smfl7.tex

Lecture 7. 31.1.2011

It was shown by M. S. BARTLETT in 1946 (see e.g. Diggle 2.5) that for large n and τ non-zero,

$$r(\tau) \sim N(0, 1/n):$$

$r(\tau)$ is approximately *normal* with mean 0 and variance $1/n$. So as $\sqrt{n}r(\tau) \sim \Phi := N(0,1)$, the *standard normal* distribution, which takes values $> 1.96 \sim 2$ in modulus with probability 5%, only values of $r(\tau)$ with

$$|r(\tau)| \geq 1.96/\sqrt{n} \sim 2/\sqrt{n}$$

differ significantly from zero.

3. AUTOREGRESSIVE PROCESSES, AR(1)

Recall that in a linear regression model, the dependent variable Y depends in a linear way on an independent variable X (or $X_1, X_2, X_3, \cdots$, or $X, X^2, X^3, \cdots$), with an error structure or noise process also present.

In a TS model, the current value X_t depends in a linear way on the previous value X_{t-1} (or on the p previous values $X_{t-1}, X_{t-2}, \cdots, X_{t-p}$), again plus noise.

First-order case: $AR(1)$. Suppose that our model is

$$X_t = \phi X_{t-1} + m + \epsilon_t, \qquad ((\epsilon_t) \quad WN)$$

for t an integer (positive, negative or zero), where (ϵ_t) is a white noise process $WN(\sigma^2)$. Take means and use $EX_t = \mu$, $E\epsilon_t = 0$:

$$\mu = \phi\mu + m.$$

So if $\phi \neq 1$,

$$\mu = m/(1-\phi),$$

and if $\phi = 1$, then $m = 0$.

For simplicity, centre at means:

$$\begin{aligned} X_t - \mu &= \phi(X_{t-1} - \mu) + m - \mu + \phi\mu + \epsilon_t \\ &= \phi(X_{t-1} - \mu) + m - \mu(1-\phi) + \epsilon_t \\ &= \phi(X_{t-1} - \mu) + \epsilon_t, \end{aligned}$$

by above. Centring at means (i.e. replacing $X_t - \mu$ by X_t) for simplicity, we have

$$X_t = \phi X_{t-1} + \epsilon_t, \tag{$*$}$$

a simpler model, with all means zero. This is called an *autoregressive model of order one*, $AR(1)$. For, it has the form of a regression model, with X_{t-1} as the 'dependent variable' and X_t as the 'independent variable': X_t is regressed on the previous X-value (earlier in time), so the *process* (X_t) is *regressed on itself* (Greek: autos = self).

Using $(*)$ recursively,

$$\begin{aligned} X_t &= \phi(\phi X_{t-2} + \epsilon_{t-1}) + \epsilon_t \\ &= \phi^2 X_{t-2} + \phi\epsilon_{t-1} + \epsilon_t \\ &= \cdots \\ &= \phi^n X_{t-n} + \textstyle\sum_{i=0}^{n-1} \phi^i \epsilon_{t-1}. \end{aligned}$$

If $|\phi| < 1$, this suggests that the first term on the RHS $\to 0$ as $n \to \infty$, giving $X_t = \sum_0^\infty \phi^i \epsilon_{t-i}$. This is true, provided we interpret the convergence of the infinite series on RHS suitably. We have

$$\sum_1^{n-1} \phi^i \epsilon_{t-i})^2 = E[(\phi^n X_{t-n})^2] = \phi^{2n} E[X_{t-n}^2] = \phi^{2n}\gamma_0,$$

where $\gamma_0 = var X_t$ for all t. Since $|\phi| < 1$, $\phi^{2n} \to 0$ as $n \to \infty$, so RHS $\to 0$ as $n \to \infty$. So LHS $\to 0$ as $n \to \infty$. This says that

$$\textstyle\sum_0^n \phi^i \epsilon_{t-i} \to X_t \qquad (n \to \infty),$$

or

$$\textstyle\sum_0^\infty \phi^i \epsilon_{t-i} = X_t,$$

in mean square (or, in L_2).

Interpreting convergence in this mean-square sense,

$$X_t = \textstyle\sum_0^\infty \phi^i \epsilon_{t-i} \tag{$**$}$$

expresses X_t on LHS as a *weighted sum* of $\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \cdots$ on RHS. This weighted sum resembles an *average* (although the weights sum to $1/(1-\phi)$, not 1 as is usual for an average), and the set $(\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \cdots)$ of white-noise variables being averaged over *moves* with t; there are *infinitely many* of them.

Hence $(**)$ is called the *infinite moving-average representation* of the $AR(1)$ process $(*)$. Note that the further we go back in time, the more the ϵ_{t-i} are down-weighted by the geometrically decreasing weights ϕ^i.

Autocovariance of $AR(1)$. Since ϵ_{t+1} is independent of (or, using the weaker definition of white noise, uncorrelated with) $\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \cdots$, it is independent of (or uncorrelated with) the linear combination $X_t = \sum_0^\infty \phi^i \epsilon_{t-i}$ of them. So ϵ_{t+1} is uncorrelated with $X_t, X_{t-1}, \cdots$. This says that X_s and ϵ_t are uncorrelated for $s < t$. Since all means are zero:

$$E(X_s \epsilon_t) = 0 \qquad (s < t).$$

Square both sides of $(*)$ and take expectations:

$$E[X_t^2] = \phi^2 E[X_{t-1}^2] + 2\phi E[X_{t-1}\epsilon_{t-1}] + E[\epsilon_t^2].$$

The second term on RHS is zero by above; $E[X_t^2] = var X_t = \gamma_0$ for all t, and $E[\epsilon_t^2] = var \epsilon_t = \sigma^2$ for all t. So

$$\gamma_0 = \phi^2 \gamma_0 + \sigma^2 :$$

$$\gamma_0 = \sigma^2/(1 - \phi^2),$$

identifying γ_0 in terms of the WN variance σ^2 and the weight ϕ.

Multiply $(*)$ by $X_{t-\tau}$ $(\tau \geq 1)$ and take expectations:

$$\gamma_\tau = \phi \gamma_{\tau-1}$$

(since ϵ_t on RHS is uncorrelated with $X_{t-\tau}$). Using this repeatedly,

$$\gamma_\tau = \phi\gamma_{\tau-1} = \phi^2 \gamma_{\tau-2} = \cdots = \phi^\tau \gamma_0 = \phi^\tau \sigma^2/(1-\phi^2) :$$

$$\gamma_\tau = \sigma^2.\phi^\tau/(1-\phi^2) \qquad (\tau \geq 0),$$

giving the autocovariance of an $AR(1)$ process as geometrically decreasing. Passing to the autocorrelation $\rho_\tau = \gamma_\tau/\gamma_0$: $\rho_\tau = \phi^\tau$ for $\tau \geq 0$). Note that $\rho_\tau = \rho_{-\tau}$ (since two random variables have the same covariance and correlation either way round), so we can re-write this as

$$\rho_\tau = \phi^{|\tau|}.$$

Recall $|\phi| < 1$ here. Two cases are worth distinguishing.

Case 1: $0 \leq \phi < 1$. Here the graph of ρ_τ is a geometric series with non-negative common ratio. Since the sample autocorrelation r_τ is an approximation to ρ_τ, the correlogram (graph of r_τ) is an approximation to this.

Successive values of X_t are positively correlated: positive values of X_t tend to be succeeded by positive values, and similarly negative by negative.
Case 2: $-1 < \phi < 0$. Here the graph is again a geometric series, but one that *oscillates in sign*, as well as damping down geometrically. Successive values of X_t are negatively correlated: positive values tend to be succeeded by negative values, and vice versa.

To summarise: the signature of an $AR(1)$ process is a correlogram that looks like an approximation to a geometric series, as in Case 1 or 2 above, depending on the sign of ϕ.

The Lag Operator.

Before proceeding, we introduce some useful notation and terminology. The *lag operator*, or *backward shift operator*, operates on sequences by shifting the index back in time by one. We write it as B:

$$BX_t = X_{t-1},$$

(though L - L for lag - is also used). Repeating this, B^2 shifts back in time by two, $B^2 X_t = X_{t-2}$, and generally

$$B^n X_t = X_{t-n} \qquad (n = 0, 1, 2, \cdots)$$

($B^0 = I$ is the identity operator: $B^0 X_t = I X_t = X_t$).

We can re-write $(*)$ in this notation as

$$X_t = \phi B X_t + \epsilon_t : \qquad (1 - \phi B) X_t = \epsilon_t.$$

Formally, this suggests

$$\begin{aligned} X_t = (1 - \phi B)^{-1} \epsilon_t &= (1 + \phi B + \phi^2 B^2 + \cdots + \phi^i B^i + \cdots) \epsilon_t \\ &= 1 + \phi \epsilon_{t-1} + \phi^2 \epsilon_{t-2} + \cdots + \phi^i \epsilon_{t-i} + \cdots \\ &= \textstyle\sum_0^\infty \phi^i \epsilon_{t-i}, \end{aligned}$$

which is $(**)$ as above, *provided* that the operator equation

$$(1 - \phi B)^{-1} = \textstyle\sum_{i=0}^\infty \phi^i B^i$$

makes sense. It does make sense, with convergence on the RHS interpreted in the *mean-square sense* as above, if $|\phi| < 1$.