Revision concepts

Univariate Time Series Models

Assumption: the series has been generated by a *stochastic process*, $\{Y_t\}$, that is, each element in the series, $Y_1, Y_2, ..., Y_t,$ is drawn randomly from a probability distribution.

Definitions:

A time series is a collection of random variables ordered in time $\{Y_1, Y_2, ..., Y_t\}$. The stochastic process $\{Y_t\}$ can be described by the joint probability distribution (T-dimensional), with mean, variance and covariances:

Mean = $E(Y_t)=\mu_t$, $t=1,2,...,T$

$V(Y_t) = E(Y_t-\mu_t)^2 = \sigma^2_t$, $t=1,2,...,T$

$Cov(Y_t,Y_s) = E[(Y_t-\mu_t)(Y_s-\mu_s)] = \gamma(t,s)$

autocovariances between Y_t and Y_s, $t,s=1,2,...,T$ and $t\neq s$. If $t=s$ we obtain the variance of Y, denoted as $\gamma(0)$.

The <u>observed</u> time series $\{y_1, y_{2,...,}y_t\}$, is regarded as a particular realisation (sample) of the stochastic process (population).

In practice we have only a <u>single realisation</u> of a time series, for example, GDP in 1997, quarter 4, to infer the unknown parameters of the stochastic process.

This procedure is valid only if the process is *ergodic*. Ergodicity implies that the sample moments of the realisation approach the population moments as T (the length of the realisation) becomes infinite.

Strict stationarity. A time series is said to be strictly stationary if the joint and conditional probability distributions of the process are unchanged if displaced in time. Thus the PDF of any set of T observations $Y_{t1}, Y_{t2}, \ldots, Y_{tT}$ must be the same as the PDF of $Y_{t1+k}, Y_{t2+k}, \ldots, Y_{tT+k}$, for any t, k.

In practice it is more common to use a weaker definition of stationarity.

Weak stationarity (or covariance stationarity). A time series is said to be weakly stationary if the mean, variance and covariances are independent of t.

Thus, for a weakly stationary time series it holds that:

$E(Y_t)$=constant=μ
$Var(Y_t)= E(Y_t-\mu)^2$ =constant= σ^2

k^{th} order autocovariance:

$$Cov(Y_t,Y_{t-k})=E[(Y_t-\mu)(Y_{t-k}-\mu)]= \gamma(k)\ ,\ \ k=...,-2,-1,0,1,2,...$$

i.e. the autocovariance between Y_t and Y_{t-k} depends on k, the time gap, and not on t, and they are symmetrical about lag zero, thus:

$$Cov(Y_t,Y_{t-k})= Cov(Y_t,Y_{t+k})=Cov(Y_{t-j},Y_{t-j-k}).$$

For k=0, $\gamma(0)=Var(Y_t)$.

If any of these conditions are not satisfied, the time series is *nonstationary*.

For a <u>covariance stationary</u> series we define:

Autocorrelation coefficients at lag k:

$$\rho(k)=\frac{Cov(Y_tY_{t-k})}{\sqrt{Var(Y_t)}\sqrt{Var(Y_{t-k})}}=\frac{Cov(Y_tY_{t-k})}{Var(Y_t)}=\frac{\gamma(k)}{\gamma(0)}, \qquad -1 < \rho(k) < 1,$$

k=...,-2,-1,0,1,2,...(time independent) $\therefore\ \rho(0) = 1$

<u>Autocorrelation function</u> (ACF): a formula giving the autocorrelation coefficients. Useful to characterise ARMA models (see later). For example:

ACF for the AR(1) process: $\rho_k = \phi\, \rho_{k-1} = \phi^{k}$

ACF for the AR(2) process : $\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2}$, k=2,3,…

<u>Correlogram:</u> a graphical representation of the ACF, a plot of ρ(k) against k.

Stationary Time series Models

(i) White-noise process (purely random process)

the sequence $\{\varepsilon_t\}$ is a white noise process if $\varepsilon_t \sim iid(0,\sigma^2)$,

$$E(\varepsilon_t) = 0$$
$$V(\varepsilon_t) = \sigma^2 < \infty$$
$$Cov(\varepsilon_t \varepsilon_s) = 0, \ \forall t \neq s$$

(ii) Autoregressive process of order p - AR(p)

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \varepsilon_t \quad \varepsilon_t \sim iid(0,\sigma^2) \quad (1)$$

y_t is regressed on past values. ϕ_1, ϕ_2,...,ϕ_p are unknown parameters. Using the lag operator $L^k y_t = y_{t-k}$ the AR(p) process can be written as:

$\phi(L)y_t = \varepsilon_t$ where $\phi(L)=1- \phi_1 L- \phi_2 L^2 -\ldots- \phi_p L^p$

This process is a pth order stochastic difference equation. There is a direct link between the stability (convergence) condition of a difference equation and the stationarity of an economic variable. If all the roots of the characteristic equation of (1):

$$b^p - \phi_1 b^{p-1} - \phi_2 b^{p-2} - \ldots - \phi_p = 0 \qquad (2)$$

lie within the unit circle (stability condition, $|b_i|<1$), the AR process is stationary. Or, equivalently, the characteristic roots of the lag polynomial:

$$1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p = 0 \qquad (3)$$

(called the inverse characteristic equation), must lie outside the unit circle ($|L_i|>1$). The polynomial $\phi(L)=1-\phi_1 L-\phi_2 L^2-\ldots-\phi_p L^p$ can always be factorised as

$$\phi(L)=1-\phi_1 L-\phi_2 L^2-\ldots-\phi_p L^p=(1-b_1 L)(1-b_2 L)\ldots(1-b_p L)$$

where the b_i may be seen as the roots of the characteristic equation (2) above.

For example, in an **AR(2) process**

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$$

The characteristic equation is $b^2 - \phi_1 b - \phi_2 = 0$, with roots

$$b_1, b_2 = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{2}$$

Now, the lag polynomial can be factorised as

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 = (1 - b_1 L)(1 - b_2 L)$$

and the roots are $L_i = 1/b_i$ $(i=1,2)$, so that if $|b_i|<1$, $|L_i|>1$.

The connection between the ϕ_i and the b_i parameters is $b_1 + b_2 = \phi_1$, and $b_1 b_2 = -\phi_2$.

In the AR(2) case complex roots will occur if $\phi_1^2 + 4\phi_2 < 0$. The roots can be written

$$b_1, b_2 = h \pm vi$$, with h and v real numbers, and $i = \sqrt{-1}$

The absolute value of each root is $|b_i| = \sqrt{h^2 + v^2} = -\phi_2$

The stationarity conditions for the AR(2) process can also be expressed in terms of the coefficients ϕ_i as follows: $\phi_2 + \phi_1 < 1$; $\phi_2 - \phi_1 < 1$; $|\phi_2| < 1$

AR(1) process

$$y_t = \phi y_{t-1} + \varepsilon_t ,$$

y_t is regressed on y_{t-1}, a first order difference equation. Using the lag operator:

$$y_t - \phi L y_t = \varepsilon_t \quad \Rightarrow \quad (1-\phi L)y_t = \varepsilon_t$$

Necessary condition for stationarity: $|\phi| < 1$, or |L|>1.

Solving the AR(1) process above for y_t:

$$y_t = \frac{\varepsilon_t}{(1-\phi L)}, \quad (1-\phi L)^{-1} = 1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \ldots,$$ an infinite sum.

So, $y_t = (1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \ldots)\varepsilon_t = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}$, an infinite MA process.

If $|\phi| < 1$, the effects of ε_{t-j} will converge to zero as $j \to \infty$, i.e. the influence of any shock will go to zero over time, at a rate that depends on the value of ϕ. So, the stationarity condition ensures that the MA form is not explosive.

It can be shown that any AR(p) model can be similarly expressed.

Stationarity and unit roots

The process is nonstationary, and said to have a unit root if L=1 solves $\phi(L)=0$, i.e. if $\phi(1)= 1- \phi= 0$. This implies $\phi=1$ in the AR(1) model.

More generally, if $\phi(L)=1- \phi_1 L - \phi_2 L^2 -...-\phi_p L^p$, then the process will contain a unit root if $\phi(1)= 1- \phi_1 - \phi_2 -...-\phi_p= 0$.

A variable with a unit root is said to be integrated of order one, written I(1), where 1 denotes a single unit root. A stationary process is written I(0).

An I(2) process contains 2 unit roots. Example AR(2): $Y_t=2Y_{t-1}-Y_{t-2}+u_t$, which can be rewritten as $(1-L)(1-L)Y_t=u_t$.

(iii) Moving Average process of order q - MA(q)

$$y_t = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q} + \varepsilon_t \qquad \varepsilon_t \sim iid(0, \sigma^2)$$

y_t is a weighted average of the $\varepsilon_t, \varepsilon_{t-1}, \ldots, \varepsilon_{t-q}$

Moving Average process of order 1 MA(1):

$$y_t = \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

or, using the lag operator:

$y_t = \theta(L)\varepsilon_t \qquad \theta(L) = 1 + \theta_1 L$

In practice we observe the y_t series, and the ε_ts can only be estimated by the equation:

$$[\theta(L)]^{-1} y_t = \varepsilon_t$$

If a non-explosive solution for ε_t is found, the MA model is said to be *invertible*, i.e. an infinite AR representation (AR(∞)) exists.

For the MA(1) model the invertibility condition is that $\theta_1<1$, or the solution of $\theta(L)=1+\theta_1 L=0$ is >1. In general, MA(q) models are called invertible when the solutions to

$$\theta(L)=1+\theta_1 L+\theta_2 L^2+\ldots+\theta_q L^q = 0$$

are all outside the unit circle.

(iv) Autoregressive Moving Average process of order p,q - ARMA(p,q)

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

This process combines an AR(p) with an MA(q) model. The AR(p) component is a linear difference equation. If all the characteristic roots of the AR component are within the unit circle (stability condition), the ARMA(p,q) model is stationary.

Wold decomposition theorem ensures that any stationary process can be written in MA form (although this might need to have an infinite order).

ARMA(1,1) $\quad y_t = \phi_1 y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$.

$(1-\phi_1 L)y_t = (1+\theta_1 L)\varepsilon_t \quad$ stationary if $|\phi_1| < 1$; invertible if $|\theta_1| < 1$

Consider now the case of *common roots* or *redundant parameters* in ARMA models

ARMA(1,1) $y_t = \phi_1 y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$.

$$(1-\phi_1 L)y_t = (1+\theta_1 L)\varepsilon_t$$

If $\theta_1 = -\phi_1$, can devide both sides by $(1+\theta_1 L)$ to obtain a white noise process $y_t = \varepsilon_t$

ARMA (2,1) $(1-\phi_1 L-\phi_2 L^2)y_t = (1+\theta_1 L)\varepsilon_t$

rewrite as $(1-b_1 L)(1-b_2 L)y_t = (1+\theta_1 L)\varepsilon_t$

If $\theta_1 = -b_1$ can devide both sides by $(1+\theta_1 L)$ to obtain an AR(1) process: $(1-b_2 L)y_t = \varepsilon_t$

In general, if there is one common root, an ARMA(p,q) can be written equivalently as an ARMA(p-1, q-1) model.

Properties of AR series

AR(1) model

$y_t = \phi y_{t-1} + \varepsilon_t$ can be written as infinite MA:

$y_t = \varepsilon_t + \phi\,\varepsilon_{t-1} + \phi^2\,\varepsilon_{t-2} + \ldots\ldots$

$= \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}$ called the Wold representation

Note: This is a special case of a more general result. The Wold's decomposition theorem states that any zero-mean covariance stationary process can be represented in the form of an infinite moving average.

Mean: $E(y_t) = E(\varepsilon_t + \phi\,\varepsilon_{t-1} + \phi^2\,\varepsilon_{t-2} + \ldots\ldots)$

$= E(\varepsilon_t) + \phi E(\varepsilon_{t-1}) + \phi^2 E(\varepsilon_{t-2}) + \ldots\ldots$

$= 0$

as $E(\varepsilon_t)=0$ by assumption. The mean is finite and time independent.

<u>Variance</u>: $\quad Var(y_t) = Var(\varepsilon_t + \phi\, \varepsilon_{t-1} + \phi^2\, \varepsilon_{t-2} + \ldots\ldots)$

$$= \sigma_\varepsilon^{\ 2} + \phi^2\, \sigma_\varepsilon^{\ 2} + \phi^4\, \sigma_\varepsilon^{\ 2} + \ldots..$$

$$= \sigma_\varepsilon^{\ 2}\, (1 + \phi^2 + \phi^4 + \ldots.)$$

$$\Rightarrow \quad \sigma_y^{\ 2} = \frac{\sigma_\varepsilon^2}{1-\phi^2} = \gamma(0)$$

as $Cov(\varepsilon_t\, ,\, \varepsilon_{t-j}) = 0$. The variance is finite and time independent.

(Auto) Covariances:

$$Cov(y_t, y_{t-1}) = Cov[\phi\, y_{t-1} + \varepsilon_t\, ,\, y_{t-1}] = \phi\, Var(y_{t-1}) + 0 = \phi\, \sigma_y^2$$

$$\therefore\ \gamma(1) = \phi\, \sigma_y^2 \qquad [\ \gamma(0) = \sigma_y^2\ ]$$

as $Var(y_t) = Var(y_{t-1}) = \sigma_y^2$ and $E(\varepsilon_t\, y_{t-1}) = 0$

$$\therefore\ \rho(1) = \frac{\gamma(1)}{\gamma(0)} = \phi$$

$$Cov(y_t\, ,\, y_{t-2}) = Cov[\phi\, y_{t-1} + \varepsilon_t\, ,\, y_{t-2}] = \phi\, Cov(y_{t-1}\, ,\, y_{t-2}) + 0$$

$\therefore\ \gamma(2) = \phi\,\gamma(1) = \phi^2\,\sigma_y^{\,2}$

$\therefore\ \rho(2) = \dfrac{\gamma(2)}{\gamma(0)} = \phi^2$ etc............................

$\gamma(k) = \phi^k\,\sigma_y^{\,2}$, and

$\rho(k) = \phi^k$ for $k = 0, 1, 2, \ldots$

Note $\rho(0) = 1$, and $\rho(k) = \phi^k \to 0$ as $k \to \infty$ if $-1 < \phi < 1$

For an AR(1) process the coefficient on $y_{t-1} = \phi$ = 1st order autocorrelation coefficient $\rho(1)$. This is only true for AR(1).

So, the autocorrelation function for the AR(1) model decays to zero, i.e. the influence of any shock goes to zero, at a rate which depends on the value of ϕ.

Mean, variance and autocovariances are all finite and time independent: the process is stationary.

For the AR(2) process it can be shown that

$$\gamma(0)=\frac{\sigma^2}{1-\phi_1\rho_1-\phi_2\rho_2}\ ; \quad \rho_1=\frac{\phi_1}{1-\phi_2}; \quad \rho_2=\frac{\phi_1^{\ 2}}{1-\phi_2}+\phi_2$$

By combining the expressions above the unconditional variance can be rewritten as

$$\gamma(0)=\frac{(1-\phi_2)\sigma^2}{(1+\phi_2)[(1-\phi_2)^2-\phi_1^2]}$$

Also note that $\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2}$, k=2,3,…, a second order difference equation, with coefficients ϕ_1 and ϕ_2.

So, stationarity conditions ensure that ACF dies out as lag increases, i.e. the ρ_k sequence must be convergent. Initial values are $\rho_0=1$ and ρ_1 above.

Properties of MA series

MA(1) $y_t = \varepsilon_t + \theta\, \varepsilon_{t-1}$

Mean $E(y_t) = 0$;

Variance $Var(y_t) = (1+\theta^2)\sigma_\varepsilon^2 = \gamma(0)$

Covariances $Cov(y_t, y_{t-1}) = Cov(\varepsilon_t + \theta\, \varepsilon_{t-1}\, ,\, \varepsilon_{t-1} + \theta\, \varepsilon_{t-2}) = \theta\, \sigma_\varepsilon^2 = \gamma(1)$.

$$\therefore\ \rho(1) = \frac{\gamma(1)}{\gamma(0)} = \frac{\theta\sigma_\varepsilon^2}{(1+\theta^2)\sigma_\varepsilon^2} = \frac{\theta}{1+\theta^2}$$

and $\gamma(2) = \gamma(3) = \ldots = 0$ which implies that $\rho(2) = \rho(3) = \ldots = 0$

For the MA(2) process it can be shown that

$$Var(y_t) = (1+\theta_1^2+\theta_2^2)\sigma_\varepsilon^2 = \gamma(0)$$

$$\rho(1) = \frac{\theta_1+\theta_1\theta_2}{1+\theta_1^2+\theta_2^2};\ \rho(2) = \frac{\theta_2}{1+\theta_1^2+\theta_2^2};\ \rho(3)=\rho(4)=..=0$$

Thus for an MA(q), $\rho_k = 0 \quad \forall \quad k > q$.

Properties of ARMA series

For the ARMA(1,1) process it can be shown that:

$$\gamma(0) = \frac{1+\theta_1^2 - 2\phi_1\theta_1}{1-\phi_1^2}\sigma^2;$$

$$\gamma(1) = \frac{(1-\phi_1\theta_1)(\phi_1-\theta_1)}{1-\phi_1^2}\sigma^2$$

$$\rho(1) = \frac{(1-\phi_1\theta_1)(\phi_1-\theta_1)}{1+\theta_1^2 - 2\phi_1\theta_1};$$

$\rho(k) = \phi_1\rho(k-1)$ for $k>1$.

Partial autocorrelation functions (PACF)

All AR processes imply ACFs that dump out. PACFs are used to help discriminating between AR processes of different orders.

PACs are calculated to adjust for the portion of autocorrelation between y_t and y_{t-k}, due to the correlation that these variables have with the intermediate lags y_{t-1}, y_{t-2}, ..., y_{t-k+1}.

The lag k partial autocorrelation is the partial regression coefficient ϕ_{kk} in the kth order autoregression:

$$y_t = \phi_{k1} y_{t-1} + \phi_{k2} y_{t-2} + ... + \phi_{kk} y_{t-k} + \varepsilon_t$$

and it measures the additional correlation between y_t and y_{t-k} after adjusting for the intermediate variables y_{t-1}, y_{t-2}, ..., y_{t-k+1}.

AR(1): $\phi_{11} = \rho_1$ $\phi_{kk} = 0$ for $k>1$

AR(2): $\phi_{11} = \rho_1$, $\phi_{22} = \dfrac{\rho_2 - \rho_1^2}{1-\rho_1^2}$ $\phi_{kk} = 0$ for $k>2$

AR(p) $\phi_{11}\neq 0$, $\phi_{22}\neq 0$,..., $\phi_{pp}\neq 0$, $\phi_{kk} = 0$ for $k>p$

PACs are zero for lags larger than the order of the process.

Summary Table: theoretical values. (The estimated sample values may or may not correspond to these population values)

	AR(1)	MA(1)
Var(y_t)	$\dfrac{\sigma_\varepsilon^2}{1-\phi^2}$	$(1+\theta^2)\sigma_\varepsilon^2$
ρ_k=corr($y_t y_{t-k}$)	ϕ^k	$\dfrac{\theta}{1+\theta^2}$ k=1 0 k=2,3,...
Stationarity (stability) condition	$-1 < \phi < 1$	No restriction on θ is needed
Invertibility condition (this means that an MA process can be written as an infinite AR process)		$-1 < \theta < 1$

Summary of correlation patterns

Process	**ACF**	**PACF**
AR(p)	Infinite: damps out	Finite: cuts off after lag p
MA(q)	Finite: cuts off after lag q	Infinite: damps out
ARMA	Infinite: damps out	Infinite: damps out