DATA 5610

Time Series Notes I

Tyler J. Brough

Department of Data Analytics & Information Systems

UtahStateUniversity

Introduction to Time Series I

Beginning Time Series Topics

Most data in economics (espcially in macroeconomics and finance) come in the form of *time series*.

Time Series: a set of repeated observations of the same random variable ordered in time.

- Example: GNP or stock returns
- Also: prices, exchange rates, interest rates, inflation (lots of others)

We can write a time series as $\{x_1, x_2, \ldots, x_T\}$ or simply as $\{x_t\}_{t=1}^{T}$.

We treat x_t as a random variable. Really nothing different from the rest of econometrics. Notice the difference is the subscript t rather than i.

If, for example, a random variable y_t is generated by

$$y_t = x_t\beta + \varepsilon_t$$

in which $E(y_t|x_t) = 0$

Then OLS provides a consistent estimate for β (just as if the subscript were "i" instead of "t").

The phrase "time series" is used to denote:

1. a sample $\{x_t\}$ such as IBM stock price from Jan. 1, 2010 to Dec. 31, 2010.
2. A probability model for that sample. i.e. a statement about the joint distribution of the random variables $\{x_t\}$.

A first model for the joint distribution of a time series $\{x_t\}$ is:

$$x_t = \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2)$$

i.e. x_t is normal and independent over time.

Typically, time series are not iid, which is what makes them interesting.

Ex: unusually high inflation today is likely to lead to unusually high inflation tomorrow.

The building block for our time series models is the ***white noise process***

$$\varepsilon_t \sim \text{ iid } N(0, \sigma_\varepsilon^2)$$

Note three implications:

1. $E(\varepsilon_t) = E(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) = E(\varepsilon_t | \text{ all info at t-1}) = 0$
2. $E(\varepsilon_t \varepsilon_{t-j}) = Cov(\varepsilon_t \varepsilon_{t-j}) = 0$
3. $Var(\varepsilon_t) = Var(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) = Var(\varepsilon_t | \text{ all info at t-1}) = \sigma_{\varepsilon_t^2}$

(1) and (2) are the absence of any serial correlation or predictability.

(3) is conditional homoscedasticity or a constant conditional variance.

By itself ε_t is pretty boring. If ε_t is abnormally high there is no tendency for ε_{t+1} to be high.

More realistic models are constructed by taking combinations of ε_t.

Basic ARMA Models

Most of the time our time series models will be created by taking linear combinations of white noise

- AR(1): $x_t = \phi x_{t-1} + \varepsilon_t$
- MA(1): $x_t = \varepsilon_t + \theta \varepsilon_{t-1}$
- AR(p): $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + \varepsilon_t$
- MA(q): $x_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}$
- ARMA(p,q): $x_t = \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$

Notice that each model is a recipe to generate a sequence $\{x_t\}$ given a sequence of realizations of the white noise process and a starting x_0 value.

All of these models are mean zero, and represent deviations of the series about a mean. For example, if a series has mean $\bar{x}$ and follows an AR(1)

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

is equivalent to

$$\begin{aligned} x_t &= (1 - \phi)\bar{x} + \phi x_{t-1} + \varepsilon_t \\ &= \mu + \phi x_{t-1} + \varepsilon_t \end{aligned}$$

where $\mu = (1 - \phi)\bar{x}$

NB: the constant absorbs the mean

Lag Operators and Polynomials

It is easiest to represent ARMA models in *lag operator* notation. The lag operator moves the index back one time unit:

$$Lx_t = x_{t-1}$$

More formally, L is an operator that takes an original time series $\{x_t\}$ and produces another, which is the same as the original only shifted backwards in time.

From the definition we can do other things:

$$
\begin{aligned}
L^2 x_t &= L(Lx_t) = Lx_{t-1} = x_{t-2} \\
L^j x_t &= x_{t-j} \\
L^{-j} x_t &= x_{t+j}
\end{aligned}
$$

We can also define lag polynomials, e.g.

$$a(L) = (a_0 L + a_1 L^1 + a_2 L^2)x_t = a_0 x_t + a_1 x_{t-1} + a_2 x_{t-2}$$

Using this notation we can rewrite the ARMA models as

- AR(1): $(1 - \phi L)x_t = \varepsilon_t$
- MA(1): $x_t = (1 + \theta L)\varepsilon_t$
- AR(p): $(1 - \phi_1 L - \phi_2 L^2 + \cdots + \phi_p L^p)x_t = \varepsilon_t$
- MA(q): $x_t = (1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q)\varepsilon_t$

ARMA models are not unique. A time series with a given joint distribution of $\{x_0, x_1, \ldots, x_T\}$ can usually be represented with a variety of ARMA models.

It is often convenient to work with different representations:

1. The shortest (or only finite length) polynomial representation is usually the easiest to work with
2. AR forms are the easiest to estimate (since OLS assumptions still apply)
3. MA forms express x_t in terms of a linear combination of independent right hand side variables. Often finding variances and covariances in this form is easiest.

AR(1) to MA(∞) by Recursive Substitution

Start with an AR(1)

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

Recursively substituting

$$x_t = \phi(\phi x_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \phi^2 x_{t-2} + \phi\varepsilon_{t-1} + \varepsilon_t$$
$$x_t = \phi^k x_{t-k} + \phi^{k-1}\varepsilon_{t-k+1} + \cdots + \phi^2\varepsilon_{t-2} + \phi\varepsilon_{t-1} + \varepsilon_t$$

Thus an AR(1) can always be expressed as an ARMA(k,k-1).

Also, if $|\phi| < 1$ so that

$$\lim_{k \to \infty} \phi^k x_{t-k} = 0$$

then

$$x_t = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}$$

AR(1) to MA(∞) with Lag Operators

Starting again with the AR(1) model:

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

The way to "invert" the AR(1) is to write

$$x_t = (1 - \phi L)^{-1} \varepsilon_t$$

What does $(1-\phi L)^{-1}$ mean? We have only defined polynomials in L so far.

We try to use the expression

$$(1-z)^{-1} = 1 + z + z^2 + z^3 + \cdots \quad \text{for} \quad |z| < 1$$

NB: this expression for z can be proven with a Taylor expansion.

Using this expansion and hoping that $|\phi| < 1$ implies $|\phi L| < 1$, suggests

$$x_t = (1-\phi L)^{-1}\varepsilon_t = (1 + \phi L + \phi^2 L^2 + \cdots)\varepsilon_t = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}$$

Note: we can't always perform this inversion. We require $|\phi| < 1$. Not all ARMA processes are invertible to a representation of x_t in terms of current and past ε_t

AR(p) to MA(∞)

Getting to an MA(∞) from an AR(1) is almost as easy either way (recursive substitution or lag operators) but in higher order models lag operators become much easier.

Let's try an AR(2):

$$
\begin{aligned}
x_t &= \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t \\
(1 - \phi_1 L - \phi_2 L^2) x_t &= \varepsilon_t
\end{aligned}
$$

We need to factor $(1 - \phi_1 L - \phi_2 L^2)$ in order to use the $(1 - z)^{-1}$ formula. So find λ_1 and λ_2 such that

$$(1 - \phi_1 L - \phi_2 L^2) = (1 - \lambda_1 L)(1 - \lambda_2 L)$$

The solution is

$$\begin{aligned} \lambda_1 \lambda_2 &= -\phi_2 \\ \lambda_1 + \lambda_2 &= \phi_1 \end{aligned}$$

Some Mathematical Details

$$
\begin{aligned}
(1 - \lambda_1 L)(1 - \lambda_2 L) &= 1 - \lambda_2 L - \lambda_1 L + \lambda_1 \lambda_2 L \\
&= 1 - (\lambda_1 + \lambda_2) L + \lambda_1 \lambda_2 L
\end{aligned}
$$

Now we need to invert

$$(1-\lambda_1 L)(1-\lambda_2 L)x_t = \varepsilon_t$$

Thus

$$\begin{aligned} x_t &= (1-\lambda_1 L)^{-1}(1-\lambda_2)^{-1}x_t = \varepsilon_t \\ x_t &= \left[\sum_{j=0}^{\infty} \lambda_1^j L^j\right]\left[\sum_{j=0}^{\infty} \lambda_2^j L^j\right]\varepsilon_t \end{aligned}$$

Multiplying out the polynomials is tedious but straight forward

$$\begin{aligned} \left[\sum_{j=0}^{\infty} \lambda_1^j L^j\right]\left[\sum_{j=0}^{\infty} \lambda_2^j L^j\right] &= (1+\lambda_1 L+\lambda_2 L^2+\cdots)(1+\lambda_2 L+\lambda_2 L^2+\cdots) \\ &= 1+(\lambda_1+\lambda_2)L+(\lambda_1^2+\lambda_1\lambda_2+\lambda_2^2)L^2+\cdots \\ &= \sum_{j=0}^{\infty}\left(\sum_{k=0}^{j}\lambda_1^k\lambda_2^{j-k}\right)L^j \end{aligned}$$

A nicer way to express an MA(∞) is to use the **partial fractions trick**. Find a and b such that

$$
\begin{aligned}
\frac{1}{(1-\lambda_1 L)(1-\lambda_2 L)} &= \frac{a}{(1-\lambda_1 L)} + \frac{b}{(1-\lambda_2 L)} \\
&= \frac{a(1-\lambda_2 L) + b(1-\lambda_1 L)}{(1-\lambda_1 L)(1-\lambda_2 L)}
\end{aligned}
$$

The right-hand side numerator must equal 1, so

$$
\begin{aligned}
a + b &= 1 \\
a\lambda_2 + b\lambda_1 &= 0
\end{aligned}
$$

The solution is

$$
b = \frac{\lambda_2}{\lambda_2 - \lambda_1}, \quad a = \frac{\lambda_1}{\lambda_1 - \lambda_2}
$$

$$\frac{1}{(1-\lambda_1 L)(1-\lambda_2 L)} = \frac{\lambda_1}{(\lambda_1-\lambda_2)}\frac{1}{(1-\lambda_1 L)} + \frac{\lambda_2}{(\lambda_2-\lambda_1)}\frac{1}{(1-\lambda_2 L)}$$

Thus we can express x_t as

$$\begin{aligned} x_t &= \frac{\lambda_1}{\lambda_1-\lambda_2}\sum_{j=0}^{\infty}\lambda_1^j\varepsilon_{t-j} + \frac{\lambda_2}{\lambda_2-\lambda_1}\sum_{j=0}^{\infty}\lambda_2^j\varepsilon_{t-j} \\ &= \sum_{j=0}^{\infty}\left(\frac{\lambda_1}{\lambda_1-\lambda_2}\lambda_1^j + \frac{\lambda_2}{\lambda_2-\lambda_1}\lambda_2^j\right)\varepsilon_{t-j} \end{aligned}$$

Note: Again not every AR(2) can be inverted. We require that the λ's satisfy $|\lambda| < 1$.

Until explicitly stated we will assume we are working with invertible ARMA models.

MA(q) to AR(∞)

This is now straight forward

$$x_t = b(L)\varepsilon_t$$

has AR(∞) representation

$$b(L)^{-1}x_t = \varepsilon_t$$