Economics 210B - Macroeconomics Winter 2008-09

Prof. Garey Ramey

Notes on Time Series Modeling

1 $MA(\infty)$ representation of $AR(p)$ processes

Definition $\{y_t\}$ *is an* $MA(\infty)$ process *(infinite-order moving average) if*

$$y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \tag{1}$$

where $\{\varepsilon_t\}$ *is white noise and* $\mu, \psi_0, \psi_1, ...$ *are arbitrary constants.*

A sufficient condition for stationarity of an $MA(\infty)$ process is square summability of $\{\psi_j\}$:

$$\sum_{j=0}^{\infty} \psi_j^2 < \infty.$$

Definition $\{y_t\}$ *is an* $AR(p)$ process *(p^{th}-order autoregressive) if*

$$\begin{aligned} y_t &= c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t \\ &= c + \sum_{i=1}^{p} \phi_i y_{t-i} + \varepsilon_t, \end{aligned} \tag{2}$$

where $\{\varepsilon_t\}$ *is white noise and* $c, \phi_1, ..., \phi_p$ *are arbitrary constants.*

Represent (2) using the lag operator:

$$(1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p) y_t = c + \varepsilon_t.$$

To obtain an $MA(\infty)$ representation, consider the equation

$$\lambda^p - \phi_1 \lambda^{p-1} - \phi_2 \lambda^{p-2} - ... - \phi_p = 0. \tag{3}$$

According to the fundamental theorem of algebra, there are p roots $\lambda_1, \lambda_2, ..., \lambda_p$ in the complex plane such that, for any λ:

$$\lambda^p - \phi_1 \lambda^{p-1} - \phi_2 \lambda^{p-2} - ... - \phi_p = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_p).$$

Note that complex roots come in conjugate pairs $\lambda_i = a + bi$, $\lambda_j = a - bi$. Divide through by λ^p and let $z = 1/\lambda$:

$$1 - \phi_1 z - \phi_2 z^2 - ... - \phi_p z^p = (1 - \lambda_1 z)(1 - \lambda_2 z) \cdots (1 - \lambda_p z).$$

By setting $z = L$ we may write

$$\begin{aligned} & (1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p) y_t \\ = \; & (1 - \lambda_1 L)(1 - \lambda_2 L) \cdots (1 - \lambda_p L) y_t = c + \varepsilon_t. \end{aligned}$$

Assume $|\lambda_i| < 1$ for all i, i.e., all roots lie inside the unit circle on the complex plane (recall $|a + bi| = a^2 + b^2$, which is the length of the vector (a, b)). Solve for y_t:

$$y_t = \frac{1}{1 - \lambda_1 L} \frac{1}{1 - \lambda_2 L} \cdots \frac{1}{1 - \lambda_p L} (c + \varepsilon_t). \tag{4}$$

The $MA(\infty)$ representation is derived from (4) in two steps.

<u>Step 1 - Constant term</u>. Note that, for any constant α:

$$\frac{1}{1 - \lambda_i L} \alpha = \sum\nolimits_{j=0}^{\infty} (\lambda_i L)^j \alpha = \sum\nolimits_{j=0}^{\infty} \lambda_i^j \alpha = \frac{\alpha}{1 - \lambda_i}.$$

Thus:

$$\frac{1}{1 - \lambda_1 L} \frac{1}{1 - \lambda_2 L} \cdots \frac{1}{1 - \lambda_p L} c = \frac{c}{(1 - \lambda_1)(1 - \lambda_2) \cdots (1 - \lambda_p)} = \frac{c}{1 - \phi_1 - ... - \phi_p} \equiv \mu. \tag{5}$$

<u>Step 2 - MA coefficients</u>. Suppose the roots of (3) are distinct, i.e., $\lambda_i \neq \lambda_k$ for all i, k. Then the product term in (4) can be expanded with partial fractions:

$$\frac{1}{1 - \lambda_1 L} \frac{1}{1 - \lambda_2 L} \cdots \frac{1}{1 - \lambda_p L} = \sum\nolimits_{i=1}^{p} \frac{\omega_i}{1 - \lambda_i L},$$

where

$$\omega_i \equiv \frac{\lambda_i^{p-1}}{\prod\limits_{\substack{k=1 \\ k \neq i}}^{p} (\lambda_i - \lambda_k)}. \tag{6}$$

Furthermore, we can write:

$$\sum_{i=1}^{p} \frac{\omega_i}{1-\lambda_i L} = \sum_{i=1}^{p} \omega_i \sum_{j=0}^{\infty} (\lambda_i L)^j = \sum_{j=0}^{\infty} \sum_{i=1}^{p} \omega_i \lambda_i^j L^j = \sum_{j=0}^{\infty} \psi_j L^j,$$

where

$$\psi_j \equiv \sum_{i=1}^{p} \omega_i \lambda_i^j. \tag{7}$$

Thus:

$$\frac{1}{1-\lambda_1 L} \frac{1}{1-\lambda_2 L} \cdots \frac{1}{1-\lambda_p L} \varepsilon_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}.$$

Finally, combine terms:

$$y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}. \tag{8}$$

The restriction $|\lambda_i| < 1$ for all i implies that $\{\psi_j\}$ satisfies square summability, and so $\{y_t\}$ is stationary.

The following proposition summarizes this analysis.

Proposition. *Suppose (3) has distinct roots $\lambda_1, ..., \lambda_p$ satisfying $|\lambda_i| < 1$ for all i. Then the $AR(p)$ process (2) is stationary and has an $MA(\infty)$ representation (8), where μ is given by (5) and ψ_i is given by (6) and (7).*

Often $AR(p)$ processes are analyzed using the equation

$$1 - \phi_1 z - \phi_2 z^2 - ... - \phi_p z^p = 0.$$

In this case, the stationarity condition is that the roots lie outside of the unit circle, since the roots of this equation are the inverses of the earlier roots.

2 $VAR(p)$ processes

a Stationarity

Definition. *$\{Y_t\}$ is a $VAR(p)$ process (p^{th}-order vector autoregressive) if*

$$Y_t = C + \sum_{i=1}^{p} \Phi_i Y_{t-i} + \varepsilon_t, \tag{9}$$

where

$$Y_t = \begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{nt} \end{bmatrix}, \quad C = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}, \quad \Phi_i = \begin{bmatrix} \phi^i_{11} & \phi^i_{12} & \cdots & \phi^i_{1n} \\ \phi^i_{21} & \phi^i_{22} & \cdots & \phi^i_{2n} \\ \vdots & \vdots & & \vdots \\ \phi^i_{n1} & \phi^i_{n2} & \cdots & \phi^i_{nn} \end{bmatrix}, \quad \varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{nt} \end{bmatrix},$$

and ε_{it} *is vector white noise:*

$$E(\varepsilon_t) = 0_{n\times 1}, \quad E(\varepsilon_t \varepsilon_t') = \Omega,$$

where Ω *is a positive definite and symmtric* $n \times n$ *matrix, and*

$$E(\varepsilon_s \varepsilon_t') = 0_{n\times n} \text{ for all } s \neq t.$$

Ω is the variance-covariance matrix of the white noise vector. Positive definiteness means that $x'\Omega x > 0$ for all nonzero n-vectors x.

Stationarity. To evaluate stationarity of a $VAR(p)$, consider the equation

$$|I_n \lambda^p - \Phi_1 \lambda^{p-1} - \Phi_2 \lambda^{p-2} - ... - \Phi_p| = 0, \tag{10}$$

where $|\cdot|$ denotes the determinant and I_n is the $n \times n$ identity matrix:

$$I_n \equiv \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

The VAR is stationary if all solutions $\lambda = \lambda_i$ to (10) satisfy $|\lambda_i| < 1$. (Note that there are np roots of (10), possibly repeated, and complex roots come in conjugate pairs.) Equivalently, the VAR is stationary if all values of z satisfying

$$|I_n - \Phi_1 z - \Phi_2 z^2 - ... - \Phi_p z^p| = 0$$

lie outside of the unit circle.

The solutions to (10) can be computed using the following $np \times np$ matrix:

$$F = \begin{bmatrix} \Phi_1 & \Phi_2 & \Phi_3 & \cdots & \Phi_{p-1} & \Phi_p \\ I_n & 0_n & 0_n & \cdots & 0_n & 0_n \\ 0_n & I_n & 0_n & \cdots & 0_n & 0_n \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0_n & 0_n & 0_n & \cdots & I_n & 0_n \end{bmatrix},$$

where 0_n is an $n \times n$ matrix of zeros. It can be shown that the eigenvalues $\lambda_1, ..., \lambda_{np}$ of F are precisely the solutions to

$$|I_n \lambda^p - \Phi_1 \lambda^{p-1} - \Phi_2 \lambda^{p-2} - ... - \Phi_p| = 0.$$

To calculate the mean of a stationary VAR, take expectation:

$$EY_t = C + \sum\nolimits_{i=1}^{p} \Phi_i EY_{t-i}.$$

Stationarity implies $EY_t = \mu$ for all t. Thus:

$$\mu = (I_n - \sum\nolimits_{i=1}^{p} \Phi_i)^{-1} C.$$

The autocovariances are given by

$$\Gamma_j \equiv E(Y_t - \mu)(Y_{t-j} - \mu)'.$$

Each Γ_j is an $n \times n$ matrix, with γ_{ik}^j giving the covariance between y_{jt} and $y_{k,t-j}$.

b $MA(\infty)$ representation

To obtain an $MA(\infty)$ representation, express (9) as

$$(I_n - \sum\nolimits_{i=1}^{p} \Phi_i L^i) Y_t = C + \varepsilon_t.$$

Stationarity allows us to invert the lag polynomial:

$$(I_n - \sum\nolimits_{i=1}^{p} \Phi_i L^i)^{-1} = \sum\nolimits_{j=0}^{\infty} \Psi_j L^j. \tag{11}$$

Thus:

$$Y_t = \mu + \sum\nolimits_{j=0}^{\infty} \Psi_j \varepsilon_{t-j}. \tag{12}$$

The values of Ψ_j, $j = 0, 1, 2, ...$ may be obtained using the method of undetermined coefficients. Write (11) as:

$$I_n = (I_n - \sum\nolimits_{i=1}^{p} \Phi_i L^i) \sum\nolimits_{j=0}^{\infty} \Psi_j L^j. \tag{13}$$

The constant terms on each side of (13) must agree. Thus:

$$I_n = \Psi_0. \tag{14}$$

Further, since there are no powers of L on the LHS, the coefficient of L^j on the RHS must equal zero for each $j > 0$:

$$0 = \Psi_j - \Psi_{j-1}\Phi_1 - \Psi_{j-2}\Phi_2 - \cdots - \Psi_{j-p}\Phi_p, \quad j = 1, 2, ... \;. \tag{15}$$

Given the coefficients Φ_i and $\Psi_0 = I_n$, (15) may be iterated to compute MA coefficients $\Psi_1, \Psi_2, \Psi_3, ...$.

Nonuniqueness. Importantly, the $MA(\infty)$ representation of a VAR is nonunique. Let H be any nonsingular $n \times n$ matrix and define

$$u_t \equiv H\varepsilon_t.$$

Note that u_t is vector white noise:

$$E(u_t) = HE(\varepsilon_t) = 0_{n\times 1},$$

$$E(u_t u_t') = HE(\varepsilon_t \varepsilon_t')H' = H\Omega H',$$

$$E(u_s u_t') = HE(\varepsilon_s \varepsilon_t')H' = 0_n,$$

and $H\Omega H'$ is positive definite since $H'x$ is nonzero whenever x is. The $MA(\infty)$ representation can be expressed as

$$Y_t = \mu + \sum\nolimits_{j=0}^{\infty} \Psi_j H^{-1} H\varepsilon_{t-j} = \mu + \sum\nolimits_{j=0}^{\infty} \Theta_j u_{t-j},$$

where $\Theta_j \equiv \Psi_j H^{-1}$.

Note that in this case u_t is not the fundamental innovation. To obtain the $MA(\infty)$ representation in terms of the fundamental innovation we must impose the normalization $\Theta_0 = I_n$, i.e., $H = I_n$.

3 Identification of shocks

a Triangular factorization

We wish to assess how fluctuations in "more exogenous" variables affect "less exogenous" ones. One way to do this is to rearrange the vector of innovations ε_t into components that derive from "exogenous shocks" to the n variables. This can be accomplished using a triangular factorization of Ω.

For any positive definite symmetric matrix Ω, there exists a unique representation of the form

$$\Omega = ADA',$$

where A is a lower triangular matrix with 1's along the principal diagonal:

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_{21} & 1 & 0 & \cdots & 0 \\ a_{31} & a_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & 1 \end{bmatrix},$$

and D is a diagonal matrix:

$$D = \begin{bmatrix} d_{11} & 0 & 0 & \cdots & 0 \\ 0 & d_{22} & 0 & \cdots & 0 \\ 0 & 0 & d_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & d_{nn} \end{bmatrix},$$

with $d_{ii} > 0$ for $i = 1, ..., n$.

Use the factorization to define a vector of <u>exogenous shocks</u>:

$$u_t \equiv A^{-1}\varepsilon_t.$$

Substitute into the $MA(\infty)$ representation to obtain an alternative "structural" representation:

$$Y_t = \mu + \varepsilon_t + \sum\nolimits_{j=1}^{\infty} \Psi_j \varepsilon_{t-j} = \mu + Au_t + \sum\nolimits_{j=1}^{\infty} \Psi_j A u_{t-j} = \mu + \sum\nolimits_{j=0}^{\infty} \Theta_j u_{t-j},$$

where

$$\Theta_0 \equiv A, \qquad \Theta_j \equiv \Psi_j A, \quad j = 1, 2, \ldots .$$

Note that the shocks $u_{1t}, \ldots, u_{nt}$ are mutually uncorrelated:

$$E(u_t u_t') = A^{-1}E(\varepsilon_t\varepsilon_t')(A^{-1})' = A^{-1}\Omega(A')^{-1} = A^{-1}ADA'(A')^{-1} = D.$$

Thus:

$$Var(u_{it}) = d_{ii}, \quad Cov(u_{it}, u_{kt}) = 0.$$

To implement this approach, we order the variables from "most exogenous" to "least exogenous." This means that innovations to y_{it} are affected by the shocks $u_{1t}, \ldots, u_{it}$, but not by $u_{i+1,t}, \ldots, u_{nt}$.

<u>Bivariate case</u>. Let $n = 2$. (12) may be expressed as

$$\begin{bmatrix} \hat{y}_{1t} \\ \hat{y}_{2t} \end{bmatrix} = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} + \sum\nolimits_{j=1}^{\infty} \begin{bmatrix} \psi_{11}^j & \psi_{12}^j \\ \psi_{21}^j & \psi_{22}^j \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-j} \\ \varepsilon_{2,t-j} \end{bmatrix},$$

where $\hat{y}_{it} \equiv y_{it} - \mu_i$. Here y_{1t} is taken to be "most exogenous." Ω is factorized using the matrices

$$A = \begin{bmatrix} 1 & 0 \\ a_{21} & 1 \end{bmatrix}, \quad D = \begin{bmatrix} d_{11} & 0 \\ 0 & d_{22} \end{bmatrix}.$$

Thus,

$$\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ a_{21} & 1 \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \begin{bmatrix} u_{1t} \\ a_{21}u_{1t} + u_{2t} \end{bmatrix}.$$

Innovations to y_{1t} are driven by the exogenous shocks u_{1t}. Innovations to y_{2t} are driven by both innovations to y_{1t} and uncorrelated shocks u_{2t}.

Furthermore, for $j > 0$:

$$\Theta_j = \Psi_j A = \begin{bmatrix} \psi_{11}^j & \psi_{12}^j \\ \psi_{21}^j & \psi_{22}^j \end{bmatrix} \begin{bmatrix} 1 & 0 \\ a_{21} & 1 \end{bmatrix} = \begin{bmatrix} \psi_{11}^j + a_{21}\psi_{12}^j & \psi_{12}^j \\ \psi_{21}^j + a_{21}\psi_{22}^j & \psi_{22}^j \end{bmatrix}.$$

Alternative $MA(\infty)$ representation:

$$\begin{bmatrix} \hat{y}_{1t} \\ \hat{y}_{2t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ a_{21} & 1 \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} + \sum\nolimits_{j=1}^{\infty} \begin{bmatrix} \psi_{11}^j + a_{21}\psi_{12}^j & \psi_{12}^j \\ \psi_{21}^j + a_{21}\psi_{22}^j & \psi_{22}^j \end{bmatrix} \begin{bmatrix} u_{1,t-j} \\ u_{2,t-j} \end{bmatrix}.$$

We can use this to assess the effects of an exogenous shock to y_{1t}. Suppose the system begins in the <u>nonstochastic steady state</u>:

$$\begin{bmatrix} u_{1,t-j} \\ u_{2,t-j} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad j = 1, 2, \ldots \quad \Rightarrow \quad \begin{bmatrix} \hat{y}_{1,t-j} \\ \hat{y}_{2,t-j} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

At time t there is a positive shock to y_{1t}, and there are no shocks following this:

$$\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} u_{1,t+j} \\ u_{2,t+j} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad j = 1, 2, \ldots .$$

Then from the above representation we have

$$\begin{bmatrix} \hat{y}_{1t} \\ \hat{y}_{2t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ a_{21} & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ a_{21} \end{bmatrix},$$

$$\begin{bmatrix} \hat{y}_{1,t+j} \\ \hat{y}_{2,t+j} \end{bmatrix} = \begin{bmatrix} \psi_{11}^j + a_{21}\psi_{12}^j & \psi_{12}^j \\ \psi_{21}^j + a_{21}\psi_{22}^j & \psi_{22}^j \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \psi_{11}^j + a_{21}\psi_{12}^j \\ \psi_{21}^j + a_{21}\psi_{22}^j \end{bmatrix}.$$

Subsequent movements in each variable are driven by the direct effect of y_{1t} and an indirect effect coming through the response of y_{2t}. These are the <u>orthogonalized impulse-response functions</u>.

We can also assess the effects of a positive shock to y_{2t}, as captured by u_{2t}. In this case the change in y_{2t} is conditioned on u_{1t}, i.e., u_{2t} indicates the movement in y_{2t} that cannot be predicted after u_{1t} is known.

$$\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} u_{1,t+j} \\ u_{2,t+j} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad j = 1, 2, \dots ,$$

$$\begin{bmatrix} \hat{y}_{1t} \\ \hat{y}_{2t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ a_{21} & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$\begin{bmatrix} \hat{y}_{1,t+j} \\ \hat{y}_{2,t+j} \end{bmatrix} = \begin{bmatrix} \psi_{11}^j + a_{21}\psi_{12}^j & \psi_{12}^j \\ \psi_{21}^j + a_{21}\psi_{22}^j & \psi_{22}^j \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \psi_{12}^j \\ \psi_{22}^j \end{bmatrix}.$$

Note that u_{1t} affects y_{2t} in period t (as long as $a_{21} \neq 0$), but u_{2t} does not affect y_{1t}. This is the sense in which y_{1t} is "more exogenous."

Empirical implementation. For a given observed sample of size T, we can obtain OLS estimates $\hat{C}$ and $\hat{\Phi}_i$, $i = 1, ..., p$ by regressing Y_t on a constant terms and p lags $Y_{t-1}, ..., Y_{t-p}$. Estimated innovations are obtained from the OLS residuals:

$$\hat{\varepsilon}_t = Y_t - \hat{C} - \sum\nolimits_{i=1}^{p} \hat{\Phi}_i Y_{t-i}.$$

The variance-covariance matrix is estimated as

$$\hat{\Omega} = \frac{1}{T} \sum\nolimits_{t=1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_t'.$$

Estimates of the MA coefficients $\hat{\Psi}_j$, $j = 1, 2, ...$ can be obtained using the formulas derived above:

$$\hat{\Psi}_0 = I_n,$$

$$\hat{\Psi}_s - \hat{\Psi}_{s-1}\hat{\Phi}_1 - \hat{\Psi}_{s-2}\hat{\Phi}_2 - \hat{\Psi}_{s-p}\hat{\Phi}_p = 0, \quad s = 1, 2, \dots .$$

Orthogonalized impulse response functions are computed as

$$\hat{\Theta}_0 = A, \qquad \hat{\Theta}_j = \hat{\Psi}_j A, \quad j = 1, 2, \dots .$$

The coefficient $\hat{\theta}_{ik}^j$, the ik-element of $\hat{\Theta}_j$, gives the response of $\hat{y}_{i,t+j}$ to a one-unit positive shock to u_{kt}.

Cholesky factorization. For any positive definite symmetric matrix Ω, there exists a unique representation of the form

$$\Omega = PP',$$

where

$$P = AD^{1/2} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_{21} & 1 & 0 & \cdots & 0 \\ a_{31} & a_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \sqrt{d_{11}} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{d_{22}} & 0 & \cdots & 0 \\ 0 & 0 & \sqrt{d_{33}} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{d_{nn}} \end{bmatrix}.$$

This is called the Cholesky factorization.

Using the Cholesky factorization, the vector of exogenous shocks may be defined as:

$$v_t \equiv P^{-1}\varepsilon_t.$$

In the structural representation, A is simply replaced by P. Moreover, $E(v_t v_t') = I_n$, i.e., $Var(v_{it}) = 1$ for all i.

b Identification via long-run restrictions

Consider the following bivariate VAR process:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} y_{1,t-1} \\ 0 \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} + \sum_{j=1}^{\infty} \Psi_j \begin{bmatrix} \varepsilon_{1,t-j} \\ \varepsilon_{2,t-j} \end{bmatrix}, \tag{16}$$

with variance-covariance matrix Ω, where

$$\Psi_j = \begin{bmatrix} \psi_{11}^j & \psi_{12}^j \\ \psi_{21}^j & \psi_{22}^j \end{bmatrix}.$$

Note that forecasted values of y_{1t} are permanently affected by innovations, while the effects on y_{2t} die out when the Ψ_j's satisfy suitable stationary restrictions. This distinction can be used to identify "permanent" versus "transitory" shocks.

Write (16) as

$$\begin{bmatrix} \Delta y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} + \sum_{j=1}^{\infty} \Psi_j \begin{bmatrix} \varepsilon_{1,t-j} \\ \varepsilon_{2,t-j} \end{bmatrix}, \tag{17}$$

where $\Delta y_t = y_t - y_{t-1}$, and assume that (17) is stationary. We wish to obtain a structural representation

$$\begin{bmatrix} \Delta y_{1t} \\ y_{2t} \end{bmatrix} = \sum_{j=0}^{\infty} \Theta_j \begin{bmatrix} u_{1,t-j} \\ u_{2,t-j} \end{bmatrix}, \tag{18}$$

where u_{1t} and u_{2t} indicate permanent and transitory shocks, respectively, and

$$\Theta_j = \begin{bmatrix} \theta_{11}^j & \theta_{12}^j \\ \theta_{21}^j & \theta_{22}^j \end{bmatrix}.$$

Assume $Cov(u_{1t}, u_{2t}) = 0$ and $Var(u_{1t}) = Var(u_{2t}) = 1$, i.e., the variances of the shocks are normalized to unity. Furthermore, since $\Theta_0 u_t = \varepsilon_t$:

$$\Theta_0 E_t(u_t u_t')\Theta_0' = E_t(\varepsilon_t \varepsilon_t') \quad \Rightarrow \quad \Theta_0 \Theta_0' = \Omega.$$

Recall that the Cholesky factorization gives a unique lower triangular matrix satisfying $PP' = \Omega$. It follows that $\Theta_0 = P\Gamma$ for some orthogonal matrix Γ, i.e., Γ satisfies $\Gamma\Gamma' = I_2$. Orthogonality implies three restrictions on Γ, so we need one more restriction to identify Θ_0.

For the fourth restriction, assume that u_{2t} has no long-run effect on the level of y_{1t}, so that u_{2t} is transitory. For this to be true, all effects on Δy_{1t} must cancel out in the long run:

$$\sum_{j=0}^{\infty} \theta_{12}^j = 0.$$

Moreover, since $\Theta_j = \Psi_j \Theta_0$:

$$\theta_{12}^j = \psi_{11}^j \theta_{12}^0 + \psi_{12}^j \theta_{22}^0.$$

Substitute and rearrange:

$$\theta_{12}^0 \sum_{j=0}^{\infty} \psi_{11}^j + \theta_{22}^0 \sum_{j=0}^{\infty} \psi_{12}^j = 0.$$

This supplies one more restriction, and thus Θ_0 is identified.

4 Granger causality

Consider two stationary processes $\{y_{1t}\}$ and $\{y_{2t}\}$. Recall that the linear projection of y_{1t} on $y_{1,t-1}, y_{1,t-2}, ...$, denoted by $y_{1t}^{P\infty}$, minimizes MSE among all linear forecast rules $\sum_{j=1}^{\infty} g_j^{\infty} y_{1,t-j}$. We are interested in whether the variable y_{2t} can be used to obtain better predictions of y_{1t}. That is, does the linear projection of y_{1t} on $y_{1,t-1}, y_{1,t-2}, ...$ <u>and</u> $y_{2,t-1}, y_{2,t-2}, ...$ give a lower MSE than $y_{1t}^{P\infty}$? If not, then we say that the variable y_{2t} does <u>not</u> Granger-cause y_{1t}.

Suppose y_{1t} and y_{2t} are given by a bivariate VAR:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \sum_{i=1}^{p} \begin{bmatrix} \phi_{11}^i & \phi_{12}^i \\ \phi_{21}^i & \phi_{22}^i \end{bmatrix} \begin{bmatrix} y_{1,t-i} \\ y_{2,t-i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}.$$

Then y_{2t} does not Granger-cause y_{1t} if the coefficient matrices are lower triangular:

$$\begin{bmatrix} \phi_{11}^i & \phi_{12}^i \\ \phi_{21}^i & \phi_{22}^i \end{bmatrix} = \begin{bmatrix} \phi_{11}^i & 0 \\ \phi_{21}^i & \phi_{22}^i \end{bmatrix}, \quad i = 1, ..., p.$$

To test for Granger causality, estimate the first equation in the VAR with and without the parameter restriction

$$y_{1t} = c_1 + \sum_{i=1}^{p} \phi_{11}^i y_{1,t-i} + \eta_{1t},$$
$$y_{1t} = c_1 + \sum_{i=1}^{p} (\phi_{11}^i y_{1,t-i} + \phi_{12}^i y_{2,t-i}) + \varepsilon_{1t}.$$

Let $\hat{\eta}_{1t}$ and $\hat{\varepsilon}_{1t}$ be the fitted residuals and let the sample size be T. Define

$$RSS_0 = \sum_{t=1}^{T} \hat{\eta}_{1t}^2, \qquad RSS_1 = \sum_{t=1}^{T} \hat{\varepsilon}_{1t}^2.$$

Then for large T the following statistic has a χ^2 distribution:

$$S = \frac{T(RSS_0 - RSS_1)}{RSS_1}$$

If S exceeds a designated critical value for a $\chi^2(p)$ variable (e.g., 5%), then we reject the null hypothesis that y_{2t} does not Granger-cause y_{1t}, i.e., y_{2t} does help in forecasting y_{1t}.

Sources

Hamilton, Time Series Analysis, 1994, chs. 3,4,11

Blanchard and Quah,"The Dynamic Effects of Aggregate Demand and Supply Disturbances," AER, Sept. 1989