14.384 Time Series Analysis, Fall 2007
Professor Anna Mikusheva
Paul Schrimpf, scribe
September 6, 2007

Lecture 1
revised on September 9, 2009

Stationarity, Lag Operator, ARMA, and Covariance Structure

Introduction

History – popular in early 90s, making comeback now.

The main difference between time series econometrics and cross-section is in dependence structure. Cross-section econometrics mainly deals with i.i.d. observations, while in time series each new arriving observation is stochastically depending on the previously observed. The dependence is our best friend and a great enemy. On one side, the dependence screw up your inferences: the regular (ordinary) Central Limit Theorem should be corrected to hold for dependent observations. That bring us to the task of correcting our procedures for dependence. On the other side, the dependence allow us to do more by exploiting it. For example, we can make forecasts (which are almost non-sense in cross-section).

Some topics may sounds counter-intuitive for you at first (if you are cross-section minded). For example, if you are working with very persistent time series, your estimates can be severely biased even if the exclusion restriction is satisfied. But time-to-time you can recover a stochastic trend super-consistently even when the exclusion restriction is not exactly satisfied. I personally find it amusing:)

Can roughly divide time series into macro and finance related stuff. Macro Time series mostly focuses on means. Macro limited by small number of observations available over long horizon. A typical data set has at best 20 years of monthly or 40 years of quarterly data, which sum up to less than 300 observations. This allows us to study linear relations between variables or model means. Financial data usually high-frequency over short period of time. This allows us to model volatility and higher moments.

Outline

Can divide course into two main parts:

1. Classics

		stationary	nonstationary
	Univariate	ARMA	unit root
	Multivariate	VARMA	cointegration

2. DSGE
   - simulated GMM
   - ML
   - Bayesian

Goals

The main objective of this course is to develop the skills needed to do empirical research in fields operating with time series data sets. The course aims to provide students with techniques and receipts for estimation and assessment of quality of economic models with time series data. Special attention will be placed on limitations and pitfalls of different methods and their potential fixes. The course will also emphasize recent developments in Time Series Analysis and will present some open questions and areas of ongoing research.

Problem Sets

Will have an empirical part – requires programming. Use whatever language you prefer. I recommend Matlab and discourage Stata. There will be a session devoted to Intro to MatLab. You need not write your programs from scratch. You can freely download programs from the web, but make sure you use them correctly and cite them. Working in groups is encouraged, but you should write your own solutions.

The final exam will be in a take-home format.

ARMA Processes

Stationarity

We need what we have observed to be stable, in some sense, so that we can make inferences and statements about the future.

Definition 1. *White noise* is a sequence of random variables $\{e_t\}$ such that $Ee_t = 0$, $Ee_te_s = 0$, $Ee_t^2 = \sigma^2$

Definition 2. A process, $\{y_t\}$, is *strictly stationary* if for each k, t and n, the distribution of $\{y_t, ..., y_{t+k}\}$ is the same as the distribution of $\{y_{t+n}, ..., y_{t+k+n}\}$

Definition 3. $\{y_t\}$, is *2nd order stationary*(or *weakly stationary*, or simply *stationary*) if Ey_t, Ey_t^2, and $\text{cov}(y_t, y_{t+k})$ do not depend on t

Examples of non-stationary

Example 4. Break:

$$y_t = \begin{cases} \beta + e_t & t \leq k \\ \beta + \lambda + e_t & t > k \end{cases}$$

Example 5. Random Walk (also known as unit root process)

$$y_t = y_{t-1} + e_t$$

One can notice that

$$Var(y_t) = Var(y_{t-1}) + Var(e_t) > Var(y_{t-1})$$

Lag operator and operations with it

Definition 6. *Lag operator* Denoted L. $Ly_t = y_{t-1}$.

1) The lag operator can be raised to powers, e.g. $L^2y_t = y_{t-2}$. We can also form polynomials of it

$$a(L) = a_0 + a_1L + a_2L^2 + ... + a_pL^p$$
$$a(L)y_t = a_0y_t + a_1y_{t-1} + a_2y_{t-2} + ... + a_py_{t-p}$$

3) Lag polynomials can be multiplied. Multiplication is commutative, $a(L)b(L) = b(L)a(L)$.

4) Some lag polynomials can be inverted. We define $(1 - \rho L)^{-1}$ by the following equality:

$$(1 - \rho L)(1 - \rho L)^{-1} \equiv 1 \tag{1}$$

The statement is: If $|\rho| < 1$, then

$$(1 - \rho L)^{-1} = \sum_{i=0}^{\infty} \rho^i L^i.$$

First we need to check that the infinite sum is a valid definition of an operator from weakly stationary sequences (in $\mathcal{L}^2$ sense) to weakly stationary sequences. This mean that if x_t is weakly stationary, then for all t a sequence $(\sum_{i=0}^J \rho^i L^i)x_t$ converges to a limit in $\mathcal{L}^2$ space as $J \to \infty$ and the result is a weakly stationary sequence. If $|\rho| < 1$, one can check that $y_{t,J} = (\sum_{i=0}^J \rho^i L^i)x_t$ is a Cauchy sequence in $\mathcal{L}^2$. That is, $E(y_{t,j} - y_{t,J})^2 \to 0$ as $\min(j, J) \to \infty$. Given that $\mathcal{L}^2$ is a complete space (all Cauchy sequences converge), there exists a limit, call it z_t: $y_{t,J} \to^{\mathcal{L}^2} z_t$. Then we have to check that z_t is weakly stationary(which is trivial).

Now we have to check the equality (1):

$$\begin{aligned}(1 - \rho L)\sum_{i=0}^{\infty} \rho^i L^i =& \sum_{i=0}^{\infty} \rho^i L^i - \sum_{i=1}^{\infty} \rho^i L^i \\ =& \rho^0 L^0 = 1\end{aligned}$$

For higher order polynomials, we can invert them by factoring, using the formula for $(1-\rho L)^{-1}$ (assuming that the roots are outside the unit circle), and then rearranging, for example:

$$\begin{aligned}1 - a_1 L - a_2 L^2 =& (1 - \lambda_1 L)(1 - \lambda_2 L) \ , \ |\lambda_i| < 1 \\ (1 - a_1 L - a_2 L^2)^{-1} =& (1 - \lambda_1 L)^{-1}(1 - \lambda_2 L)^{-1} \\ =& (\sum_{i=0}^{\infty} \lambda_1^i L^i)(\sum_{i=0}^{\infty} \lambda_2^i L^i) \\ =& \sum_{j=0}^{\infty} L^j \left(\sum_{k=0}^{j} \lambda_1^k \lambda_2^{j-k} \right)\end{aligned}$$

Another (perhaps more easy) way to approach the same problem is do a partial fraction decomposition

$$\begin{aligned}\frac{1}{(1 - \lambda_1 x)(1 - \lambda_2 x)} =& \frac{a}{1 - \lambda_1 x} + \frac{b}{1 - \lambda_2 x} \\ & \Rightarrow \\ a =& \frac{\lambda_1}{\lambda_1 - \lambda_2} \ , \ b = \frac{\lambda_2}{\lambda_2 - \lambda_1} \\ a^{-1}(L) =& a \sum_{i=0}^{\infty} \lambda_1^i L^i + b \sum_{i=0}^{\infty} \lambda_2^i L^i\end{aligned}$$

This trick only works when the λ_i are unique. The formula is slightly different otherwise.

Note: the λ_i are the inverse of the roots of the lag polynomials. To invert a polynomial, we needed $|\lambda_i| < 1$, i.e., the roots of the polynomial are outside of unit circle.

Simple Processes

Autoregressive (AR)

$$\begin{aligned}AR(1)\colon\ & y_t = \rho y_{t-1} + e_t \ , \ |\rho| < 1 \\ & (1 - \rho L) y_t = e_t \\ AR(p)\colon\ & a(L) y_t = e_t, \text{ where a(L) is order } p\end{aligned}$$

Moving average (MA)

$$\begin{aligned}MA(1)\colon\ & y_t = e_t + \theta e_{t-1} \\ & y_t = (1 + \theta L) e_t \\ MA(q)\colon\ & y_t = b(L) e_t, \text{ where b(L) is order } q\end{aligned}$$

ARMA

$$ARMA(p,q)\colon\ a(L)y_t = b(L)e_t,$$

where $a(L)$ is order p and $b(L)$ is order q, and $a(L)$ and $b(L)$ are relatively prime.

An ARMA representation is not unique. For example, an $AR(1)$ (with $|\rho| < 1$) is equal to an $MA(\infty)$, as we saw above. We can see it from the formula for inversion:

$$y_t = \sum_{j=0}^{\infty} \rho^j e_{t-j}$$

Aside: you can also get this formula by repeatedly using the definition of AR(1):

$$y_t = \rho y_{t-1} + e_t = \rho(\rho y_{t-2} + e_{t-1}) + e_t =$$
$$... = \sum_{j=0}^{k-1} \rho^j e_{t-j} + \rho^k y_{t-k}$$

and noticing that $\rho^k y_{t-k} \to^{\mathcal{L}^2} 0$ as $k \to \infty$ In fact, this is more generally true. Any $AR(p)$ with roots outside the unit circle has an MA representation. These processes are called stationary (because there is a weakly stationary version of them).

Any MA process with roots outside unit circle can be written as $AR(\infty)$, such processes called *invertible*. If $y_t = b(L)e_t$ ia an invertible MA process, then $e_t = b(L)^{-1}y_t$. That is, the "errors" are laying in a space of observations and can be recovered from y's (another name for this: errors are *fundamental*).

Covariance structure

Definition 7. *auto-covariance* $\gamma_k \equiv \text{cov}(y_t, y_{t+k})$

Definition 8. *auto-correlation* $\rho_k \equiv \frac{\gamma_k}{\gamma_0}$

$AR(1)$ **example**

$$y_t = \rho y_{t-1} + e_t$$

Observe $Var(y_t) = \rho^2 Var(y_{t-1}) + \sigma^2$, and $Var(y_t) = Var(y_{t-1}) = \gamma_0$, so $\gamma_0 = \frac{\sigma^2}{1-\rho^2}$. Also, it is easy to see by induction that $\gamma_k = \frac{\rho^k \sigma^2}{1-\rho^2}$.

Another way to see this is from the MA representation:

$$y_t = \sum_{i=0}^{\infty} \rho^i e_i$$
$$\Rightarrow$$
$$\gamma_0 = \sum \rho^{2i}\sigma^2 = \frac{\sigma^2}{1-\rho^2}$$
$$\gamma_k = cov(\sum_{i=0}^{\infty} \rho^i e_{t-i}, \sum_{j=0}^{\infty} \rho^j e_{t+k-j}) = \sum_{i,j: i=j-k}^{\infty} \rho^{i+j}\sigma^2 = \frac{\rho^k \sigma^2}{1-\rho^2}$$

More generally, if $y_t = \sum_{i=0}^{\infty} c_i e_{t-i}$ then

$$\begin{aligned}\operatorname{cov}(y_t, y_{t+k}) =& \operatorname{cov}\left(\sum_{i=0}^{\infty} c_i e_{t-i}, \sum_{i=0}^{\infty} c_i e_{t+k-i}\right) \\ =& \sigma^2 \sum_{j=0}^{\infty} c_j c_{j+k}\end{aligned}$$

MA representation and covariance stationarity $y_t = \sum_{i=0}^{\infty} c_i e_{t-i}$ so y_t has finite variance, and in fact is covariance stationary, if $\sum_{j=0}^{\infty} c_j^2 < \infty$. It is often easier to prove things with the stronger assumption of absolute summability, $\sum_{j=0}^{\infty} |c_j| < \infty$ (or stronger still $\sum_{j=0}^{\infty} j|c_j| < \infty$).

Covariance function

Definition 9. *covariance function* $\gamma(\xi) = \sum_{i=-\infty}^{\infty} \gamma_i \xi^i$, where ξ is a complex number.

Lemma 10. Covariance function of MA *For an MA,* $y_t = c(L)e_t$, $\gamma(\xi) = \sigma^2 c(\xi) c(\xi^{-1})$.

Proof.

$$\begin{aligned}c(\xi)c(\xi^{-1}) =& (\sum_{i=0}^{\infty} c_i \xi^i)(\sum_{i=0}^{\infty} c_i \xi^{-i}) \\ =& \sum_{j,l=0}^{\infty} c_j c_l \xi^{j-l} \\ =& \sum_{k=-\infty}^{\infty} \xi^k \sum_{j=0}^{\infty} c_j c_{j+k}\end{aligned}$$

□

Lemma 11. Covariance function of ARMA *For an ARMA,* $a(L)y_t = b(L)e_t$, $\gamma(\xi) = \sigma^2 \frac{b(\xi)b(\xi^{-1})}{a(\xi)a(\xi^{-1})}$

The last statement will be very helpful for spectrum.