Brough Lecture Notes: Beginning Time Series Topics II

Finance 5330: Financial Econometrics Tyler J. Brough Last Updated: January 25, 2019

Time Series Continued

Summary of allowed lag polynomial manipulations

1. We can multiply them

$$a(L)b(L) = (a_0 + a_1 L + \cdots)(b_0 + b_1 L + \cdots) = a_0 b_0 + (a_0 b_1 + b_0 a_1)L + \cdots$$

2. They commute

$$a(L)b(L) = b(L)a(L)$$

3. We can raise them to positive integer powers

$$a(L)^2 = a(L)a(L)$$

4. We can invert them, by factoring them and inverting each term

$$\begin{aligned} a(L) &= (1 - \lambda_1 L)(1 - \lambda_2 L) \quad \cdots \\ a(L)^{-1} &= (1 - \lambda_1 L)^{-1}(1 - \lambda_2 L)^{-1} \\ &= \sum_{j=0}^{\infty} \lambda_1^j L^j \sum_{j=0}^{\infty} \lambda_2^j L^j \\ &= c_1(1 - \lambda_1 L)^{-1} + c_2(1 - \lambda_2 L)^{-1} \quad \cdots \end{aligned}$$

We'll look at roots greater than and/or equal to one, fractional powers, and non-polynomial functions of lag operators later.

Multivariate ARMA Models

The multivariate case is similar, reinterpreting our variables as vectors and matrices:

$$x_t = \begin{bmatrix} y_t \\ z_t \end{bmatrix}$$

The building block is the multivariate white noise process, $\varepsilon_t \sim iidN(0, \Sigma)$, which we write as follows

$$\varepsilon_t = \begin{bmatrix} \delta_t \\ \nu_t \end{bmatrix}$$

with

$$\begin{aligned} E(\varepsilon_t) &= 0 \\ E(\varepsilon_t \varepsilon_t') = \Sigma &= \begin{bmatrix} \sigma_\delta^2 & \sigma_{\delta\nu}^2 \\ \sigma_{\nu\delta}^2 & \sigma_\nu^2 \end{bmatrix} \end{aligned}$$

and

$$E(\varepsilon_t \varepsilon_{t-j}') = 0$$

The AR(1) is $x_t = \phi x_{t-1} + \varepsilon_t$, or

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \phi_{yy} & \phi_{yz} \\ \phi_{zy} & \phi_{zz} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \delta_t \\ \nu_t \end{bmatrix}$$

Or

$$\begin{aligned} y_t &= \phi_{yy} y_{t-1} + \phi_{yz} z_{t-1} + \delta_t \\ z_t &= \phi_{zy} y_{t-1} + \phi_{zz} z_{t-1} + \nu_t \end{aligned}$$

NB: this is a Vector autoregressive model of order 1, or a VAR(1) model.

NB: both lagged y and lagged z appear in each equation.

Thus, the VAR(1) captures cross-variable dynamics.

Ex: It could capture the fact that if volume is higher in one trading period, volatility tends to be higher the following trading period; as well as the fact theat if volatility is high one period volume tends to be high the next period.

We can write the VAR(1) in lag operator notation:

$$(I - \Phi L)x_t = \varepsilon_t$$

or

$$A(L)x_t = B(L)\varepsilon_t$$

where:

- $A(L) = I - \Phi_1 L - \Phi_2 L^2 - \cdots$

- $B(L) = I + \Theta_1 L + \Theta_2 L^2 + \cdots$

NB: $\Phi_j = \begin{bmatrix} \phi_{j,yy} & \phi_{j,yz} \\ \phi_{j,zy} & \phi_{j,zz} \end{bmatrix}$ and similarly for Θ_j.

We can invert multivariate ARMA models.

For example, the MA(∞) representation can be obtained from the VAR(1) as

$$(I - \Phi L)x_t = \varepsilon_t \quad \Leftrightarrow \quad (I - \Phi L)^{-1}\varepsilon_t = \sum_{j=0}^{\infty} \Phi^j \varepsilon_{t-j}$$

Autocorrelation and Autocovariance Functions

Autocovariance of a series x_t is

$$\gamma_j = Cov(x_t, x_{t-j})$$

Hence,

$$\gamma_j = E(x_t x_{t-j})$$

NB: $\gamma_0 = Var(x_t)$

NB: Recall $Cov(x_t, x_{t-j}) = E[(x_t - E(x_t))(x_{t-j} - E(x_{t-j}))]$ but $E(x_t) = 0$ for our purposes.

Autocorrelation is:

$$\rho_j = \frac{\gamma_j}{Var(x_t)} = \frac{\gamma_j}{\gamma_0}$$

Autocovariance and Autocorrelation of ARMA Processes

White noise: since we assume $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$, it's clear that

$$\gamma_0 = \sigma_{\varepsilon_t}^2, \quad \gamma_j = 0 \quad \text{for all} \quad j \neq 0$$

$$\rho_0 = 1, \quad \rho_j = 0 \quad \text{for all} \quad j \neq 0$$

MA(1)

$$x_t = \varepsilon_t + \theta \varepsilon_{t-1}$$

Autocovariance:

$$
\begin{aligned}
\gamma_0 &= Var(x_t) = Var(\varepsilon_t + \theta\varepsilon_{t-1}) \\
&= \sigma_\varepsilon^2 + \theta^2\sigma_\varepsilon^2 \\
&= (1+\theta^2)\sigma_\varepsilon^2 \\
\gamma_1 &= E(x_t x_{t-1}) = E[(\varepsilon_t + \theta\varepsilon_{t-1})(\varepsilon_{t-1} + \theta\varepsilon_{t-2})] \\
&= E[\theta\varepsilon_{t-1}^2] \\
&= \theta\sigma_\varepsilon^2 \\
\gamma_2 &= E(x_t x_{t-2}) = E[(\varepsilon_t + \theta\varepsilon_{t-1})(\varepsilon_{t-1} + \theta\varepsilon_{t-3})] = 0 \\
\gamma_3 &= 0
\end{aligned}
$$

Autocorrelation

$$
\begin{aligned}
\rho_1 &= \frac{\theta}{1+\theta^2} \\
\rho_2 &= 0
\end{aligned}
$$

MA(2)

$$x_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2}$$

Autocovariance:

$$
\begin{aligned}
\gamma_0 &= E[(\varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2})(\varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2})] \\
&= (1+\theta_1^2+\theta_2^2)\sigma_\varepsilon^2 \\
\gamma_1 &= E[(\varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2})(\varepsilon_{t-1} + \theta_1\varepsilon_{t-2} + \theta_2\varepsilon_{t-3})] \\
&= (\theta_1 + \theta_1\theta_2)\sigma_\varepsilon^2
\end{aligned}
$$

$\gamma_3, \gamma_4, \ldots = 0$

Autocorrelation:

$$
\begin{aligned}
\rho_0 &= 1 \\
\rho_1 &= \frac{\theta_1 + \theta_1\theta_2}{(1+\theta_1^2+\theta_2^2)} \\
\rho_2 &= \frac{\theta_2}{(1+\theta_1^2+\theta_2^2)}
\end{aligned}
$$

$\rho_3, \rho_4, \ldots = 0$

MA(q), MA(∞)

By now the pattern is clear: MA(q) processes have q autocorrelations different from zero. Also, if

$$x_t = \theta(L)\varepsilon_t = \sum_{j=0}^{\infty}(\theta_j L^j)\varepsilon_t$$

then

$$\gamma_0 = Var(x_t) = \left(\sum_{j=0}^{\infty}\theta_j^2\right)\sigma_\varepsilon^2$$
$$\gamma_k = \sum_{j=0}^{\infty}\theta_j\theta_{j+k}\sigma_\varepsilon^2$$

NB: $\theta_0 = 1$

NB: The lesson is that calculation of 2nd moments for MA processes is easy. Because covariance terms $E(\varepsilon_j\varepsilon_k)$ drop out!

AR(1)

Two ways to proceed:

1. Invert the MA(∞) and use the above

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

$$\gamma_0 = \left(\sum_{j=0}^{\infty}\phi^{2j}\right)\sigma_\varepsilon^2 = \frac{1}{1-\phi^2}\sigma_\varepsilon^2; \quad \rho_0 = 1$$
$$\gamma_1 = \left(\sum_{j=0}^{\infty}\phi^j\phi^{j+1}\right)\sigma_\varepsilon^2 = \phi\left(\sum_{j=0}^{\infty}\phi^{2j}\right)\sigma_\varepsilon^2 = \frac{\phi}{1-\phi^2}\sigma_\varepsilon^2; \quad \rho_1 = \phi$$

and continuing

$$\gamma_k = \frac{\phi^k}{1-\phi^2}\sigma_\varepsilon^2; \quad \rho_k = \phi^k$$

2. Another way useful in it's own right

$$\gamma_1 = E(x_t x_{t-1}) = E[(\phi x_{t-1}+\varepsilon_t)(x_{t-1})] = \phi\sigma_x^2; \quad \rho = \phi$$
$$\gamma_2 = E(x_t x_{t-2}) = E[(\phi^2 x_{t-2}+\phi\varepsilon_{t-1}+\varepsilon_t)(x_{t-2})] = \phi^k\sigma_x^2; \quad \rho_k = \phi^k$$

AR(p) Yule-Walker Equations

This latter method is the easiest way to proceed for AR(p)'s.

Let's look at an AR(3), then you can generalize.

Multiplying both sides by $x_t, x_{t-1}, \cdots$ taking expectations, then dividing by γ_0 we obtain

$$
\begin{aligned}
1 &= \phi\rho_1 + \phi_2\rho_2 + \phi_3\rho_3 + \frac{\sigma_\varepsilon^2}{\gamma_0} \\
\rho_1 &= \phi_1 + \phi_2\rho_1 + \phi_3\rho_2 \\
\rho_2 &= \phi_1\rho_1 + \phi_2 + \phi_3\rho_1 \\
\rho_3 &= \phi_1\rho_2 + \phi_2\rho_1 + \phi_3 \\
\rho_k &= \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \phi_3\rho_{k-3}
\end{aligned}
$$

The 2nd, 3rd, and 4th equations can be solved for ρ_1, ρ_2, ρ_3.

Then each remaining equation gives ρ_k in terms of ρ_{k-1} and ρ_{k-2}, so we can solve for all of the ρ's.

NB: The ρ's follow the same difference equation as the original x's

The first equation can be solved for the variance

$$
\sigma_x^2 = \gamma_0 = \frac{\sigma_\varepsilon^2}{1 - [\phi_1\rho_1 + \phi_2\rho_2 + \phi_3\rho_3]}
$$

Stationarity

In calculating the moments of ARMA processes, the moments did not depend on calendar time

$$
\begin{aligned}
E(x_t) &= E(x_s) \qquad \text{for all } t \text{ and } s \\
E(x_t x_{t-j}) &= E(x_s x_{s-j}) \quad \text{for all } t \text{ and } s
\end{aligned}
$$

These properties are true for the invertible ARMA models, but they reflect a deeper property.

A process $\{x_t\}$ is **strongly stationary** or **strictly stationary** if the joint probability distribution function of $\{x_{t-s}, \cdots, x_t, \cdots, x_{t+s}\}$ is independent of t for all s.

A process $\{x_t\}$ is **weakly stationary** or **covariance stationary** if $E(x_t)$, $E(x_t^2)$ are finite and $E(x_t x_{t-j})$ depends only on j and not on t.

NB:

1. Strong stationary does not imply weak stationarity. $E(x_t^2)$ must be finite.
    - Ex: on iid Cauchy process is strongly, but not covariance stationary.
2. Strong stationarity plus $E(x_t), E(x_t^2) < \infty \Rightarrow$ weak stationarity

3. Weak stationarity does not $\Rightarrow$ strong stationarity. If the process is not normal, other moments $(E(x_t x_{t-j} x_{t-k}))$ *might* depend on t, so the process might not be strongly stationary.

4. Weak stationarity plus normality $\Rightarrow$ strong stationarity