Economics 303: Advanced Macroeconomics I

Tuesday-Fridays 10:30 - Noon

Chancellor's Suite

George J. Hall

Brandeis University

Outline

- Our Strategy for Modeling Macroeconomics
- A First Look at the Data
- Stochastic Linear Difference Equations
  - ARMA Models
  - Matrix or State Space Notation

What is Macroeconomics?

- It is the study of aggregate production, consumption, investment and prices in large economic systems comprised of many participants (whom we will call agents).

Why Study Macroeconomics?

- The questions are too big and too important to ignore.

Questions

- Why do some countries have so much (and are such wonderful places to live) while other countries have so little (and are such wretched places to live)?
- Why do some economies grow at 3% per year, while other grow at 8% per year and still others grow at 0.3% per year?
- Why do recessions (or short-term reductions in output) occur in countries that exhibit persistent growth?
- What is the cause of moderate and high inflation?

Policy

- What is the likely impact of stabilization policies on economic activity?
- What is the best approach to monetary policy in moderate inflation economies?

The Scientific Method

1. Write down a model.
2. Solve the model for a refutable hypothesis.
3. Take the model to data.
4. If the model is consistent with the data, accept the hypothesis and keep the model. If not, reject the hypothesis and discard the model.

QUESTION: But what if we can reject any model?

Strategy we will follow in this course

1. State a set of facts about the way an economic system works.
2. Construct a model economy. That is "a theory."
3. Solve the model. Simulate the model.
4. Compare the implications of the model to the facts we wish to understand.
5. Use the model to address questions (e.g. policy experiments) we are interested in.

The better the model does in step 4, the more likely we are to trust the answers the model provides in step 5.

Theory as Data Reduction

- A good theory compresses data with minimal loss of information.
- An example from the physical sciences (from C. Sims).
  - Tycho Brahe accumulated large amount of reliable data on the movements of the planets. In doing so he could accurately predict the motion of the planets.
  - But then Kepler came along and notices that planets move on elliptical orbits with the sun at a focus, allowing the movements of the planets to be described by a small set of coordinates.
  - Newton then found the inverse-square law, thus developing a formula that compressed the data further and gave predictions to the paths of new planets.

Some Facts

- Most data in macro come in the form of time-series.
  - a set of repeated observations of the same variable.

$$y_1, y_2, ..., y_T$$

  - GDP, unemployment, inflation, ...
  - the subscripts denote time.
- To describe the "basic facts" macroeconomics wishes to understand, we must summarize (or compress) the relevant information contained in macro time-series.
- Today we will look at industrial production (IP) and real gross domestic product (GDP).

Some Points About the Data

- Substantial Growth in the U.S. economy over a long period of time.
- For industrial production, the growth rate is 3.8 percent per year (1921-2006). IP in 2006 was 25 times larger than it was in 1921.
- For real gross domestic product, the growth rate is 3.3 percent per year (1947-2006).
- Population growth has been about 1.4 per year. Thus output per person has been growing about 1.9 percent.
- "Rule of 72" Given a growth rate, you can find the number of years it takes for income to double by dividing 72 by the growth rate.

$$\frac{72}{1.9} = 38$$

Output per person doubles every generation.

Trends versus Cycles

- The traditional business cycle approach characterizes the economy as growing along a smooth trend from which it is disturbed by cyclical fluctuations.
  - Increases in productivity move the trend, temporary shocks (e.g. demand, money) push the economy temporarily away from the trend.
  - Deviations from the trend are stationary.
  - Deviations are persistent.
- But why should changes in productivity lead to smooth changes in output?
  - Perhaps shocks are permanent.
  - Output modeled as a non-stationary process
  - First-differences of output are stationary.
- One model for the long run, with another for the short-run?

Stochastic Line Difference Equations

- Inputs: shocks or innovations (e.g. productivity, demand, ...)
- Output: macro time series (e.g. production, inflation, employment, ...)
- Output is related to input via stochastic linear difference equation.
- Simple example.
  - one input series: $w_1,\ w_2,\ w_3,\ ...$
  - one output series: $x_1,\ x_2,\ x_3,\ ...$
  - one initial condition: x_0

Let:

$$x_{t+1} = 0.95x_t + w_{t+1}$$

where the w_{t+1}'s take the value 0.5 if a coin flip shows heads and -0.5 if a coin flip shows tails.

Sample Paths

- Three sample paths.
- The three sample paths differ because of differences in the outcomes of the coin tosses
- Yet there are characteristics that are common to all (most?) sample paths generated by this process.
- Note that any x_t is a weighted sum of the initial condition x_0 and all the past shocks $(w_1,\ w_2, ..., w_t)$.

ARMA Models

- White Noise
- Basic ARMA models
- Lag operators

White Noise

- Building block for linear time-series models.
- The w_ts follow a white noise process if the $w_t \sim i.i.d. N(0, \sigma_w^2)$.
- The normality assumption can be relaxed.

Properties of White Noise

- Unpredictable (1)

$$\begin{aligned} E(w_{t+1}|J_t) &= E(w_{t+1}|\text{all information known at date } t) \\ &= E(w_{t+1}|w_t, w_{t-1}, w_{t-2}, ...) \\ &= 0. \end{aligned}$$

- Unpredictable (2)

$$E(w_t w_{t-j}) = \text{Cov}(w_t w_{t-j}) = 0$$

- Conditional homoskedasticity

$$\begin{aligned} \text{Var}(w_{t+1}|J_t) &= \text{Var}(w_{t+1}|\text{all information known at date } t) \\ &= \text{Var}(w_{t+1}|w_t, w_{t-1}, ...) \\ &= \sigma_w^2 \end{aligned}$$

- Normality is not essential
- Property 1 is essential: martingale difference property.

Basic ARMA Models

We can use white noise w_t to build more interesting models

AR(1): $x_{t+1} = \phi x_t + w_{t+1}$

MA(1): $x_{t+1} = w_{t+1} + \theta w_t$

ARMA(1,1): $x_{t+1} = \phi_0 x_t + w_{t+1} + \theta_0 w_t$

AR(p): $x_{t+1} = \Sigma_{k=0}^{p-1} \phi_k x_{t-k} + w_{t+1}$

MA(q): $x_{t+1} = w_{t+1} + \Sigma_{k=0}^{q-1} \theta_k w_{t-k}$

ARMA(p,q): $x_{t+1} = \Sigma_{k=0}^{p-1} \phi_k x_{t-k} + w_{t+1} + \Sigma_{k=0}^{q-1} \theta_k w_{t-k}$

Basic ARMA Models

- All of them correspond to stochastic difference equations where x_t is a linear combination of past values of x's and current and past values of the shocks (or innovations).
- All these models have mean zero, they are used to represent the deviations from the mean value of x (call it $\bar{x}$) or more generally, the deviation from some deterministic trend.

Basic ARMA Models

- For example, if the deviations of x_t from its mean follow an AR(1):

$$x_{t+1} - \bar{x} = \phi(x_t - \bar{x}) + w_{t+1}$$

and we have

$$x_{t+1} = (1 - \phi)\bar{x} + \phi x_t + w_{t+1}$$

- Or if the deviation from a linear time trend follows an AR(1)

$$x_t - (a + bt) = \phi(x_{t-1} - a - b(t-1)) + w_t$$

and we have

$$x_t = c + dt + \phi x_{t-1} + w_t$$

with $c = (1 - \phi)a + \phi b$ and $d = (1 - \phi)b$.

- Constants absorb means. Linear time trends absorb linear trends, ... In what follows we ignore any mean or deterministic trend.

Lag operators and polynomials

- Lag operator moves the time-index back one unit:

$$Lx_t \equiv x_{t-1}$$

- Hence

$$L^2 x_t \equiv L(Lx_t) = Lx_{t-1} = x_{t-2}$$

- In general, for $j = 0, 1, 2, 3, ...$

$$\begin{aligned} L^j &= x_{t-j} \\ L^{-j} &= x_{t+j} \end{aligned}$$

- For example, we can rewrite an AR(1) process as

$$(1 - \phi L)x_t = w_t$$

and defining the lag-polynomial $a(L) \equiv 1 - \phi L$ we have

$$a(L)x_t = w_t$$

Representing an AR(1) as an MA(∞)

- Assume x_t follows an AR(1)

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

- Applying this expression recursively:

$$\begin{aligned} x_t &= \phi(\phi x_{t-2} + w_{t-1}) + w_t \\ &= \phi^2 x_{t-2} + \phi w_{t-1} + w_t \\ &\vdots \\ &= \phi^k x_{t-k} + \phi^{k-1} w_{t+k-1} + \ldots + \phi w_{t-1} + w_t \end{aligned}$$

- If $|\phi| < 1$, so that $\lim_{k \to \infty} \phi^k x_{t-k} = 0$, it follows that

$$x_t = \sum_{k \geq 0} \phi^k w_{t-k}$$

- Thus an AR(1) can be represented as an MA(∞)

Representing an AR(1) as an MA(∞)

- Next we repeat the derivation using lag-polynomials.
- To do this we use the following identity for lag polynomials:

$$\frac{1}{1-\phi L} = \sum_{k\geq 0} \phi^k L^k$$

- This identity can be proved formally and is analogous to the well known geometric series expression for real (and complex) numbers z that satisfy $|z| < 1$:

$$\frac{1}{1-z} = \sum_{k\geq 0} z^k$$

Representing an AR(1) as an MA(∞)

- From

$$(1-\phi L)x_t = w_t$$

  we have

$$\begin{aligned} x_t &= \frac{1}{1-\phi L} w_t \\ &= \sum_{k\geq 0} \phi^k L^k w_t \\ &= \sum_{k\geq 0} \phi^k w_{t-k}. \end{aligned}$$

- The condition that $|\phi < 1$ is important to apply the above "trick."

Matrix Representation

- We can write any ARMA model in matrix notation

$$x_{t+1} = Ax_t + Cw_{t+1}$$

where

$$\begin{aligned} x_t &= n \times 1 \text{ vector. We will call it the ``state.''} \\ A &= n \times n \text{ matrix of coefficients} \\ C &= n \times m \text{ matrix of coefficients} \\ w_{t+1} &= m \times 1 \text{ vector of shocks} \end{aligned}$$

- We assume

1. w_{t+1} is an i.i.d. process satisfying $w_{t+1} \sim N(0, I)$
2. $E(w_{t+j}|J_t) = 0$ and $Ew_{t+1}w_{t+1} = I$.

Examples

- Second-order auto-regression

$$z_{t+1} = \alpha + \phi_1 z_t + \phi_2 z_{t-1} + \epsilon_t$$

where $\epsilon_t \sim N(0, \sigma_\epsilon^2)$

can be written as:

$$\begin{bmatrix} 1 \\ z_{t+1} \\ z_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \alpha & \phi_1 & \phi_2 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ z_t \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} 0 \\ \sigma_\epsilon \\ 0 \end{bmatrix} [w_{t+1}]$$

Examples

- First-order moving average

$$z_{t+1} = w_{t+1} + \phi w_t$$

where $w_{t+1} \sim N(0,1)$

can be written as:

$$\begin{bmatrix} z_{t+1} \\ w_{t+1} \end{bmatrix} = \begin{bmatrix} 0 & \phi \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_t \\ w_t \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} w_{t+1} \end{bmatrix}$$

Prediction

Note that

$$E_t x_{t+j} = A^j x_t$$

Covariance Stationary

A stochastic process $\{x_t\}$ is said to be *covariance stationary* if it satisfies the following two properties:

1. the mean is independent of time, $Ex_t = Ex_0$ for all t, and
2. the sequence of autocovariance matrices $E(x_{t+j} - Ex_{t+j})(x_t - Ex_t)'$ depend on the separation dates $j = ... - 2, -1, 0, 1, 2...$ but not on t.

Second Moments

- Assume x_0 is drawn from a distribution with mean $\mu_0 = Ex_0$ and covariance $\Sigma_0 = E(x - Ex_0)(x - Ex_0)'$.
- $\mu_{t+1} = A\mu_t$.
- So $x_{t+1} - \mu_{t+1} = A(x_t - \mu_t) + Cw_{t+1}$.
- Thus

$$E(x_{t+1} - \mu)(x_{t+1} - \mu)' = AE(x_t - \mu)(x_t - \mu)'A' + CC'$$

or

$$\Sigma = A\Sigma A' + CC'$$

This type of equation is known as a *discrete Lyapunov* equation.

- Since,

$$(x_{t+j} - \mu_{t+j}) = A^j(x_t - \mu_t) + Cw_{t+j} + ... + A^{j-1}Cw_{t+1}$$

we get

$$E(x_{t+j} - \mu)(x_t - \mu)' = A^j\Sigma$$

Impulse Response Functions

- As we did before, we can represent the x's a linear combinations of the past shocks. Using the lag operator,

$$(I - AL)x_{t+1} = Cw_{t+1}$$

hence we get

$$x_{t+1} = \sum_{k=0}^{\infty} A^k C w_{t+1-k}$$