Lecture Notes on Vector Autoregression (VAR)[1]

Econ 624

March 11, 2021

[1] These notes are for instructional purposes only and are not to be distributed outside of the classroom.

VAR

- Consider the p^{th}- order vector autoregression (or $VAR(p)$ process)

$$\underset{m\times 1}{Y_t} = \underset{m\times 1}{\mu} + \underset{m\times m}{A_1}\,\underset{m\times 1}{Y_{t-1}} + \cdots + \underset{m\times m}{A_p}\,\underset{m\times 1}{Y_{t-p}} + \underset{m\times 1}{\varepsilon_t}$$

  where

$$\{\varepsilon_t\} \equiv i.i.d.\,(0, \Sigma_\varepsilon) \text{ with } \Sigma_\varepsilon > 0.$$

- **Remark:** Vector autoregression (VAR) is one of the workhorse models in emprical analysis of multiple time series. Empirical studies in economics rarely consider the VARMA (Vector Autoregression and Moving Average) model. Instead, people implicitly assume that a VAR of high enough order acts as a sufficient filter to transform the data into an $i.i.d.$ sequence, or, more generally, a martingale difference sequence.

VAR

- **Companion Form:** The idea here is to try to give a more convenient representation for higher-order vector autoregression. Define

$$\underset{mp\times 1}{\widetilde{Y}_t} = \begin{pmatrix} Y_t \\ Y_{t-1} \\ \vdots \\ Y_{t-p+1} \end{pmatrix}, \quad \underset{mp\times 1}{\widetilde{\mu}} = \begin{pmatrix} \mu \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\underset{mp\times mp}{A} = \begin{pmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I_m & 0 & \cdots & 0 & 0 \\ 0 & I_m & \ddots & \vdots & 0 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & I_m & 0 \end{pmatrix}, \quad \underset{mp\times 1}{\widetilde{\varepsilon}_t} = \begin{pmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

- **Companion Form (con't):** Using these notations, it is clear that we can write

$$\begin{aligned}\widetilde{Y}_t &= \begin{pmatrix} Y_t \\ Y_{t-1} \\ \vdots \\ Y_{t-p+1} \end{pmatrix} \\ &= \begin{pmatrix} \mu + A_1 Y_{t-1} + \cdots + A_p Y_{t-p} \\ Y_{t-1} \\ \vdots \\ Y_{t-p+1} \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}\end{aligned}$$

- **Companion Form (con't):** or

$$\widetilde{Y}_t = \begin{pmatrix} Y_t \\ Y_{t-1} \\ \vdots \\ Y_{t-p+1} \end{pmatrix}$$

$$= \begin{pmatrix} \mu \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I_m & 0 & \cdots & 0 & 0 \\ 0 & I_m & \ddots & \vdots & 0 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & I_m & 0 \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ Y_{t-2} \\ \vdots \\ Y_{t-p} \end{pmatrix}$$

$$+ \begin{pmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

VAR

- **Companion Form (con't):** It follows that we can represent a $VAR(p)$ process in a more convenient $VAR(1)$ form, i.e.,

$$\widetilde{Y}_t = \widetilde{\mu} + A\widetilde{Y}_{t-1} + \widetilde{\varepsilon}_t.$$

- For the companion form, we assume the following condition.
Stability Condition: Assume that all the eigenvalues are distinct and have modulus less than 1, or, equivalently,

$$\det\left(I_{mp} - Az\right) = 0 \Longrightarrow |z| > 1.$$

- **Remark:** Note that the assumption that all the eigenvalues are distinct, or have algebraic multiplicity equaling 1, is stronger than necessary but is made for convenience.

- **Further Remarks:**

(i) Note that

$$0 = \det\left(I_{mp} - Az\right) = z^{mp} \det\left(z^{-1} I_{mp} - A\right) = z^{mp} \det\left(\lambda I_{mp} - A\right)$$

by setting $\lambda = z^{-1}$. Since $z = 0$ is clearly not a solution of the determinantal equation

$$\det\left(I_{mp} - Az\right) = 0,$$

it follows that the assumption

$$\det\left(I_{mp} - Az\right) = 0 \Longrightarrow |z| > 1.$$

is equivalent to the assumption that

$$\det\left(\lambda I_{mp} - A\right) = 0 \Longrightarrow |\lambda| = \left|\frac{1}{z}\right| = \frac{1}{|z|} < 1.$$

(ii) Note that in empirical applications $A_1, A_2, ..., A_p$ are typically matrices whose elements are assumed to be real-valued, so that

$$\underset{mp\times mp}{A} = \begin{pmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I_m & 0 & \cdots & 0 & 0 \\ 0 & I_m & \ddots & \vdots & 0 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & I_m & 0 \end{pmatrix}$$

is a matrix, whose elements are real-valued. Moreover, A is a square matrix but it is not symmetric. A sufficient condition for this type of matrix to be diagonalizable is if all its eigenvalues are distinct, which is what we assume here.

- **Further Remarks (con't):**

(ii) In this case, we would have

$$A = T \Lambda T^{-1}$$

where

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_{mp} \end{pmatrix}$$

is an diagonal matrix whose diagonal elements are the eigenvalues of A and T is a nonsingular matrix whose columns are the eigenvectors of A.

- **Further Remarks (con't):**

(ii) It further follows that in this case

$$\begin{aligned} A^2 &= T\Lambda T^{-1} T\Lambda T^{-1} = T\Lambda^2 T^{-1} \\ A^3 &= T\Lambda T^{-1} T\Lambda T^{-1} T\Lambda T^{-1} = T\Lambda^2 T^{-1} T\Lambda T^{-1} = T\Lambda^3 T^{-1} \\ &\vdots \\ A^j &= T\Lambda^j T^{-1}. \end{aligned}$$

VAR

- **Further Remarks (con't):**

(iii) Since $|\lambda_k| < 1$ for every $k \in \{1, ..., mp\}$, this suggests that we can invert the matrix lag operator $I_{mp} - AL$ to obtain the vector moving average (VMA) representation

$$\begin{aligned} \widetilde{Y}_t &= (I_{mp} - AL)^{-1} (\widetilde{\mu} + \widetilde{\varepsilon}_t) \\ &= (I_{mp} - AL)^{-1} \widetilde{\mu} + \sum_{j=0}^{\infty} A^j \widetilde{\varepsilon}_{t-j}. \end{aligned}$$

Note that the matrix $I_{mp} - A$ is nonsingular since, by assumption, $z = 1$ is not a root of the determinantal equation $\det(I_{mp} - Az) = 0$. Hence,

$$\det(I_{mp} - A) \neq 0$$

and, thus, $I_{mp} - A$ is nonsingular.

(iv) The next thing we want to show is that $\det(I_{mp} - Az) = \det(I_m - A_1 z - \cdots - A_p z^p)$.

- **Further Remarks (con't):**

(iv) To see this, first write

$$
\begin{aligned}
& \det\left(I_{mp} - Az\right) \\
= \quad & \det\left\{ \begin{pmatrix} I_m & 0 & \cdots & 0 & 0 \\ 0 & I_m & \ddots & & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & & \ddots & I_m & 0 \\ 0 & 0 & \cdots & 0 & I_m \end{pmatrix} \right. \\
& \left. - \begin{pmatrix} A_1 z & A_2 z & \cdots & A_{p-1} z & A_p z \\ I_m z & 0 & \cdots & 0 & 0 \\ 0 & I_m z & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & I_m z & 0 \end{pmatrix} \right\}
\end{aligned}
$$

- **Further Remarks (con't):**

(iv) so that

$$
\begin{aligned}
& \det\left(I_{mp} - Az\right) \\
= \;& \det\left\{ \begin{pmatrix} I_m - A_1 z & -A_2 z & \cdots & -A_{p-1} z & -A_p z \\ -I_m z & I_m & 0 & \cdots & 0 \\ 0 & -I_m z & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & I_m & 0 \\ 0 & \cdots & 0 & -I_m z & I_m \end{pmatrix} \right\} \\
= \;& \det\left\{ \begin{matrix} B_{11}(z) & B_{12}(z) \\ B_{21}(z) & B_{22}(z) \end{matrix} \right\}
\end{aligned}
$$

where $\underset{m\times m}{B_{11}}(z) = I_m - A_1 z$,

$\underset{m\times m(p-1)}{B_{12}}(z) = \begin{pmatrix} -A_2 z & \cdots & -A_{p-1} z & -A_p z \end{pmatrix}$,

- **Further Remarks (con't):**

(iv)

$$\underset{m(p-1)\times m}{B_{21}(z)} = \begin{pmatrix} -I_m z \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

and

$$\underset{m(p-1)\times m(p-1)}{B_{22}(z)} = \begin{pmatrix} I_m & 0 & \cdots & 0 \\ -I_m z & \ddots & \ddots & \vdots \\ \ddots & \ddots & I_m & 0 \\ \cdots & 0 & -I_m z & I_m \end{pmatrix}$$

- **Further Remarks (con't):**

(iv) Next, note that

$$\begin{aligned} & \det \begin{pmatrix} B_{11}(z) & B_{12}(z) \\ B_{21}(z) & B_{22}(z) \end{pmatrix} \\ = \; & \det \left(B_{22}(z)\right) \det \left\{ B_{11}(z) - B_{12}(z) B_{22}(z)^{-1} B_{21}(z) \right\} \end{aligned}$$

Now,

$$\det \left(B_{22}(z)\right) = \det \left\{ \begin{pmatrix} I_m & 0 & \cdots & 0 \\ -I_m z & \ddots & \ddots & \vdots \\ \ddots & \ddots & I_m & 0 \\ \cdots & 0 & -I_m z & I_m \end{pmatrix} \right\} = 1$$

since $B_{22}(z)$ is a block lower triangular matrix, so its determinant is just the product of the determinants of the diagonal blocks.

- **Further Remarks (con't):**

(iv) To calculate $\det\left\{B_{11}(z) - B_{12}(z)\, B_{22}(z)^{-1} B_{21}(z)\right\}$, we first note that

$$B_{22}(z)^{-1} = \begin{pmatrix} I_m & 0 & \cdots & 0 \\ I_m z & \ddots & \ddots & \vdots \\ \vdots & \ddots & I_m & 0 \\ I_m z^{(p-2)} & \cdots & I_m z & I_m \end{pmatrix}$$

To show this, we shall give a simple argument using mathematical induction. To proceed, first consider the case $p = 2$. In this case,

$$B_{22}(z) = \begin{pmatrix} I_m & 0 \\ -I_m z & I_m \end{pmatrix}$$

- **Further Remarks (con't):**

(iv) Applying the formula for the inverse of block lower triangular matrices, i.e.,

$$\begin{pmatrix} A & 0 \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & 0 \\ -D^{-1}CA^{-1} & D^{-1} \end{pmatrix}$$

we have

$$\begin{aligned} \begin{pmatrix} I_m & 0 \\ -I_m z & I_m \end{pmatrix}^{-1} &= \begin{pmatrix} I_m & 0 \\ -I_m \left(-I_m z\right) I_m & I_m \end{pmatrix} \\ &= \begin{pmatrix} I_m & 0 \\ I_m z & I_m \end{pmatrix} \end{aligned}$$

- **Further Remarks (con't):**

(iv) Now, suppose that this holds for an $mq \times mq$ matrix $B_{22}^{(q)}(z)$, i.e.,

$$\underset{mq\times mq}{B_{22}^{(q)}(z)^{-1}} = \begin{pmatrix} I_m & 0 & \cdots & 0 \\ -I_m z & \ddots & \ddots & \vdots \\ \ddots & \ddots & I_m & 0 \\ \cdots & 0 & -I_m z & I_m \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} I_m & 0 & \cdots & 0 \\ I_m z & \ddots & \ddots & \vdots \\ \vdots & \ddots & I_m & 0 \\ I_m z^{(q-1)} & \cdots & I_m z & I_m \end{pmatrix}$$

- **Further Remarks (con't):**

(iv) Then, an $m(q+1) \times m(q+1)$ matrix $B_{22}^{(q+1)}(z)$ would have the have the following partitioned form

$$B_{22}^{(q+1)}(z) = \begin{pmatrix} B_{22}^{(q)}(z) & 0 \\ C & D \end{pmatrix}$$

where $D = I_m$ and

$$\underset{m \times mq}{C} = \left[\begin{array}{cc} \underset{m \times m(q-1)}{0} & \underset{m \times m}{-I_m z} \end{array} \right].$$

so that, upon applying the formula for the inverse of block lower triangular matrices given previously, we get

$$B_{22}^{(q+1)}(z)^{-1} = \begin{pmatrix} B_{22}^{(q)}(z)^{-1} & 0 \\ -D^{-1}CB_{22}^{(q)}(z)^{-1} & D^{-1} \end{pmatrix}$$

- **Further Remarks (con't):**

(iv) Now,

$$\begin{aligned}
-D^{-1}CB_{22}^{(q)}(z)^{-1} &= -I_m \left[\underset{m\times m}{0} \quad \cdots \quad \underset{m\times m}{0} \quad \underset{m\times m}{-I_m z} \right] \\
&\quad \times \begin{pmatrix} I_m & 0 & \cdots & 0 \\ I_m z & \ddots & \ddots & 0 \\ \vdots & \ddots & I_m & 0 \\ I_m z^{(q-1)} & \cdots & I_m z & I_m \end{pmatrix} \\
&= -I_m \left[\underset{m\times m}{-I_m z^q} \quad \cdots \quad \underset{m\times m}{-I_m z^2} \quad \underset{m\times m}{-I_m z} \right] \\
&= \left[\underset{m\times m}{I_m z^q} \quad \cdots \quad \underset{m\times m}{I_m z^2} \quad \underset{m\times m}{I_m z} \right]
\end{aligned}$$

- **Further Remarks (con't):**

(iv) so that

$$B_{22}^{(q+1)}(z)^{-1} = \begin{pmatrix} B_{22}^{(q)}(z)^{-1} & 0 \\ -D^{-1}CB_{22}^{(q)}(z)^{-1} & D^{-1} \end{pmatrix}$$

$$= \begin{pmatrix} I_m & 0 & \cdots & 0 & 0 \\ I_m z & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & I_m & 0 & \vdots \\ I_m z^{(q-1)} & \cdots & I_m z & I_m & 0 \\ I_m z^q & \cdots & I_m z^2 & I_m z & I_m \end{pmatrix}$$

- **Further Remarks (con't):**

(iv) Making use of this formula for the inverse of $B_{22}(z)$, we get

$$\begin{aligned} & B_{11}(z) - B_{12}(z) B_{22}(z)^{-1} B_{21}(z) \\ = \; & I_m - A_1 z \\ & - \left\{ \begin{pmatrix} -A_2 z & \cdots & -A_{p-1} z & -A_p z \end{pmatrix} \right. \\ & \left. \times \begin{pmatrix} I_m & 0 & \cdots & 0 \\ I_m z & \ddots & \ddots & \vdots \\ \vdots & \ddots & I_m & 0 \\ I_m z^{(p-2)} & \cdots & I_m z & I_m \end{pmatrix} \begin{pmatrix} -I_m z \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right\} \end{aligned}$$

- **Further Remarks (con't):**

(iv) By straightforward multiplication, we then obtain

$$
\begin{aligned}
& B_{11}(z) - B_{12}(z) B_{22}(z)^{-1} B_{21}(z) \\
= \quad & I_m - A_1 z - \begin{pmatrix} -A_2 z & \cdots & -A_{p-1} z & -A_p z \end{pmatrix} \begin{pmatrix} -I_m z \\ -I_m z^2 \\ \vdots \\ -I_m z^{(p-1)} \end{pmatrix} \\
= \quad & I_m - A_1 z - A_2 z^2 - \cdots - A_{p-1} z^{(p-1)} - A_p z^p
\end{aligned}
$$

- **Further Remarks (con't):**

(iv) from which it follows that

$$\begin{aligned}\det\left(I_{mp}-Az\right) &= \det\left(B_{22}\left(z\right)\right) \\ &\quad \times\det\left\{B_{11}\left(z\right)-B_{12}\left(z\right)B_{22}\left(z\right)^{-1}B_{21}\left(z\right)\right\} \\ &= \det\left\{B_{11}\left(z\right)-B_{12}\left(z\right)B_{22}\left(z\right)^{-1}B_{21}\left(z\right)\right\} \\ &= \det\left(I_m-A_1z-\cdots-A_pz^p\right)\end{aligned}$$

Hence, the stability condition can be equivalently stated as

$$\det\left(I_m-A_1z-\cdots-A_pz^p\right)=0\Longrightarrow|z|>1.$$

- **Further Remarks (con't):**

(v) To recover the vector moving average (VMA) representation for Y_t, we define

$$\underset{m\times mp}{J} = \begin{pmatrix} I_m & 0 & \cdots & 0 \end{pmatrix}$$

and note that

$$J\widetilde{Y}_t = \begin{pmatrix} I_m & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} Y_t \\ Y_{t-1} \\ \vdots \\ Y_{t-p+1} \end{pmatrix} = Y_t$$

VAR

- **Further Remarks (con't):**

(v) In addition, note that

$$
\begin{aligned}
J'J\widetilde{\varepsilon}_{t-j} &= \begin{pmatrix} I_m \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{pmatrix} I_m & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} \varepsilon_{t-j} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} I_m & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} \varepsilon_{t-j} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} \varepsilon_{t-j} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \widetilde{\varepsilon}_{t-j}.
\end{aligned}
$$

VAR

- **Further Remarks (con't):**

(v) Moreover,

$$J\tilde{\varepsilon}_{t-j} = \begin{pmatrix} I_m & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} \varepsilon_{t-j} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \varepsilon_{t-j}.$$

It follows that

$$\begin{aligned} \sum_{j=0}^{\infty} JA^j \tilde{\varepsilon}_{t-j} &= \sum_{j=0}^{\infty} JA^j J' J\tilde{\varepsilon}_{t-j} \\ &= \sum_{j=0}^{\infty} JA^j J' \varepsilon_{t-j}. \end{aligned}$$

- **Further Remarks (con't):**

(v) Hence,

$$\begin{aligned} Y_t &= J\widetilde{Y}_t \\ &= J\left(I_{mp} - A\right)^{-1}\widetilde{\mu} + \sum_{j=0}^{\infty} JA^j \widetilde{\varepsilon}_{t-j} \\ &= J\left(I_{mp} - A\right)^{-1}\widetilde{\mu} + \sum_{j=0}^{\infty} JA^j J' \varepsilon_{t-j} \\ &= \delta + \sum_{j=0}^{\infty} \Psi_j \varepsilon_{t-j}, \end{aligned}$$

where $\delta = J\left(I_{mp} - A\right)^{-1}\widetilde{\mu}$ and $\Psi_j = JA^j J'$ and where the matrix $I_{mp} - A$ is nonsingular because by assumption $\det\left(I_{mp} - Az\right) = 0 \Longrightarrow |z| > 1$ so that $z = 1$ is not a root and, thus, $\det\left(I_{mp} - A\right) \neq 0$.

VAR

- **Further Remarks (con't):**

(v) Next, note that, since

$$J'J\widetilde{\mu} = \begin{pmatrix} I_m \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{pmatrix} I_m & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} \mu \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} I_m & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} \mu \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} \mu \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \widetilde{\mu},$$

- **Further Remarks (con't):**

(v) Moreover,

$$J\widetilde{\mu} = \begin{pmatrix} I_m & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} \mu \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \mu,$$

so that we have

$$\begin{aligned} J\left(I_{mp} - A\right)^{-1} \widetilde{\mu} &= J\left(I_{mp} - A\right)^{-1} J' J\widetilde{\mu} \\ &= J\left(I_{mp} - A\right)^{-1} J' \mu. \end{aligned}$$

- **Further Remarks (con't):**

(v) Next, note that since $J = \begin{pmatrix} I_m & 0 & \cdots & 0 \end{pmatrix}$, $J\left(I_{mp} - A\right)^{-1} J'$ gives the $m \times m$ submatrix in the upper left-hand corner of the matrix $\left(I_{mp} - A\right)^{-1}$. Write $I_{mp} - A$ in partitioned form as

$$\begin{pmatrix} I_m - A_1 & -A_2 & \cdots & -A_{p-1} & -A_p \\ -I_m & I_m & 0 & \cdots & 0 \\ 0 & -I_m & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & I_m & 0 \\ 0 & \cdots & 0 & -I_m & I_m \end{pmatrix} = \begin{pmatrix} B & C \\ D & E \end{pmatrix}$$

VAR

- **Further Remarks (con't):**

(v) where $B = I_m - A_1$, $C = \begin{pmatrix} -A_2 & \cdots & -A_{p-1} & -A_p \end{pmatrix}$,

$$D = \begin{pmatrix} -I_m \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \text{ and}$$

$$E = \begin{pmatrix} I_m & 0 & \cdots & 0 \\ -I_m & \ddots & \ddots & \vdots \\ \ddots & \ddots & I_m & 0 \\ \cdots & 0 & -I_m & I_m \end{pmatrix}.$$

- **Further Remarks (con't):**

(v) Note that a general $mp \times mp$ matrix written in partitioned form

$$\begin{pmatrix} \underset{m \times m}{B} & \underset{m \times m(p-1)}{C} \\ \underset{m(p-1) \times m}{D} & \underset{m(p-1) \times m(p-1)}{E} \end{pmatrix}$$

has inverse given by

$$\begin{pmatrix} F^{-1} & -F^{-1}CE^{-1} \\ -E^{-1}DF^{-1} & E^{-1} + E^{-1}DF^{-1}CE^{-1} \end{pmatrix}$$

where

$$F = B - CE^{-1}D$$

VAR

- **Further Remarks (con't):**

(v) Previously, we have shown that

$$\begin{pmatrix} I_m & 0 & \cdots & 0 \\ -I_m z & \ddots & \ddots & \vdots \\ \ddots & \ddots & I_m & 0 \\ \cdots & 0 & -I_m z & I_m \end{pmatrix}^{-1} = \begin{pmatrix} I_m & 0 & \cdots & 0 \\ I_m z & \ddots & \ddots & \vdots \\ \vdots & \ddots & I_m & 0 \\ I_m z^{(p-2)} & \cdots & I_m z & I_m \end{pmatrix}$$

Setting $z = 1$, we, thus, deduce that

$$E^{-1} = \begin{pmatrix} I_m & 0 & \cdots & 0 \\ -I_m & \ddots & \ddots & \vdots \\ \ddots & \ddots & I_m & 0 \\ \cdots & 0 & -I_m & I_m \end{pmatrix}^{-1} = \begin{pmatrix} I_m & 0 & \cdots & 0 \\ I_m & \ddots & \ddots & \vdots \\ \vdots & \ddots & I_m & 0 \\ I_m & \cdots & I_m & I_m \end{pmatrix}$$

- **Further Remarks (con't):**

(v) Applying the above formula, we see that

$$
\begin{aligned}
& J\left(I_{mp}-A\right)^{-1} J' \\
= \; & F^{-1} = \left(B - CE^{-1}D\right)^{-1} \\
= \; & \left[\left(I_m - A_1\right) - \left\{ \begin{pmatrix} -A_2 & \cdots & -A_{p-1} & -A_p \end{pmatrix} \right.\right. \\
& \left.\left. \times \begin{pmatrix} I_m & 0 & \cdots & 0 \\ I_m & I_m & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ I_m & \cdots & I_m & I_m \end{pmatrix} \begin{pmatrix} -I_m \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right\}\right]^{-1}
\end{aligned}
$$

- **Further Remarks (con't):**

(v) Straightforward multiplication then shows that

$$
\begin{aligned}
& J\left(I_{mp} - A\right)^{-1} J' \\
= \quad & F^{-1} = \left(B - CE^{-1}D\right)^{-1} \\
= \quad & \left[\left(I_m - A_1\right) - \begin{pmatrix} -A_2 & \cdots & -A_{p-1} & -A_p \end{pmatrix} \begin{pmatrix} -I_m \\ -I_m \\ \vdots \\ -I_m \end{pmatrix} \right]^{-1} \\
= \quad & \left(I_m - A_1 - A_2 - \cdots - A_{p-1} - A_p\right)^{-1}
\end{aligned}
$$

- **Further Remarks (con't):**

(v) Hence,

$$\begin{aligned}\delta &= J\left(I_{mp}-A\right)^{-1}\widetilde{\mu} \\ &= J\left(I_{mp}-A\right)^{-1}J'J\widetilde{\mu} \\ &= J\left(I_{mp}-A\right)^{-1}J'\mu \\ &= \left(I_m - A_1 - A_2 - \cdots - A_{p-1} - A_p\right)^{-1}\mu\end{aligned}$$

Note again that $I_m - A_1 - A_2 - \cdots - A_{p-1} - A_p$ is nonsingular because we assume that $z = 1$ is not a root of the determinantal equation

$$\det\left(I_m - A_1 z - \cdots - A_p z^p\right) = 0$$

- To proceed, note first tht given a $VAR\left(p\right)$ process with an intercept

$$Y_t = \mu + A_1 Y_{t-1} + \cdots + A_p Y_{t-p} + \varepsilon_t,$$

we can transform it as follows

$$\begin{aligned}
& Y_t - \delta \\
= \ & \mu - \delta + A_1 Y_{t-1} + \cdots + A_p Y_{t-p} + \varepsilon_t \\
= \ & \mu - \left(I_m - A_1 - \cdots - A_p\right)^{-1} \mu + \left(A_1 + \cdots + A_p\right) \delta \\
& - \left(A_1 + \cdots + A_p\right) \delta + A_1 Y_{t-1} + \cdots + A_p Y_{t-p} + \varepsilon_t \\
= \ & \mu - \left(I_m - A_1 - \cdots - A_p\right)^{-1} \mu + \left(A_1 + \cdots + A_p\right) \delta \\
& + A_1 \left(Y_{t-1} - \delta\right) + \cdots + A_p \left(Y_{t-p} - \delta\right) + \varepsilon_t
\end{aligned}$$

Estimation of VAR

- $$\begin{aligned}
&= \mu - (I_m - A_1 - \cdots - A_p)^{-1}\mu \\
&\quad + (A_1 + \cdots + A_p)(I_m - A_1 - \cdots - A_p)^{-1}\mu \\
&\quad + A_1 (Y_{t-1} - \delta) + \cdots + A_p (Y_{t-p} - \delta) + \varepsilon_t \\
&= (I_m - A_1 - \cdots - A_p)(I_m - A_1 - \cdots - A_p)^{-1}\mu \\
&\quad - (I_m - A_1 - \cdots - A_p)^{-1}\mu \\
&\quad + (A_1 + \cdots + A_p)(I_m - A_1 - \cdots - A_p)^{-1}\mu \\
&\quad + A_1 (Y_{t-1} - \delta) + \cdots + A_p (Y_{t-p} - \delta) + \varepsilon_t \\
&= (I_m - A_1 - \cdots - A_p)^{-1}\mu - (I_m - A_1 - \cdots - A_p)^{-1}\mu \\
&\quad - (A_1 + \cdots + A_p)(I_m - A_1 - \cdots - A_p)^{-1}\mu \\
&\quad + (A_1 + \cdots + A_p)(I_m - A_1 - \cdots - A_p)^{-1}\mu \\
&\quad + A_1 (Y_{t-1} - \delta) + \cdots + A_p (Y_{t-p} - \delta) + \varepsilon_t \\
&= A_1 (Y_{t-1} - \delta) + \cdots + A_p (Y_{t-p} - \delta) + \varepsilon_t
\end{aligned}$$

Estimation of VAR

- Define $\underline{Y}_t = Y_t - \delta$ and we can rewrite this $VAR\,(p)$ process in the alternative form

$$\underline{Y}_t = A_1 \underline{Y}_{t-1} + \cdots + A_p \underline{Y}_{t-p} + \varepsilon_t$$

  Hence, to simplify notation, we shall in our subsequent discussion assume that $\delta = 0$, so that $\underline{Y}_t = Y_t$.

Estimation of VAR

- Next, we transpose the VAR equation to obtain

$$
\begin{aligned}
Y_t' &= Y_{t-1}'A_1' + \cdots + Y_{t-p}'A_p' + \varepsilon_t' \\
&= \begin{pmatrix} Y_{t-1}' & Y_{t-2}' & \cdots & Y_{t-p}' \end{pmatrix} \begin{pmatrix} A_1' \\ A_2' \\ \vdots \\ A_p' \end{pmatrix} + \varepsilon_t' \\
&= X_t' B + \varepsilon_t'
\end{aligned}
$$

where $\underset{1\times mp}{X_t'} = \begin{pmatrix} Y_{t-1}' & Y_{t-2}' & \cdots & Y_{t-p}' \end{pmatrix}$ and

$$
\underset{mp\times m}{B} = \begin{pmatrix} A_1' \\ A_2' \\ \vdots \\ A_p' \end{pmatrix}
$$

Estimation of VAR

- Define

$$\underset{T\times m}{Y}=\begin{pmatrix} Y_1' \\ Y_2' \\ \vdots \\ Y_T' \end{pmatrix},\ \underset{T\times mp}{X}=\begin{pmatrix} X_1' \\ X_2' \\ \vdots \\ X_T' \end{pmatrix},\text{ and }\underset{T\times m}{E}=\begin{pmatrix} \varepsilon_1' \\ \varepsilon_2' \\ \vdots \\ \varepsilon_p' \end{pmatrix}$$

and we can write the model more succinctly as

$$Y = XB + E$$

Moreover, by vectorizing both sides of the equation above using the identity

$$vec\left(ABC\right) = \left(C' \otimes A\right) vec\left(B\right),$$

VAR

- we have

$$\begin{aligned} y &= vec\left(Y\right) \\ &= \left(I_m \otimes X\right) vec\left(B\right) + vec\left(E\right) \\ &= \left(I_m \otimes X\right) \beta + \xi, \end{aligned}$$

where

$$\underset{mT\times 1}{vec\left(Y\right)} = \begin{pmatrix} Y_{1\cdot} \\ Y_{2\cdot} \\ \vdots \\ Y_{m\cdot} \end{pmatrix}, \underset{m^2 p\times 1}{\beta} = vec\left(B\right), \text{ and } \underset{mT\times 1}{\xi} = \begin{pmatrix} \varepsilon_{1\cdot} \\ \varepsilon_{2\cdot} \\ \vdots \\ \varepsilon_{m\cdot} \end{pmatrix}$$

and where

$$\underset{T\times 1}{Y_{i\cdot}} = \begin{pmatrix} Y_{i1} \\ Y_{i2} \\ \vdots \\ Y_{iT} \end{pmatrix} \text{ and } \underset{T\times 1}{\varepsilon_{i\cdot}} = \begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \vdots \\ \varepsilon_{iT} \end{pmatrix} \text{ for } i = 1, ..., m.$$

- **Question:** To estimate β in the regression $y = (I_m \otimes X)\beta + \xi$, should we use generalized least squares (GLS) or ordinary least squares (OLS)? To answer this question, observe first that, by assumption,

$$E(\xi) = \begin{pmatrix} E[\varepsilon_{1\cdot}] \\ E[\varepsilon_{2\cdot}] \\ \vdots \\ E[\varepsilon_{m\cdot}] \end{pmatrix} = 0$$

In addition,

$$\begin{aligned} VC(\xi) &= E[\xi\xi'] \\ &= \begin{pmatrix} E[\varepsilon_{1\cdot}\varepsilon'_{1\cdot}] & E[\varepsilon_{1\cdot}\varepsilon'_{2\cdot}] & \cdots & E[\varepsilon_{1\cdot}\varepsilon'_{m\cdot}] \\ E[\varepsilon_{2\cdot}\varepsilon'_{1\cdot}] & E[\varepsilon_{2\cdot}\varepsilon'_{2\cdot}] & \cdots & E[\varepsilon_{2\cdot}\varepsilon'_{m\cdot}] \\ \vdots & \vdots & \ddots & \vdots \\ E[\varepsilon_{m\cdot}\varepsilon'_{1\cdot}] & E[\varepsilon_{m\cdot}\varepsilon'_{2\cdot}] & \cdots & E[\varepsilon_{m\cdot}\varepsilon'_{m\cdot}] \end{pmatrix} \end{aligned}$$

VAR

- Now, let $\sigma_{ij} = E\left[\varepsilon_{i1}\varepsilon_{j1}\right]$ denote the $(i,j)^{th}$ element of Σ_{ε}. For $i,j = 1, ..., m$; we have

$$
\begin{aligned}
E\left[\varepsilon_{i\cdot}\varepsilon_{j\cdot}'\right] &= E\left[\begin{pmatrix}\varepsilon_{i1}\\ \varepsilon_{i2}\\ \vdots\\ \varepsilon_{iT}\end{pmatrix}\begin{pmatrix}\varepsilon_{j1} & \varepsilon_{j2} & \cdots & \varepsilon_{jT}\end{pmatrix}\right] \\
&= \begin{pmatrix} E\left[\varepsilon_{i1}\varepsilon_{j1}\right] & E\left[\varepsilon_{i1}\varepsilon_{j2}\right] & \cdots & E\left[\varepsilon_{i1}\varepsilon_{jT}\right] \\ E\left[\varepsilon_{i2}\varepsilon_{j1}\right] & E\left[\varepsilon_{i2}\varepsilon_{j2}\right] & \cdots & E\left[\varepsilon_{i2}\varepsilon_{jT}\right] \\ \vdots & \vdots & \ddots & \vdots \\ E\left[\varepsilon_{iT}\varepsilon_{j1}\right] & E\left[\varepsilon_{iT}\varepsilon_{j2}\right] & \cdots & E\left[\varepsilon_{iT}\varepsilon_{jT}\right] \end{pmatrix} \\
&= \begin{pmatrix} \sigma_{ij} & 0 & \cdots & 0 \\ 0 & \sigma_{ij} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_{ij} \end{pmatrix} = \sigma_{ij} I_T .
\end{aligned}
$$

- It follows that

$$\begin{aligned} VC(\xi) &= E\left[\xi\xi'\right] \\ &= \begin{pmatrix} \sigma_{11} I_T & \sigma_{12} I_T & \cdots & \sigma_{1m} I_T \\ \sigma_{21} I_T & \sigma_{22} I_T & \cdots & \sigma_{2m} I_T \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1} I_T & \sigma_{m2} I_T & \cdots & \sigma_{mm} I_T \end{pmatrix} \\ &= (\Sigma_\varepsilon \otimes I_T) \end{aligned}$$

Note that, by the fact that the error covariance matrix Σ_ε is symmetric, we have that $\sigma_{ij} = \sigma_{ji}$ for all $i \neq j$.

VAR

- Next, assume initially that Σ_ε is known. In this case the formula for the GLS estimator of β in the regression $vec\left(Y\right) = \left(I_m \otimes X\right)\beta + \xi$ is given by

$$
\begin{aligned}
& \widehat{\beta}_{GLS} \\
= \; & \left[\left(I_m \otimes X\right)'\left(\Sigma_\varepsilon \otimes I_T\right)^{-1}\left(I_m \otimes X\right)\right]^{-1} \\
& \times \left(I_m \otimes X\right)'\left(\Sigma_\varepsilon \otimes I_T\right)^{-1} vec\left(Y\right) \\
= \; & \left[\left(I_m \otimes X\right)'\left(\Sigma_\varepsilon^{-1} \otimes I_T\right)\left(I_m \otimes X\right)\right]^{-1} \\
& \times \left(I_m \otimes X\right)'\left(\Sigma_\varepsilon^{-1} \otimes I_T\right) vec\left(Y\right) \\
& \left(\text{since } \left(A \otimes B\right)^{-1} = \left(A^{-1} \otimes B^{-1}\right) \text{ assuming that both}\right. \\
& \left. A \text{ and } B \text{ are non-singular and, thus, invertible}\right)
\end{aligned}
$$

- It follows that

$$
\begin{aligned}
& \widehat{\beta}_{GLS} \\
= \; & \left[\left(I_m \otimes X'\right)\left(\Sigma_\varepsilon^{-1} \otimes I_T\right)\left(I_m \otimes X\right)\right]^{-1} \\
& \times \left(I_m \otimes X'\right)\left(\Sigma_\varepsilon^{-1} \otimes I_T\right) vec\left(Y\right) \\
& \left(\text{since } \left(A \otimes B\right)' = \left(A' \otimes B'\right)\right) \\
= \; & \left[\Sigma_\varepsilon^{-1} \otimes X'X\right]^{-1}\left(\Sigma_\varepsilon^{-1} \otimes X'\right) vec\left(Y\right) \\
& \left(\text{since } \left(A \otimes B\right)\left(C \otimes D\right) = \left(AC \otimes BD\right)\right) \\
= \; & \left[\Sigma_\varepsilon \otimes \left(X'X\right)^{-1}\right]\left(\Sigma_\varepsilon^{-1} \otimes X'\right) vec\left(Y\right) \\
= \; & \left(I_m \otimes \left(X'X\right)^{-1} X'\right) vec\left(Y\right)
\end{aligned}
$$

VAR

- It follows that

$$
\begin{aligned}
& \widehat{\beta}_{GLS} \\
= & \left(I_m \otimes (X'X)^{-1} X'\right) vec\,(Y) \\
= & \begin{pmatrix} (X'X)^{-1} X' & 0 & \cdots & 0 \\ 0 & (X'X)^{-1} X' & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & (X'X)^{-1} X' \end{pmatrix} \begin{pmatrix} Y_{1\cdot} \\ Y_{2\cdot} \\ \vdots \\ Y_{m\cdot} \end{pmatrix} \\
= & \begin{pmatrix} (X'X)^{-1} X' Y_{1\cdot} \\ (X'X)^{-1} X' Y_{2\cdot} \\ \vdots \\ (X'X)^{-1} X' Y_{m\cdot} \end{pmatrix} = \widehat{\beta}_{OLS}.
\end{aligned}
$$

- **Remark:** Hence, GLS = OLS in this case. Note that this is just a special case of the result from seemingly unrelated regression (SUR) where GLS is the same as equation-by-equation OLS if the same set of regressors enters into every equation.

Asymptotic Normality of OLS Estimator of VAR Parameters

- **Assumptions:** Suppose that

1. $$\det\left(I_m - A_1 z - \cdots - A_p z^p\right) = 0 \Longrightarrow |z| > 1$$
2. $\{\varepsilon_t\} \equiv i.i.d.\,(0, \Sigma_\varepsilon)$, where there exists positive constant C such that $0 < 1/C \leq \lambda_{\min}\left(\Sigma_\varepsilon\right) \leq \lambda_{\max}\left(\Sigma_\varepsilon\right) \leq C < \infty$.

Asymptotic Normality of OLS Estimator of VAR Parameters

- Under Assumptions 1 and 2, one can show that

$$\sqrt{T}\left(\widehat{\beta}_{OLS}-\beta\right)\xrightarrow{d}N\left(0,V\right)\text{ as }T\rightarrow\infty,$$

where

$$\beta=vec\left(B\right)\text{ and }V=\left(\Sigma_{\varepsilon}\otimes M^{-1}\right)$$

and where

$$M=p\lim_{T\rightarrow\infty}\frac{X'X}{T}$$

with

$$\underset{mp\times m}{B}=\begin{pmatrix}A_1'\\A_2'\\\vdots\\A_p'\end{pmatrix},\ \underset{T\times mp}{X}=\begin{pmatrix}X_1'\\X_2'\\\vdots\\X_T'\end{pmatrix},\text{ and }\underset{mp\times 1}{X_t}=\begin{pmatrix}Y_{t-1}\\Y_{t-2}\\\vdots\\Y_{t-p}\end{pmatrix}.$$

Asymptotic Normality of OLS Estimator of VAR Parameters

- **Remark:** Note that we can write

$$
\begin{aligned}
\widehat{\beta}_{OLS} &= \left(I_m \otimes (X'X)^{-1} X'\right) vec\,(Y) \\
&= \left(I_m \otimes (X'X)^{-1} X'\right) [(I_m \otimes X)\,\beta + \xi] \\
&= \left(I_m \otimes (X'X)^{-1} X'X\right) \beta + \left(I_m \otimes (X'X)^{-1} X'\right) \xi \\
&= (I_m \otimes I_{mp})\,\beta + \left(I_m \otimes (X'X)^{-1} X'\right) \xi \\
&= \beta + \left(I_m \otimes (X'X)^{-1}\right) (I_m \otimes X')\,\xi
\end{aligned}
$$

Asymptotic Normality of OLS Estimator of VAR Parameters

- **Remark (con't):** It follows that

$$\begin{aligned}\sqrt{T}\left(\widehat{\beta}_{OLS}-\beta\right) &= \left[I_m \otimes \left(\frac{X'X}{T}\right)^{-1}\right]\frac{\left(I_m \otimes X'\right)\xi}{\sqrt{T}} \\ &= \left[I_m \otimes M^{-1}\right]\frac{\left(I_m \otimes X'\right)\xi}{\sqrt{T}} + o_p\left(1\right).\end{aligned}$$

Asymptotic Normality of OLS Estimator of VAR Parameters

- **Remark (con't):** Now, applying a CLT for strictly stationary and ergodic process and the Cramér-Wold device, we have

$$\frac{\left(I_m \otimes X'\right)\xi}{\sqrt{T}} \xrightarrow{d} N\left(0, \Gamma\right)$$

where

$$\begin{aligned}
\Gamma &= p\lim_{T\to\infty} \frac{\left(I_m \otimes X'\right) E\left[\xi\xi'\right]\left(I_m \otimes X\right)}{T} \\
&= p\lim_{T\to\infty} \frac{\left(I_m \otimes X'\right)\left(\Sigma_\varepsilon \otimes I_T\right)\left(I_m \otimes X\right)}{T} \\
&= \left(\Sigma_\varepsilon \otimes p\lim_{T\to\infty} \frac{X'X}{T}\right) \\
&= \left(\Sigma_\varepsilon \otimes M\right)
\end{aligned}$$

Asymptotic Normality of OLS Estimator of VAR Parameters

- **Remark (con't):** It follows by a generalized Cramér convergence theorem that

$$\sqrt{T}\left(\widehat{\beta}_{OLS}-\beta\right)\xrightarrow{d}\left(I_m\otimes M^{-1}\right)N\left(0,\Sigma_\varepsilon\otimes M\right)\equiv N\left(0,\Sigma_\varepsilon\otimes M^{-1}\right).$$

Impulse Response Analysis

- Again, we write $VAR\,(p)$ model in companion form, i.e.,

$$\widetilde{Y}_t = A\widetilde{Y}_{t-1} + \widetilde{\varepsilon}_t$$

where

$$\underset{mp\times 1}{\widetilde{Y}_t} = \begin{pmatrix} Y_t \\ Y_{t-1} \\ \vdots \\ Y_{t-p+1} \end{pmatrix},$$

$$\underset{mp\times mp}{A} = \begin{pmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I_m & 0 & \cdots & 0 & 0 \\ 0 & I_m & \ddots & \vdots & 0 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & I_m & 0 \end{pmatrix}, \quad \underset{mp\times 1}{\widetilde{\varepsilon}_t} = \begin{pmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Impulse Response Analysis

- Consider the following thought experiment:

$$\underset{m\times 1}{Y_t} = 0 \text{ for all } t \in \mathbb{Z} \text{ such that } t < 0,$$
$$\underset{m\times 1}{\varepsilon_t} = 0 \text{ for all } t \in \mathbb{Z} \text{ such that } t > 0,$$
$$\underset{m\times 1}{Y_0} = \underset{m\times 1}{\varepsilon_0} = \underset{m\times 1}{e_\ell} \text{ - i.e., a unit shock in the } \ell^{th} \text{ component at } t = 0$$

where

$$\underset{m\times 1}{e_\ell} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

i.e., a vector with 1 in the ℓ^{th} component and 0 elsewhere.

Impulse Response Analysis

- In the companion form, this can be written as

$$\underset{mp\times 1}{\widetilde{Y}_0} = \begin{pmatrix} Y_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \underset{mp\times 1}{\widetilde{\varepsilon}_0} = \begin{pmatrix} \varepsilon_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} e_\ell \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Impulse Response Analysis

Tracing the effect of this unit shock over time, we see that

- At time $t = 1$,

$$\underset{mp\times 1}{\widetilde{Y}_1} = \underset{mp\times mp}{A}\,\underset{mp\times 1}{\widetilde{Y}_0}$$

- At time $t = 2$,

$$\widetilde{Y}_2 = A\widetilde{Y}_1 = A^2\widetilde{Y}_0$$

  Continuing on, we have

- At time $t = j$, where j is a positive integer

$$\widetilde{Y}_j = A^j\widetilde{Y}_0$$

- Next, note that since

$$\widetilde{Y}_0 = \widetilde{\varepsilon}_0 = \begin{pmatrix} e_\ell \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

it is easy to see that $A^j \widetilde{Y}_0$ is just the ℓ^{th} column of A^j.

- Again, let

$$\underset{m \times mp}{J} = \begin{pmatrix} I_m & 0 & \cdots & 0 \end{pmatrix}$$

and let

$$\Psi_j = JA^j J' \text{ for } j = 0, 1, 2, ...$$

be the cofficient matrices in the $VMA(\infty)$ representation

$$Y_t = \sum_{j=0}^{\infty} \Psi_j \varepsilon_{t-j}.$$

Impulse Response Analysis

- Next, note that since

$$\widetilde{Y}_0 = \widetilde{\varepsilon}_0 = \begin{pmatrix} e_\ell \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

it is easy to see that $A^j \widetilde{Y}_0$ is just the ℓ^{th} column of A^j.

- Again, let

$$\underset{m \times mp}{J} = \begin{pmatrix} I_m & 0 & \cdots & 0 \end{pmatrix}$$

and let

$$\Psi_j = JA^j J' \text{ for } j = 0, 1, 2, \ldots.$$

be the cofficient matrices in the $VMA(\infty)$ representation

$$Y_t = \sum_{j=0}^{\infty} \Psi_j \varepsilon_{t-j}.$$

Impulse Response Analysis

- Let $\psi_{k\ell,j}$ denote the $(k,\ell)^{th}$ element of Ψ_j. Note that $\psi_{k\ell,j}$ represents the reaction of the k^{th} variable to a unit shock to the ℓ^{th} variable j periods ago, provided, of course, the effect is not contaminated by other shocks to the system in the interim. Thus, the coefficients of the $VMA(\infty)$ representation give the impulse response coefficients. Because the ε_t are just the one-step ahead forecast errors of the VAR process, the shocks considered here may be regarded as forecast error and the impulse responses are sometimes referred to as forecast error responses.
- **Proposition (Zero Impulse Responses):** If $\{Y_t\}$ is a m-variate stable $VAR(p)$ process; then, for $k \neq \ell$,

$$\psi_{k\ell,j} = 0 \text{ for } j = 1, 2, ...$$

is equivalent to

$$\psi_{k\ell,j} = 0 \text{ for } j = 1, 2, ..., p(m-1).$$

- **Accumulated Impulse Response:** The accumulated impulse response over n periods is given by

$$\sum_{j=0}^{n} \Psi_j$$

This quantity is also sometimes called the n^{th} interim multiplier.

- **Accumulated Impulse Response (con't):** Taking $n \to \infty$, we get the long-run effect or total multiplier

$$\lim_{n\to\infty} \sum_{j=0}^{n} \Psi_j = \sum_{j=0}^{\infty} \Psi_j.$$

Note further that

$$\sum_{j=0}^{\infty} \Psi_j = \left(I_m - A_1 - \cdots - A_p\right)^{-1},$$

where again we know that the inverse exists because of the condition that

$$\det\left(I_m - A_1 z - \cdots - A_p z^p\right) = 0 \Longrightarrow |z| > 1.$$

In addition, observe that the total multiplier in this case can be easily estimated by

$$\left(I_m - \widehat{A}_1 - \cdots - \widehat{A}_p\right)^{-1}.$$

Responses to Orthogonal Impulses

- The previous analysis is a bit problematic in that it assumes that a shock occurs only in one variable at a time. Such an assumption may be reasonable if ε_{kt} is uncorrelated with $\varepsilon_{\ell t}$ for all $k \neq \ell$ and for all t. However, in general, we assume ε_t to have a variance-covariance matrix Σ_ε which is not restricted to be a diagonal matrix, so that ε_{kt} is not assumed to be uncorrelated with $\varepsilon_{\ell t}$ for all $k \neq \ell$.
- Since Σ_ε is assumed to be a positive definite matrix, a way around this problem is to consider the Cholesky decomposition

$$\Sigma_\varepsilon = PP'$$

where P is a lower triangular matrix with positive diagonal elements, i.e.,

$$P = \begin{pmatrix} p_{11} & 0 & \cdots & 0 \\ p_{21} & p_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{pmatrix} \text{ with } p_{kk} > 0 \text{ for } k \in \{1, ..., m\}.$$

Responses to Orthogonal Impulses

- Note that
$$\det(P) = \prod_{k=1}^{m} p_{kk} > 0$$
so that P is, of course, nonsingular. Using this decomposition, we can then rewrite the $VMA(\infty)$ representation as
$$Y_t = \sum_{j=0}^{\infty} \Psi_j P P^{-1} \varepsilon_{t-j} = \sum_{j=0}^{\infty} \Theta_j u_{t-j},$$
where $\Theta_j = \Psi_j P$ and $u_{t-j} = P^{-1}\varepsilon_{t-j}$. Under this transformation,
$$\begin{aligned} E\left[u_t u_t'\right] &= E\left[P^{-1}\varepsilon_t \varepsilon_t' P'^{-1}\right] \\ &= P^{-1} E\left[\varepsilon_t \varepsilon_t'\right] P'^{-1} \\ &= P^{-1} \Sigma_\varepsilon P'^{-1} \\ &= P^{-1} P P' P'^{-1} \\ &= I_m \end{aligned}$$
so that u_{kt} is uncorrelated with $u_{\ell t}$ for all $k \neq \ell$ and for all t.

Responses to Orthogonal Impulses

- Now, relating u_t to ε_t, we see that

$$\varepsilon_t = Pu_t$$

or

$$\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{mt} \end{pmatrix} = \begin{pmatrix} p_{11} & 0 & \cdots & 0 \\ p_{21} & p_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{pmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \\ \vdots \\ u_{mt} \end{pmatrix}$$

so that we have

$$\begin{aligned} \varepsilon_{1t} &= p_{11}u_{1t} \\ \varepsilon_{2t} &= p_{21}u_{1t} + p_{22}u_{2t} = \frac{p_{21}}{p_{11}}\varepsilon_{1t} + p_{22}u_{2t} \\ &\vdots \\ \varepsilon_{mt} &= p_{m1}u_{1t} + \cdots + p_{mm}u_{mt} = \frac{p_{m1}}{p_{11}}\varepsilon_{1t} + \cdots + p_{mm}u_{mt} \end{aligned}$$

Responses to Orthogonal Impulses

- Hence, the way in which the Cholesky decomposition is carried out assumes a particular causal ordering, so that, in this case, the implicit assumption is that movement in ε_{1t} can cause movement in ε_{2t} contemporaneously but not the other way around and so on.
- Note further that we can estimate Σ_ε using

$$\widehat{\Sigma}_\varepsilon = \frac{\left(Y - X\widehat{B}\right)'\left(Y - X\widehat{B}\right)}{T}$$

and perform a Cholesky decomposition on $\widehat{\Sigma}_\varepsilon$ to obtain

$$\widehat{\Sigma}_\varepsilon = \widehat{P}\widehat{P}'.$$

Responses to Orthogonal Impulses

- Moreover, Ψ_j can be estimated by

$$\widehat{\Psi}_j = J\widehat{A}^j J'$$

with

$$\widehat{A} = \begin{pmatrix} \widehat{A}_1 & \widehat{A}_2 & \cdots & \widehat{A}_{p-1} & \widehat{A}_p \\ I_m & 0 & \cdots & 0 & 0 \\ 0 & I_m & \ddots & \vdots & 0 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & I_m & 0 \end{pmatrix}$$

so that

$$\widehat{\Theta}_j = \widehat{\Psi}_j \widehat{P}.$$