Econometrics, Lecture 9A
Univariate Stochastic Processes

Ron Smith
EMS, Birkbeck, University of London

Autumn 2020

Last time

- We looked at diagnostic tests for serial correlation, heteroskedasticity, non-normality, non-linearity and structural breaks.
- Many were Lagrange Multiplier type tests using the residuals from the estimated models.
- For all these tests the null hypothesis was that the model was well specified, no problems, so a p value greater than 0.05, meant the model was OK at the 5% signficance level.
- We have covered the main elements of estimation and inference that we will use, mainly in the context of the LRM. Now we are going to apply these elements to lots of different models.
- Starting with univariate time series models.

Stochastic processes

- We have observations on some economic variable, $y_t, t = 1, 2, ..., T$, e.g. log GDP. Treat each y_t as a random variable with a density function, $f_t(y_t)$ and we observe one realisation from the distribution for that period.
- A family of random variables indexed by time is called a stochastic process, an observed sample is called a realisation of the stochastic process.
- A stochastic process is said to be 'strongly stationary' if its distribution is constant through time, i.e. $f_t(y_t) = f(y_t)$.
- It is first order stationary if it has a constant mean.
- It is second order, or weakly or covariance stationary if also has constant variances and constant covariances between y_t and y_{t-i}.

Correlogram

- Constant autocovariances, $\gamma_i = Cov(y_t, y_{t-i})$ are only a function of i (the distance apart of the observations) not t, the time they are observed.
- Assuming constant mean $E(y_t) = E(y)$, the autocovariances

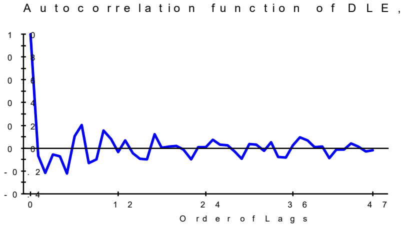
$$\gamma_i = Cov(y_t, y_{t-i}) = E(y_t - E(y))(y_{t-i} - E(y))$$

  then summarise the dependence between the observations,
- They are often represented by the autocorrelation function or correlogram, the vector (graph against i) of the autocorrelations

$$r_i = Cov(y_t, y_{t-i}) / Var(y)$$

  Asuming $Var(y_t)$ is constant
- If the series is stationary, the correlogram converges to zero quickly.

Correlogram of change in log earnings



Order of integration

- The order of integration is the number of times a series must be differenced to make it stationary (after perhaps removing deterministic elements like a linear trend, seasonals).
- So a series, y_t, is said to be Integrated of order zero, I(0), if y_t is stationary; integrated of order one, I(1), if $\Delta y_t = y_t - y_{t-1}$ is stationary; integrated of order two, I(2), if

$$\begin{aligned} \Delta^2 y_t &= \Delta y_t - \Delta y_{t-1} \\ &= (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) \\ &= y_t - 2y_{t-1} + y_{t-2} \end{aligned}$$

  is stationary.
- Notice that $\Delta^2 y_t \neq \Delta_2 y_t = y_t - y_{t-2}$.

Lag operator & White noise processes

In examining dynamics it will be useful to use the Lag Operator, L, sometimes known as the backward shift operator B.

$$
\begin{aligned}
Ly_t &= y_{t-1}; \quad L^2 y_t = y_{t-2}; \quad etc \\
\Delta y_t &= (1-L)y_t, \\
\Delta^2 y_t &= (1-L)^2 y_t.
\end{aligned}
$$

A stochastic process is said to be White Noise if it satisfies our usual assumptions for a well behaved error term

$$
\begin{aligned}
E(\varepsilon_t) &= 0; \\
E(\varepsilon_t^2) &= \sigma^2; \\
E(\varepsilon_t \varepsilon_{t-i}) &= 0, \ i \neq 0
\end{aligned}
$$

We will use ε_t below to denote white noise processes.

Autoregressive processes

A first order (one lag) autoregressive process (AR1) takes the form:

$$\begin{aligned} y_t &= \rho y_{t-1} + \varepsilon_t, \\ y_t - \rho y_{t-1} &= \varepsilon_t \\ y_t(1 - \rho L) &= \varepsilon_t, \end{aligned}$$

then by repeated substitution,

$$\begin{aligned} y_t &= \rho\left(\rho y_{t-2} + \varepsilon_{t-1}\right) + \varepsilon_t \\ y_t &= \rho\left(\rho\left(\rho y_{t-3} + \varepsilon_{t-2}\right) + \varepsilon_{t-1}\right) + \varepsilon_t \end{aligned}$$

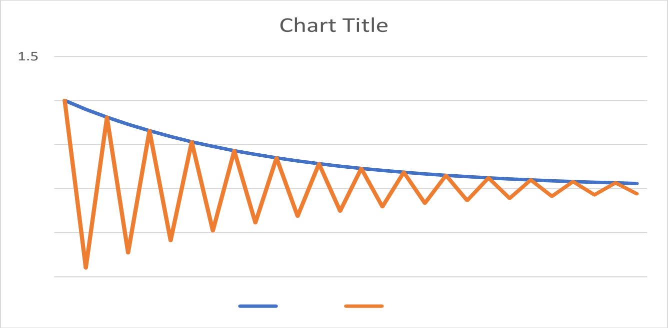
it is stationary if $\mid \rho \mid < 1$, so $\rho^n \rightarrow 0$, giving the sum of a geometric progression:

$$\begin{aligned} y_t &= \varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} + \rho^3\varepsilon_{t-3}.... \qquad (1) \\ y_t &= (1 - \rho L)^{-1}\varepsilon_t, \end{aligned}$$

Since $E(\varepsilon_t) = 0$, then $E(y_t) = 0$.

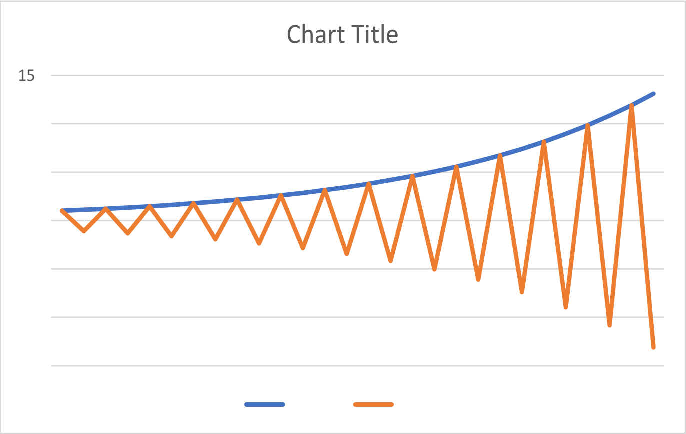
Stable case

AR1 diference equation Y(t)=0.9Y(t-1) and Y(t)=-0.9Y(t-1),



Unstable case

AR1 diference equation Y(t)=1.1Y(t-1) and Y(t)=-1.1Y(t-1),,



AR1 with intercept

$$\begin{aligned} y_t &= \alpha + \rho y_{t-1} + \varepsilon_t, \\ y_t(1-\rho L) &= \alpha + \varepsilon_t, \end{aligned}$$

with $E(y_t) = 0$ then by repeated substitution as before,

$$\begin{aligned} y_t &= \alpha + \rho\left(\alpha + \rho y_{t-2} + \varepsilon_{t-1}\right) + \varepsilon_t \\ y_t &= \alpha + \rho\left(\alpha + \rho\left(\alpha + \rho y_{t-3} + \varepsilon_{t-3}\right) + \varepsilon_{t-1}\right) + \varepsilon_t \end{aligned}$$

and if $\mid \rho \mid < 1$, the geometric progression is:

$$\begin{aligned} y_t &= \alpha + \rho\alpha + \rho^2\alpha + \ldots + \varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} + \rho^3\varepsilon_{t-3}\ldots. \\ y_t &= (1-\rho)^{-1}\alpha + (1-\rho L)^{-1}\varepsilon_t, \end{aligned}$$

$$E\left(y_t\right) = \frac{\alpha}{1-\rho}$$

Variance

Since $Var(y_t) = E(y_t - E(y_t))^2$, with $\mid \rho \mid < 1$, and since:

$$\begin{aligned} y_t - E(y_t) &= \varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} + \rho^3\varepsilon_{t-3} \ldots. \\ E(y_t - E(y_t))^2 &= E\left(\varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} \ldots\right)\left(\varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} \ldots\right) \\ &= E\left(\varepsilon_t^2 + \rho^2\varepsilon_{t-1}^2 + \rho^4\varepsilon_{t-2}^2 \ldots + \varepsilon_t\rho\varepsilon_{t-1} + \ldots\right) \\ &= \sigma^2 + \rho^2\sigma^2 + \rho^4\sigma^2 + \ldots + 0 + \ldots. \end{aligned}$$

sum of a geometric progression in ρ^2 so variance of y_t

$$Var(y_t^2) = \sigma^2/(1-\rho^2).$$

Covariances and autocorrelations

$Cov(y_t y_{t-1}) = E(y_t - E(y_t))(y_{t-1} - E(y_{t-1}))$, but mean is constant so $E(y_t) = E(y_{t-1})$

$$\begin{aligned} Cov(y_t y_{t-1}) &= E\left(\varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2}\ldots\right)\left(\varepsilon_{t-1} + \rho\varepsilon_{t-2} + \rho^2\varepsilon_{t-3}\ldots\right) \\ &= E(\rho\varepsilon_{t-1}^2 + \rho^3\varepsilon_{t-2}^2 + \ldots) \\ &= \rho\sigma^2 + \rho^3\sigma^2 + \rho^3\sigma^2 + .. \\ &= \rho\left(\sigma^2 + \rho^2\sigma^2 + \rho^4\sigma^2 + \ldots\right) \\ &= \rho Var(y_t) \end{aligned}$$

The correlation between y_t and y_{t-1},

$$r_1 = \frac{Cov(y_t y_{t-1})}{Var(y_t)} = \frac{\rho Var(y_t)}{Var(y_t)} = \rho$$

By a similar argument the correlation between y_t and $y_{t-1}$is given by $r_i = \rho^i$ so declines exponentially.

Estimation

- If the process is stationary, and ε_t is not serially correlated the parameters of the AR model can be estimated consistently by Least Squares,
- The estimates will not be unbiased (y_{t-1} is uncorrelated with ε_t but not independent since it is correlated with ε_{t-1});
- The estimate of ρ will be biased downwards but the bias declines with T.
- If ε_t is serially correlated the estimate of ρ will also be inconsistent, (y_{t-1} is correlated with ε_t) the bias does not decline with T

Stability

- A p th order autoregression (ARp) is:

$$y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \ldots + \rho_p y_{t-p} + \varepsilon_t$$

$$y_t - \rho_1 y_{t-1} - \rho_2 y_{t-2} - \ldots - \rho_p y_{t-p} = \varepsilon_t$$

$$(1 - \rho_1 L - \rho_2 L^2 - \ldots - \rho_p L^p) y_t = \varepsilon_t .$$

  The last expression, a p th order polynomial in the lag operator, can be written $A^p(L)$.
- y_t is stationary if all the roots (solutions), z_i, of $1 - \rho_1 z - \rho_2 z^2 - \ldots - \rho_p z^p = 0$ lie outside the unit circle (are greater than one in absolute value).
- The unit circle is defined on the plane which has $i = \sqrt{-1}$ on the y axis and the real line on the x axis. If the roots have imaginary parts (will be $\pm$) the process will cycle.

Stability 2

- If a root lies on the unit circle, some $z_i = 1$, the process is said to exhibit a unit root.
- The condition is sometimes expressed in terms of the inverse roots, which must lie inside the unit circle.
- Consider the case of an AR1 process

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

  For stability, the solution to $(1 - \rho z) = 0$, must be greater than unity in absolute value, since this implies $z = 1/\rho$ this requires $-1 < \rho < 1$.
- For an $AR2$ the real parts of solution to the two solutions to the quadratic $(1 - \rho_1 z - \rho_2 z^2)$ must be greater than unity.
- Usually we just check that $\sum \widehat{\rho}_i < 1$ for stationarity.

Random Walks

- If $\rho = 1$, there is a unit root, and the AR1 becomes a random walk: $y_t = y_{t-1} + \varepsilon_t$; or $\Delta y_t = \varepsilon_t$.
- Substituting back

$$y_t = \varepsilon_t + \varepsilon_{t-1} + ...\varepsilon_1 + y_0;$$

  so shocks have permanent effects.
- Although $E(y_t) = y_0$, constant,
  $Var(y_t) = E(y_t - y_0)^2 = t\sigma^2$, not constant
- A random walk with drift is of the form: $\Delta y_t = \alpha + \varepsilon_t$.

$$y_t = \alpha t + \varepsilon_t + \varepsilon_{t-1} + ...\varepsilon_1 + y_0;$$

  so $E(y_t) = y_0 + \alpha t$ trended.
- In both these cases, Δy_t is stationary, I(0), but y_t is non-stationary, I(1).
- Random walks appear very often in economics, e.g. the efficient market hypothesis implies that, to a first approximation, asset prices should be random walks.

Moving Average processes 1

- A first order moving average process (MA1) takes the form

$$y_t = \varepsilon_t + \mu\varepsilon_{t-1};$$

- A q th order moving average (MAq):

$$y_t = \varepsilon_t + \mu_1\varepsilon_{t-1} + \mu_2\varepsilon_{t-2} + \ldots + \mu_q\varepsilon_{t-q};$$

$$y_t = (1 + \mu_1 L + \mu_2 L^2 + \ldots + \mu_q L^q)\varepsilon_t = B^q(L)\varepsilon_t.$$

- Covariances for MA1

$$\begin{aligned}
\gamma_0 &= E(\varepsilon_t + \mu\varepsilon_{t-1})(\varepsilon_t + \mu\varepsilon_{t-1}), \\
&= E(\varepsilon_t^2 + \mu^2\varepsilon_{t-1}^2 + 2\varepsilon_t\mu\varepsilon_{t-1}) = \sigma^2(1+\mu^2). \\
\gamma_1 &= E(\varepsilon_t + \mu\varepsilon_{t-1})(\varepsilon_{t-1} + \mu\varepsilon_{t-2}) = \sigma^2\mu. \\
\gamma_2 &= E(\varepsilon_t + \mu\varepsilon_{t-1})(\varepsilon_{t-2} + \mu\varepsilon_{t-3}) = 0
\end{aligned}$$

- For MAq $Cov(y_t, y_{t-i}) = 0$, for $i > q$.

Moving Average processes 2

- A finite order moving average process is always stationary. Any stationary process can be represented by a (possibly infinite) moving average process.
- Notice that the AR1 is written as an infinite MA process in (1).
- The parameters of the MA model cannot be estimated by OLS, but maximum likelihood estimators are available.
- If a MA process is invertible we can write it as an AR, i.e. $B^q(L)^{-1}y_t = \varepsilon_t$.
- If we take a white noise process $y_t = \varepsilon_t$ and difference it we get

$$\Delta y_t = \varepsilon_t - \varepsilon_{t-1}$$

a non-invertible moving average process with a unit coefficient.

Next time

- Put the AR, MA and differencing together to get the autoregressive, integrated moving average model, ARIMA.
- Consider how to determine the order of integration.
- Unit root tests: Dickey Fuller and Augmented Dickey Fuller.