformly asymptotically correct? Explain.

6. *Empirical exercise. PPP puzzle.* Purchasing power parity (PPP) is "an empirical proposition that, once converted to a common currency, national price levels should be equal" (Rogoff, 1996). Even though almost no one believes in absolute PPP, most think that the real exchange rate tend toward PPP in the very long run. Main puzzle, however, is that the speed of convergence (measured as a half-life of a shock and estimated to be between three to five years) is too slow to be explained by nominal rigidities.

This exercise is aimed to answer two questions: 1) is there any long-run convergence of PPP (rephrase: does real exchange rate possess a unit root); 2) what is the half-life of real exchange rate?

Take any currency and calculate a time series for log of real exchange rate (exchange rates and CPI for various countries provided on the course webpage). I call it x_t.

(a) Test whether x_t has a unit root. Use augmented Dickey-Fuller (with lag length chosen according BIC) and Phillips-Perron test. Do this in two versions: including constant and including a linear trend. Do the data show evidence of a unit root?

(b) Use Stock (Journal of Monetary Economics, 1991) method to construct a confidence interval for the local to unity parameter. Assume that x_t is AR(1) with a constant and use local to unity approximation for t-statistic. Tables are reported in Stock(1991).

(c) Assume that $x_t = c + \rho x_{t-1} + e_t$. The half-life is defined as a time needed for half of the shock to die. That is half-life $= \frac{-\log 2}{\log \rho}$. Transform the confidence set you received in (b) to a confidence interval for the half-life of deviations from PPP. How persistent are the shocks?

7. *Simulated GMM.* Suppose we wish to estimate an MA(2) process

$$y_t = \mu + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2}$$

where the e_t are iid $N(0, \sigma^2)$ random variables. Although estimation is possible using ML, explain how you could estimate the parameters of the model using indirect inference. Also indicate how the models specification can be tested using indirect inference.

8. *Kalman filter of a long-run trend.* Consider a model of a constant long-run trend

$$\alpha_t = \alpha_{t-1}, \quad y_t = \alpha_t + v_t, \quad v_t \sim i.i.d. N(0, \sigma^2).$$

(a) Write down Kalman filter for the model with starting values $\alpha_{1|0} = a$ and $P_{0|1} = p_0$, then

$$P_{t|t} = \frac{p_0}{1 + tp_0/\sigma^2}.$$

Show also that the contribution of each additional observation to $\alpha_{T|T}$ is negligible as T increases.

(b) Show that as time horizon increases Kalman filter converges to a value independent of a and p_0. What is this value?

(c) What is the value of $\alpha_{T|T}$ if we have uninformative prior $p_0 \to \infty$?

Readings:

Christiano, Eichenbaum and Vigfusson (2004) "What happens after a technology shock?" unpublished manuscript.

Gali " Technology, Employment and the Business Cycle: Do technology shocks explain aggregate fluctuations?", AER 1999.

Romano J.P. and Wolf M.(2001): "Subsampling Intervals in Autoregressive Models with Linear Time Trend," *Econometrica*, 69(5), 1283-1314.

Stock, J. (1991). "Confidence intervals for the largest autoregressive root in US macroeconomic time series," Journal of Monetary Economics 28, 435-459.