Covariance-Stationary Vector Processes

DENG Yanfei
dengyf@fudan.edu.cn
www.scholat.com/idengyf.en

December 7, 2019

1 Vector Autoregressions (VARs)

1.1 AR(p)→AR(1)←Univariable

Recall Lec 1-2.

1.2 VAR(p)→VAR(1)←Multivariable

A pth-order vector autoregression[1]

$$\begin{cases} y_t \overset{\text{AR(p)}}{=} \underbrace{c}_{(1\times1)} + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \underbrace{\phi_i}_{(1\times1)} y_{t-i} + \cdots + \phi_p y_{t-p} + \underbrace{\epsilon_t}_{1\times1} \begin{cases} \mathbb{E}\epsilon_t = 0 \\ \mathbb{E}(\epsilon_t\epsilon_\tau) \begin{cases} \sigma^2 & \text{for } t=\tau \\ 0 & \text{otherwise} \end{cases} \end{cases} \\ \mathbf{y}_t \overset{\text{VAR(p)}}{=} \underbrace{\mathbf{c}}_{(n\times1)} + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \mathbf{\Phi}_2 \mathbf{y}_{t-2} + \cdots + \underbrace{\mathbf{\Phi}_i}_{(n\times n)} \mathbf{y}_{t-i} + \cdots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \underbrace{\boldsymbol{\epsilon}_t}_{n\times1} \begin{cases} \mathbb{E}\boldsymbol{\epsilon}_t = \mathbf{0} \\ \mathbb{E}(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_\tau') \begin{cases} \underbrace{\boldsymbol{\sigma}^2}_{(n\times n)} & \text{for } t=\tau \\ \mathbf{0} & \text{otherwise} \end{cases} \end{cases} \end{cases}$$

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{nt} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} & \cdots & \phi_{1n}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} & \cdots & \phi_{2n}^{(1)} \\ \vdots & \vdots & \cdots & \vdots \\ \phi_{n1}^{(1)} & \phi_{n2}^{(1)} & \cdots & \phi_{nn}^{(1)} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{n,t-1} \end{bmatrix} + \cdots + \begin{bmatrix} \phi_{11}^{(p)} & \phi_{12}^{(p)} & \cdots & \phi_{1n}^{(p)} \\ \phi_{21}^{(p)} & \phi_{22}^{(p)} & \cdots & \phi_{2n}^{(p)} \\ \vdots & \vdots & \cdots & \vdots \\ \phi_{n1}^{(p)} & \phi_{n2}^{(p)} & \cdots & \phi_{nn}^{(p)} \end{bmatrix} \begin{bmatrix} y_{1,t-p} \\ y_{2,t-p} \\ \vdots \\ y_{n,t-p} \end{bmatrix} + \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \vdots \\ \epsilon_{nt} \end{bmatrix}$$

Note that each regression has the same explanatory variables.

$$\begin{aligned} (\mathbf{I}_n - \mathbf{\Phi}_1 L - \mathbf{\Phi}_2 L^2 - \cdots - \mathbf{\Phi}_p L^p)\mathbf{y}_t &= \mathbf{c} + \boldsymbol{\epsilon}_t, \\ \Leftrightarrow \quad \mathbf{A}(L)\mathbf{y}_t &= \mathbf{c} + \boldsymbol{\epsilon}_t, \\ \Leftrightarrow \quad \mathbf{y}_t &= \mathbf{A}(L)^{-1}(\mathbf{c} + \boldsymbol{\epsilon}_t) \\ &= \frac{(-1)^{i+j}|\mathbf{A}(L)_{ji}|}{|\mathbf{A}(L)|}(\mathbf{c} + \boldsymbol{\epsilon}_t). \end{aligned}$$

[1] VAR is a system in which each variable is regressed on a constant and p of its own lags as well as on p lags of each of the other variables in the VAR.

$$\begin{aligned}\mathbb{E}\mathbf{y}_t \equiv \boldsymbol{\mu} &= \mathbf{c} + \boldsymbol{\Phi}_1\boldsymbol{\mu} + \boldsymbol{\Phi}_2\boldsymbol{\mu} + \cdots + \boldsymbol{\Phi}_p\boldsymbol{\mu}\\ \boldsymbol{\mu} &= (\mathbf{I}_n - \boldsymbol{\Phi}_1 - \cdots - \boldsymbol{\Phi}_p)^{-1}\mathbf{c}\\ \mathbf{y}_t - \boldsymbol{\mu} &= \boldsymbol{\Phi}_1(\mathbf{y}_{t-1} - \boldsymbol{\mu}) + \boldsymbol{\Phi}_2(\mathbf{y}_{t-2} - \boldsymbol{\mu}) + \cdots + \boldsymbol{\Phi}_p(\mathbf{y}_{t-p} - \boldsymbol{\mu}) + \boldsymbol{\epsilon}_t.\end{aligned}$$

$$\begin{bmatrix}\mathbf{y}_t-\boldsymbol{\mu}\\ \mathbf{y}_{t-1}-\boldsymbol{\mu}\\ \mathbf{y}_{t-2}-\boldsymbol{\mu}\\ \vdots\\ \mathbf{y}_{t-(p-1)}-\boldsymbol{\mu}\end{bmatrix} = \begin{bmatrix}\boldsymbol{\Phi}_1 & \boldsymbol{\Phi}_2 & \cdots & \boldsymbol{\Phi}_{p-1} & \boldsymbol{\Phi}_p\\ \mathbf{I}_n & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_n & \cdots & \mathbf{0} & \mathbf{0}\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I}_n & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{y}_{t-1}-\boldsymbol{\mu}\\ \mathbf{y}_{t-2}-\boldsymbol{\mu}\\ \mathbf{y}_{t-3}-\boldsymbol{\mu}\\ \vdots\\ \mathbf{y}_{t-p}-\boldsymbol{\mu}\end{bmatrix} + \begin{bmatrix}\boldsymbol{\epsilon}_t\\ \mathbf{0}\\ \mathbf{0}\\ \vdots\\ \mathbf{0}\end{bmatrix} \xLeftrightarrow{\text{VAR(1)}} \mathbf{Y}_t = \mathbf{F}\mathbf{Y}_{t-1} + \underbrace{\boldsymbol{\nu}_t}_{(np\times 1)} \begin{cases}\mathbb{E}\boldsymbol{\nu}_t = \mathbf{0}\\ \mathbb{E}(\boldsymbol{\nu}_t\boldsymbol{\nu}_\tau') \begin{cases}\boldsymbol{\Sigma}^2 \equiv \begin{bmatrix}\boldsymbol{\sigma}^2 & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0}\\ \vdots & \vdots & \cdots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0}\end{bmatrix} & \text{for } t=\tau\\ \mathbf{0} & \text{otherwise}\end{cases}\end{cases}$$

$$\begin{aligned}\mathbf{Y}_t &= \mathbf{F}\mathbf{Y}_{t-1} + \boldsymbol{\nu}_t,\\ \Rightarrow \mathbf{Y}_t &= \mathbf{F}^t\mathbf{Y}_0 + \sum_{i=0}^{t-1}\mathbf{F}^i\boldsymbol{\nu}_{t-i},\\ \xRightarrow{\text{or}} \mathbf{Y}_{t+h} &= \mathbf{F}^{h+1}\mathbf{Y}_{t-1} + \sum_{i=0}^{h}\mathbf{F}^{h-i}\boldsymbol{\nu}_{t+i},\end{aligned}$$

$$\xRightarrow{\text{The first n rows of the VAR(1)}} \mathbf{y}_{t+h} - \boldsymbol{\mu} = \underbrace{\mathbf{F}_{11}^{h+1}(\mathbf{y}_{t-1} - \boldsymbol{\mu}) + \cdots + \mathbf{F}_{1p}^{h+1}(\mathbf{y}_{t-p} - \boldsymbol{\mu})}_{\text{initial conditions}} + \mathbf{F}_{11}^{h}\boldsymbol{\epsilon}_t + \cdots + \mathbf{F}_{11}\boldsymbol{\epsilon}_{t+h-1} + \boldsymbol{\epsilon}_{t+h},$$

$$\xRightarrow{h\to\infty} \mathbf{y}_{t+h} = \boldsymbol{\mu} + \sum_{i=0}^{\infty}\mathbf{F}^i\boldsymbol{\epsilon}_{t-i} \equiv \boldsymbol{\mu} + \sum_{i=0}^{\infty}\boldsymbol{\Psi}_i\boldsymbol{\epsilon}_{t-i} \equiv \boldsymbol{\mu} + \boldsymbol{\Psi}(L)\boldsymbol{\epsilon}_t \quad \leftarrow \quad \text{Vector MA}(\infty).$$

Note that $\mathbf{F}_{11}$ indicates $(1\times n, 1\times n)$, $\mathbf{F}_{12}$ indicates $(1\times n, (n+1)\times 2n)$, $\mathbf{F}_{1p}$ indicates $(1\times n, n(p-1)\times np)$.

2 Stationarity

2.1 Strong stationarity

If the joint probability distribution (various moments including the first- and second- moment etc.) function of $\mathbf{y}_{t-h}, \ldots, \mathbf{y}_t, \ldots, \mathbf{y}_{t+h}$ is independent of t for all h, then the vector process $\{\mathbf{y}_t\}$ is strongly/strictly stationary. SS is useful, e.g., a nonlinear function of a SS vector is SS.

2.2 Weak stationarity

AR lag polynomials are invertible & MA lag polynomials are square summable.

(1) If and only if the impluse-response function $(\sum_{h=0}^{\infty}\beta^h\frac{\partial\mathbf{y}_{t+h}}{\partial\epsilon_t} = \sum_{h=0}^{\infty}\beta^h\mathbf{F}_{11}^h)$ eventually decays exponentially.

$\Updownarrow$

(2) If the eigenvalues of $\mathbf{F}$ in VAR(1) all lie inside the unit circle ($|\mathbf{F} - \boldsymbol{\lambda}\mathbf{I}| = \mathbf{0} \;\Leftrightarrow\; \mathbf{I}_n\boldsymbol{\lambda}^p - \boldsymbol{\Phi}_1\boldsymbol{\lambda}^{p-1} - \boldsymbol{\Phi}_2\boldsymbol{\lambda}^{p-2} - \cdots - \boldsymbol{\Phi}_p = 0$) or if all roots of z satisfying $\mathbf{I}_n - \boldsymbol{\Phi}_1 z - \boldsymbol{\Phi}_2 z^2 - \cdots - \boldsymbol{\Phi}_p z^p = 0$ lie outside the unit circle(i.e., the lag polynomial $\mathbf{A}(L) = \mathbf{I}_n - \boldsymbol{\Phi}_1 L - \boldsymbol{\Phi}_2 L^2 - \cdots - \boldsymbol{\Phi}_p L^p$ is **invertible**), then the original VAR(p) process turns out to be weakly stationary/convariance-stationary.

$\Updownarrow$

(3) Weak stationarity does not require the vector MA polynomial $\mathbf{B}(L) = \mathbf{I}_n + \boldsymbol{\Theta}_1 L + \boldsymbol{\Theta}_2 L^2 + \cdots + \boldsymbol{\Theta}_q L^q$ to be invertible.

$\Updownarrow$

(4) If neither the mean nor the variance depend on time t (i.e., they are finite) and the autocovariances $\mathbb{E}(\mathbf{y}_t\mathbf{y}_{t-h})$ depend only on h but not t,, then the stochastic VAR process is said to be covariance-stationary.

$$
\begin{aligned}
\mathbb{E}\mathbf{y}_t &= \mathbb{E}\mathbf{y}_{t-h} = \boldsymbol{\mu}, \\
\text{var}(\mathbf{y}_t) &\equiv \mathbb{E}(\mathbf{y}_t-\boldsymbol{\mu})(\mathbf{y}_t-\boldsymbol{\mu})' = \mathbb{E}(\mathbf{y}_{t-h}-\boldsymbol{\mu})(\mathbf{y}_{t-h}-\boldsymbol{\mu})' \equiv \text{var}(\mathbf{y}_{t-h}) = \boldsymbol{\sigma^2} \xleftarrow{\boldsymbol{\mu}=\mathbf{c}} \mathbf{y}_t = \mathbf{c}+\boldsymbol{\epsilon}_t, \\
\text{cov}(\mathbf{y}_t,\mathbf{y}_{t-h}) &\equiv \mathbb{E}[(\mathbf{y}_t-\boldsymbol{\mu})(\mathbf{y}_{t-h}-\boldsymbol{\mu})'] = \mathbb{E}[(\mathbf{y}_{t-j}-\boldsymbol{\mu})(\mathbf{y}_{t-j-h}-\boldsymbol{\mu})'] \equiv \text{cov}(\mathbf{y}_{t-j},\mathbf{y}_{t-j-h}) = \boldsymbol{\gamma}_h \\
\text{autocorrelation}(\mathbf{y}_t,\mathbf{y}_{t-h}) &\equiv \boldsymbol{\rho}_h \equiv \frac{\boldsymbol{\gamma}_h}{\boldsymbol{\gamma}_0} = \frac{\mathbb{E}(\mathbf{y}_t-\boldsymbol{\mu})(\mathbf{y}_{t-h}-\boldsymbol{\mu})'}{\mathbb{E}(\mathbf{y}_t-\boldsymbol{\mu})(\mathbf{y}_{t-0}-\boldsymbol{\mu})'} \stackrel{?}{=} \frac{\text{cov}(\mathbf{y}_t,\mathbf{y}_{t-h})}{\text{var}(\mathbf{y}_t)}.
\end{aligned}
$$

2.3 Weak stationarity Restrictions

Multivariate auto- and cross- correlations.

1. Vector MA(q)

$$
\mathbf{y}_t = \mathbf{c} + \underbrace{\boldsymbol{\epsilon}_t}_{(n\times 1)} + \boldsymbol{\Theta}_1\boldsymbol{\epsilon}_{t-1} + \boldsymbol{\Theta}_2\boldsymbol{\epsilon}_{t-2} + \cdots \underbrace{\boldsymbol{\Theta}_i}_{n\times n}\boldsymbol{\epsilon}_{t-i} + \cdots + \boldsymbol{\Theta}_q\boldsymbol{\epsilon}_{t-q},
$$

$$
\begin{aligned}
\boldsymbol{\mu} \equiv \mathbb{E}\mathbf{y}_t &= \mathbf{c}, \\
\boldsymbol{\gamma}_0 \equiv \text{var}(\mathbf{y}_t) &= \mathbb{E}[(\mathbf{y}_t-\mathbf{c})(\mathbf{y}_t-\mathbf{c})'] \\
&= \mathbb{E}[(\boldsymbol{\epsilon}_t+\boldsymbol{\Theta}_1\boldsymbol{\epsilon}_{t-1}+\cdots+\boldsymbol{\Theta}_q\boldsymbol{\epsilon}_{t-q})(\boldsymbol{\epsilon}_t+\boldsymbol{\Theta}_1\boldsymbol{\epsilon}_{t-1}+\cdots+\boldsymbol{\Theta}_q\boldsymbol{\epsilon}_{t-q})'] \\
&= \mathbb{E}(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t') + \boldsymbol{\Theta}_1\mathbb{E}(\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1}')\boldsymbol{\Theta}_1' + \cdots + \boldsymbol{\Theta}_q\mathbb{E}(\boldsymbol{\epsilon}_{t-q}\boldsymbol{\epsilon}_{t-q}')\boldsymbol{\Theta}_q' \\
&= \boldsymbol{\sigma^2} + \boldsymbol{\Theta}_1\boldsymbol{\sigma^2}\boldsymbol{\Theta}_1' + \cdots + \boldsymbol{\Theta}_q\boldsymbol{\sigma^2}\boldsymbol{\Theta}_q', \\
\boldsymbol{\gamma}_1 \equiv \text{cov}(\mathbf{y}_t,\mathbf{y}_{t-1}) &= \mathbb{E}[(\mathbf{y}_t-\mathbf{c})(\mathbf{y}_{t-1}-\mathbf{c})'] \\
&= \mathbb{E}[(\boldsymbol{\epsilon}_t+\boldsymbol{\Theta}_1\boldsymbol{\epsilon}_{t-1}+\cdots+\boldsymbol{\Theta}_q\boldsymbol{\epsilon}_{t-q})(\boldsymbol{\epsilon}_{t-1}+\boldsymbol{\Theta}_1\boldsymbol{\epsilon}_{t-2}+\cdots+\boldsymbol{\Theta}_q\boldsymbol{\epsilon}_{t-q-1})'] \\
&= \boldsymbol{\Theta}_1\mathbb{E}(\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1}') + \boldsymbol{\Theta}_2\mathbb{E}(\boldsymbol{\epsilon}_{t-2}\boldsymbol{\epsilon}_{t-2}')\boldsymbol{\Theta}_1' + \cdots + \boldsymbol{\Theta}_q\mathbb{E}(\boldsymbol{\epsilon}_{t-q}\boldsymbol{\epsilon}_{t-q}')\boldsymbol{\Theta}_{q-1}' \\
&= \boldsymbol{\Theta}_1\boldsymbol{\sigma^2} + \boldsymbol{\Theta}_2\boldsymbol{\sigma^2}\boldsymbol{\Theta}_1' + \cdots + \boldsymbol{\Theta}_q\boldsymbol{\sigma^2}\boldsymbol{\Theta}_{q-1}', \\
\boldsymbol{\gamma}_2 \equiv \text{cov}(\mathbf{y}_t,\mathbf{y}_{t-2}) &= \mathbb{E}[(\mathbf{y}_t-\mathbf{c})(\mathbf{y}_{t-2}-\mathbf{c})'] \\
&= \mathbb{E}[(\boldsymbol{\epsilon}_t+\boldsymbol{\Theta}_1\boldsymbol{\epsilon}_{t-1}+\cdots+\boldsymbol{\Theta}_q\boldsymbol{\epsilon}_{t-q})(\boldsymbol{\epsilon}_{t-2}+\boldsymbol{\Theta}_1\boldsymbol{\epsilon}_{t-3}+\cdots+\boldsymbol{\Theta}_q\boldsymbol{\epsilon}_{t-q-2})'] \\
&= \boldsymbol{\Theta}_2\mathbb{E}(\boldsymbol{\epsilon}_{t-2}\boldsymbol{\epsilon}_{t-2}') + \boldsymbol{\Theta}_3\mathbb{E}(\boldsymbol{\epsilon}_{t-3}\boldsymbol{\epsilon}_{t-3}')\boldsymbol{\Theta}_1' + \cdots + \boldsymbol{\Theta}_q\mathbb{E}(\boldsymbol{\epsilon}_{t-q}\boldsymbol{\epsilon}_{t-q}')\boldsymbol{\Theta}_{q-2}' \\
&= \boldsymbol{\Theta}_2\boldsymbol{\sigma^2} + \boldsymbol{\Theta}_3\boldsymbol{\sigma^2}\boldsymbol{\Theta}_1' \cdots + \boldsymbol{\Theta}_q\boldsymbol{\sigma^2}\boldsymbol{\Theta}_{q-2}', \\
&\vdots \\
\boldsymbol{\gamma}_h \equiv \text{cov}(\mathbf{y}_t,\mathbf{y}_{t-h}) &= \mathbb{E}[(\mathbf{y}_t-\mathbf{c})(\mathbf{y}_{t-h}-\mathbf{c})'] \\
&= \mathbb{E}[(\boldsymbol{\epsilon}_t+\boldsymbol{\Theta}_1\boldsymbol{\epsilon}_{t-1}+\cdots+\boldsymbol{\Theta}_q\boldsymbol{\epsilon}_{t-q})(\boldsymbol{\epsilon}_{t-h}+\boldsymbol{\Theta}_1\boldsymbol{\epsilon}_{t-h-1}+\cdots+\boldsymbol{\Theta}_q\boldsymbol{\epsilon}_{t-q-h})'] \\
&= \boldsymbol{\Theta}_h\mathbb{E}(\boldsymbol{\epsilon}_{t-h}\boldsymbol{\epsilon}_{t-h}') + \boldsymbol{\Theta}_{h+1}\mathbb{E}(\boldsymbol{\epsilon}_{t-h-1}\boldsymbol{\epsilon}_{t-h-1}')\boldsymbol{\Theta}_1' + \cdots + \boldsymbol{\Theta}_q\mathbb{E}(\boldsymbol{\epsilon}_{t-q}\boldsymbol{\epsilon}_{t-q}')\boldsymbol{\Theta}_{q-h}' \\
&= \boldsymbol{\Theta}_h\boldsymbol{\sigma^2} + \boldsymbol{\Theta}_{h+1}\boldsymbol{\sigma^2}\boldsymbol{\Theta}_1' + \cdots + \boldsymbol{\Theta}_q\boldsymbol{\sigma^2}\boldsymbol{\Theta}_{q-h}' \quad \text{for } h=0,1,2,\ldots,q. \quad \boldsymbol{\Theta}_0 \equiv \mathbf{I}_n. \\
\xRightarrow{\boldsymbol{\gamma}_h=\boldsymbol{\gamma}_{-h}'} &= \quad \boldsymbol{\sigma^2}\boldsymbol{\Theta}_{-h}' + \boldsymbol{\Theta}_1\boldsymbol{\sigma^2}\boldsymbol{\Theta}_{-h+1}' + \cdots + \boldsymbol{\Theta}_{q+h}\boldsymbol{\sigma^2}\boldsymbol{\Theta}_q' \quad \text{for } h=0,-1,-2,\ldots,-q. \\
&= \mathbf{0} \quad \text{for } |h|>q.
\end{aligned}
$$

2. Vector MA(∞)

I refer the reader to Hamilton (1994, ch.10, p.262)

3. VAR(1) with 2 variables y_t and π_t

$$\mathbf{z}_t = \mathbf{F}\mathbf{z}_{t-1} + \epsilon_t, \quad \text{where } \mathbf{z}_t = \begin{bmatrix} y_t \\ \pi_t \end{bmatrix}, \quad \boldsymbol{\epsilon}_t = \begin{bmatrix} \epsilon_{yt} \\ \epsilon_{\pi t} \end{bmatrix} \sim \text{i.i.d. } \mathcal{N}(\mathbf{0}, \sigma^{\mathbf{2}}) \begin{cases} \mathbb{E}\boldsymbol{\epsilon}_t = 0, \\ \mathbb{E}(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t') = \boldsymbol{\sigma}^{\mathbf{2}} = \begin{bmatrix} \sigma^2_{\epsilon_y} & \sigma_{\epsilon_y \epsilon_\pi} \\ \sigma_{\epsilon_y \epsilon_\pi} & \sigma^2_{\epsilon_\pi} \end{bmatrix} \overset{\perp}{=\!=} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathrm{I}, \\ \mathbb{E}(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_{t-h}') = 0. \end{cases}$$

$$\left.\begin{aligned} y_t &= \phi_{yy} y_{t-1} + \phi_{y\pi} \pi_{t-1} + \epsilon_{yt} \\ \pi_t &= \phi_{\pi y} y_{t-1} + \phi_{\pi\pi} \pi_{t-1} + \epsilon_{\pi t} \end{aligned}\right\} \Rightarrow \begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} \phi_{yy} & \phi_{y\pi} \\ \phi_{\pi y} & \phi_{\pi\pi} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ \pi_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{yt} \\ \epsilon_{\pi t} \end{bmatrix}$$

$$(\mathbf{I} - \mathbf{F}L)\mathbf{z}_t = \epsilon_t \quad \Leftrightarrow \quad \mathbf{z}_t = (\mathbf{I} - \mathbf{F}L)^{-1}\epsilon_t = \sum_{i=0}^{\infty} \mathbf{F}^i \epsilon_{t-i}.$$

$$\mathbb{E}\mathbf{z}_t \equiv \boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

$$\text{var}(\mathbf{z}_t, \mathbf{z}_t) = \mathbb{E}(\mathbf{z}_t \mathbf{z}_t') \equiv \boldsymbol{\gamma}_0 = \mathbb{E}\left\{ \begin{bmatrix} y_t \\ \pi_t \end{bmatrix} \times \begin{bmatrix} y_t & \pi_t \end{bmatrix} \right\} = \mathbb{E}\left\{ \begin{bmatrix} y_t^2 & y_t\pi_t \\ \pi_t y_t & \pi_t^2 \end{bmatrix} \right\} = \begin{bmatrix} \text{var}(y_t) & \text{cov}(y_t\pi_t) \\ \text{cov}(y_t\pi_t) & \text{var}(\pi_t) \end{bmatrix} \leftarrow \text{var/cov}$$

$$\text{cov}(\mathbf{z}_t, \mathbf{z}_{t-h}) = \mathbb{E}(\mathbf{z}_t \mathbf{z}_{t-h}') \equiv \boldsymbol{\gamma}_h = \begin{bmatrix} \mathbb{E}(y_t y_{t-h}) & \mathbb{E}(y_t \pi_{t-h}) \\ \mathbb{E}(\pi_t y_{t-h}) & \mathbb{E}(\pi_t \pi_{t-h}) \end{bmatrix} \leftarrow \text{covariances and cross-covariances}$$

$$\text{corr}(\mathbf{z}_t, \mathbf{z}_{t-h}) \equiv \boldsymbol{\rho}_h = \begin{bmatrix} \frac{\mathbb{E}(y_t y_{t-h})}{\sigma_y^2} & \frac{\mathbb{E}(y_t \pi_{t-h})}{\sigma_y \sigma_\pi} \\ \frac{\mathbb{E}(\pi_t y_{t-h})}{\sigma_y \sigma_\pi} & \frac{\mathbb{E}(\pi_t \pi_{t-h})}{\sigma_\pi^2} \end{bmatrix}$$

Note that $\boldsymbol{\gamma}_h \neq \boldsymbol{\gamma}_{-h}$ but $\boldsymbol{\gamma}_h = \boldsymbol{\gamma}_{-h}' \Leftrightarrow \mathbb{E}(\mathbf{z}_t \mathbf{z}_{t-h}') = \gamma_h = \gamma_{-h}' = [\mathbb{E}(\mathbf{z}_t \mathbf{z}_{t+h}')]'$ or $\boldsymbol{\gamma}_h' = \boldsymbol{\gamma}_{-h}$.

For example, the (1, 2) element of γ_j gives the covariance between y_{1t} and $y_{2,t-j}$, and the the (1, 2) element of γ_{-j} gives the covariance between y_{1t} and $y_{2,t+j}$. Obviousely, they are different.

To derive $\boldsymbol{\gamma}_h' = \boldsymbol{\gamma}_{-h}$, notice that

$$\begin{aligned} \boldsymbol{\gamma}_h &= \mathbb{E}[(\mathbf{y}_{t+h} - \boldsymbol{\mu})(\mathbf{y}_{(t+h)-h} - \boldsymbol{\mu})'], \\ &= \mathbb{E}[(\mathbf{y}_{t+h} - \boldsymbol{\mu})(\mathbf{y}_t - \boldsymbol{\mu})'], \\ \boldsymbol{\gamma}_h' &= \mathbb{E}[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t+h} - \boldsymbol{\mu})'] = \boldsymbol{\gamma}_{-h}. \end{aligned}$$

4. VAR(p)→VAR(1) with n variables

$$\mathbf{Y}_t = \mathbf{F}\mathbf{Y}_{t-1} + \boldsymbol{\nu}_t \overset{\text{iteration}}{=} \mathbf{F}^t \mathbf{Y}_0 + \sum_{i=0}^{t-1} \mathbf{F}^i \boldsymbol{\nu}_{t-i} = \sum_{i=0}^{\infty} \mathbf{F}^i \boldsymbol{\nu}_{t-i},$$

$$\boldsymbol{\mu} \equiv \mathbb{E}\mathbf{Y}_t = \mathbf{0},$$

$$\begin{aligned} \boldsymbol{\Gamma}_0 \equiv \mathbb{E}(\mathbf{Y}_t \mathbf{Y}_t') &= \mathbb{E}\left\{ \begin{bmatrix} \mathbf{y}_t - \boldsymbol{\mu} \\ \mathbf{y}_{t-1} - \boldsymbol{\mu} \\ \vdots \\ \mathbf{y}_{t-(p-1)} - \boldsymbol{\mu} \end{bmatrix} \times \begin{bmatrix} (\mathbf{y}_t - \boldsymbol{\mu})' & (\mathbf{y}_{t-1} - \boldsymbol{\mu})' & \cdots & (\mathbf{y}_{t-(p-1)} - \boldsymbol{\mu})' \end{bmatrix} \right\} \\ &= \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{p-1} \\ \gamma_1' & \gamma_0 & \cdots & \gamma_{p-2} \\ \vdots & \vdots & \cdots & \vdots \\ \gamma_{p-1}' & \gamma_{p-2}' & \cdots & \gamma_0 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} \boldsymbol{\Gamma}_0 \equiv \mathbb{E}(\mathbf{Y}_t \mathbf{Y}_t') &= \mathbb{E}[(\mathbf{F}\mathbf{Y}_{t-1} + \boldsymbol{\nu}_t)(\mathbf{F}\mathbf{Y}_{t-1} + \boldsymbol{\nu}_t)'] \\ &= \mathbf{F}\mathbb{E}(\mathbf{Y}_{t-1}\mathbf{Y}_{t-1}')\mathbf{F}' + \mathbb{E}(\boldsymbol{\nu}_t \boldsymbol{\nu}_t') \\ &= \mathbf{F}\boldsymbol{\Gamma}_0\mathbf{F}' + \boldsymbol{\Sigma}^2 \end{aligned}$$

Sloving the above equation by the vec operator, e.g., $\text{vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A}) \cdot \text{vec}(\mathbf{B})$:

$$\begin{aligned}
\mathbf{\Gamma}_0 &= \underbrace{\mathbf{F}}_{(np\times np)} \mathbf{\Gamma}_0 \mathbf{F}' + \underbrace{\mathbf{\Sigma}^2}_{(np\times np)}, \\
\Rightarrow \quad \text{vec}(\mathbf{\Gamma}_0) &= \text{vec}(\mathbf{F}\mathbf{\Gamma}_0\mathbf{F}') + \text{vec}(\mathbf{\Sigma}^2), \\
\Rightarrow \quad \text{vec}(\mathbf{\Gamma}_0) &= (\underbrace{\mathbf{F}\otimes\mathbf{F}}_{(np)^2\times(np)^2})\text{vec}\mathbf{\Gamma}_0 + \text{vec}(\mathbf{\Sigma}^2), \\
\xRightarrow{\text{nonsingular}} \text{vec}(\mathbf{\Gamma}_0) &= (\mathbf{I} - \mathbf{F}\otimes\mathbf{F})^{-1}\text{vec}(\mathbf{\Sigma}^2), \\
\Rightarrow \quad \text{vec}\begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{p-1} \\ \gamma_1' & \gamma_0 & \cdots & \gamma_{p-2} \\ \vdots & \vdots & \cdots & \vdots \\ \gamma_{p-1}' & \gamma_{p-2}' & \cdots & \gamma_0 \end{bmatrix} &= (\mathbf{I} - \mathbf{F}\otimes\mathbf{F})^{-1}\text{vec}\begin{bmatrix} \boldsymbol{\sigma^2} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix},
\end{aligned}$$

Note that the matrix $(\mathbf{I} - \mathbf{F}\otimes\mathbf{F})$ is nonsingular as long as unity is not an gienvalue of $\mathbf{F}\otimes\mathbf{F}$ whose eigenvaluses are all of the from $\lambda_i\lambda_j$. Since λ_i and λ_j are eigenvalues of $\mathbf{F}$, and all of them are inside the unit circle, meaning that $\lambda_i\lambda_j$ are also inside the unit circle, deriving that it is indeed nonsingular.

$$\begin{aligned}
\mathbf{\Gamma}_1 \equiv \mathbb{E}(\mathbf{Y}_t\mathbf{Y}_{t-1}') &= \mathbb{E}[(\mathbf{F}\mathbf{Y}_{t-1} + \boldsymbol{\nu}_t)\mathbf{Y}_{t-1}'] \\
&= \mathbf{F}\mathbb{E}(\mathbf{Y}_{t-1}\mathbf{Y}_{t-1}') + \mathbb{E}(\boldsymbol{\nu}_t\mathbf{Y}_{t-1}') \\
&= \mathbf{F}\mathbf{\Gamma}_0, \\
\mathbf{\Gamma}_2 \equiv \mathbb{E}(\mathbf{Y}_t\mathbf{Y}_{t-2}') &= \mathbb{E}[(\mathbf{F}\mathbf{Y}_{t-1} + \boldsymbol{\nu}_t)\mathbf{Y}_{t-2}'] \\
&= \mathbf{F}\mathbb{E}(\mathbf{Y}_{t-1}\mathbf{Y}_{t-2}') + \mathbb{E}(\boldsymbol{\nu}_t\mathbf{Y}_{t-2}') \\
&= \mathbf{F}\mathbf{\Gamma}_1, \\
&\vdots \\
\mathbf{\Gamma}_h \equiv \mathbb{E}(\mathbf{Y}_t\mathbf{Y}_{t-h}') &= \mathbb{E}[(\mathbf{F}\mathbf{Y}_{t-1} + \boldsymbol{\nu}_t)\mathbf{Y}_{t-h}'] \\
&= \mathbf{F}\mathbb{E}(\mathbf{Y}_{t-1}\mathbf{Y}_{t-h}') + \mathbb{E}(\boldsymbol{\nu}_t\mathbf{Y}_{t-h}') \\
&= \mathbf{F}\mathbf{\Gamma}_{h-1} \quad \text{for } h \geq 1; \\
&= \mathbf{F}^h\mathbf{\Gamma}_0 \quad \leftarrow \text{ iteration.}
\end{aligned}$$

The autocovariance $\boldsymbol{\gamma}_h$ $(h = p, p+1, p+2, \ldots)$ of the original vector $\mathbf{y}_t$ in VAR(p) process is given by the first n rows and n columns of $\mathbf{F}$:

$$\boldsymbol{\gamma}_h = \mathbf{\Phi}_1\boldsymbol{\gamma}_{h-1} + \mathbf{\Phi}_2\boldsymbol{\gamma}_{h-2} + \cdots + \mathbf{\Phi}_p\boldsymbol{\gamma}_{h-p}.$$

3 Vector Autoregressions

Their popularity for analyzing the dynamics of economic systems is due to Sims's (1980) influential work.See also CEE (1999).

Recall the following white noise process:

$$\begin{cases} \mathbb{E}u_t = 0 \\ \mathbb{E}u_t u_t' = \Omega \\ \mathbb{E}u_i u_j' = 0, \text{ if } i \neq j \end{cases}$$

$$Z_t = b + B_1 Z_{t-1} + B_2 Z_{t-2} + \cdots + B_q Z_{t-q} + u_t, \quad (u_t \to \text{white noise process})$$

$$\begin{bmatrix} Z_t \\ Z_{t-1} \\ \vdots \\ Z_{t-q+1} \end{bmatrix} = \begin{bmatrix} b \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} B_1 & B_2 & \cdots & B_{q-1} & B_q \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} Z_{t-1} \\ Z_{t-2} \\ \vdots \\ Z_{t-q} \end{bmatrix} + \begin{bmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\begin{aligned}
\vec{Z}_t &= \vec{b} + B\vec{Z}_{t-1} + \vec{u}_t \quad \leftarrow \text{canonical form of the VAR} \\
\vec{Z}_{t-1} &= \vec{b} + B\vec{Z}_{t-2} + \vec{u}_{t-1} \\
&\vdots \\
\vec{Z}_{t-s} &= \vec{b} + B\vec{Z}_{t-s-1} + \vec{u}_{t-s} \quad \text{for } s = 1, 2, \cdots, k. \\
\Rightarrow \vec{Z}_t &= \vec{b} + B(\vec{b} + B\vec{Z}_{t-2} + \vec{u}_{t-1}) + \vec{u}_t \\
&= \vec{b} + B\vec{b} + B^2\vec{Z}_{t-2} + B\vec{u}_{t-1} + \vec{u}_t \\
&= \vec{b} + B\vec{b} + B^2(\vec{b} + B\vec{Z}_{t-3} + \vec{u}_{t-2}) + B\vec{u}_{t-1} + \vec{u}_t \\
&= \vec{b} + B\vec{b} + B^2\vec{b} + B^3\vec{Z}_{t-3} + B^2\vec{u}_{t-2} + B\vec{u}_{t-1} + \vec{u}_t \\
&\vdots \\
&= \vec{b} + B\vec{b} + B^2\vec{b} + \cdots + B^k\vec{b} + B^{k+1}\vec{Z}_{t-k-1} + B^k\vec{u}_{t-k} + \cdots + B\vec{u}_{t-1} + \vec{u}_t. \\
\mathbb{E}\vec{Z}_t &= \vec{b} + B\vec{b} + B^2\vec{b} + \cdots\cdots\cdots + B^k\vec{b} + B^{k+1}\vec{Z}_{t-k-1} \quad \leftarrow \text{the mean function} \\
B\mathbb{E}\vec{Z}_t &= B\vec{b} + B^2\vec{b} + B^3\vec{b} + \cdots + B^k\vec{b} + B^{k+1}\vec{b} + B^{k+2}\vec{Z}_{t-k-1} \\
\Rightarrow (I - B)\mathbb{E}\vec{Z}_t &= \vec{b} - B^{k+1}\vec{b} + (B^{k+1} - B^{k+2})\vec{Z}_{t-k-1}. \\
\xRightarrow{k\to\infty, B^{k+1}=0} \mathbb{E}\vec{Z}_t &= (I - B)^{-1}b \equiv \mu, \quad \text{assume the inverse exist.} \\
\vec{Z}_t - \mathbb{E}\vec{Z}_t &= B^k\vec{u}_{t-k} + \cdots + B\vec{u}_{t-1} + \vec{u}_t \\
\Rightarrow \vec{Z}_t &= \mathbb{E}\vec{Z}_t + B^k\vec{u}_{t-k} + \cdots + B\vec{u}_{t-1} + \vec{u}_t \\
&= \mu + B^k\vec{u}_{t-k} + \cdots + B\vec{u}_{t-1} + \vec{u}_t. \\
\Leftrightarrow \vec{Z}_t &= \mu + \vec{u}_t + B\vec{u}_{t-1} + \cdots + B^k\vec{u}_{t-k} \\
\xRightarrow{k\to\infty} \vec{Z}_t &= \mu + \vec{u}_t + B\vec{u}_{t-1} + B^2\vec{u}_{t-2} + \cdots \\
\vec{Z}_t &= \mu + \vec{u}_t + \Psi_1\vec{u}_{t-1} + \Psi_2\vec{u}_{t-2} + \cdots \\
\vec{Z}_{t+k} &= \mu + \vec{u}_{t+k} + \Psi_1\vec{u}_{t+k-1} + \Psi_2\vec{u}_{t+k-2} + \cdots + \Psi_k\vec{u}_{t+k-k}, \\
\Rightarrow \quad \frac{\partial\vec{Z}_{t+k}}{\partial\vec{u}_t'} &= \mathbf{\Psi}_k \supset \psi_{i,j}(k) = \frac{\partial\vec{Z}_{i,t+k}}{\partial\vec{u}_{j,t}} \quad \leftarrow \text{the impulse response function } \Leftrightarrow \Delta\vec{Z}_{t+k} = \mathbf{\Psi}_k\Delta\vec{u}.
\end{aligned}$$

It was stated above that the condition for a **stable** VAR is that all eigenvalues of the coefficient matrix B lie inside the unit circle. The eigenvalue (λ) of matrix B is defined as $\det(B - \lambda I) = 0$. If the eigenvectors are linearly independent (i.e., all eigenvalues are distint), the spectral decomposition can be applied:

$$B = T\Lambda T^{-1}, \quad \text{where } \Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \qquad \text{following } \Gamma = \begin{bmatrix} 1 & ? & \cdots & ? \\ 0 & 1 & \cdots & ? \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$B^2 = BB = T\Lambda T^{-1} \times T\Lambda T^{-1} = T\Lambda^2 T^{-1} \xrightarrow{\text{generalised to}} B^s = T\Lambda^s T^{-1}.$$

If the inverse of T does not exist, then we can resort to an alternative way to derive the decomposition which is the Schur decomposition.

$$T = \begin{bmatrix} 2 & 3+4i \\ 1-2i & 5 \end{bmatrix} \rightarrow T' = \begin{bmatrix} 2 & 1-2i \\ 3+4i & 5 \end{bmatrix} \xrightarrow{\text{if } T'=T^{-1}} \xRightarrow{\text{the Schur decomposition}} B = T\Gamma T'.$$

$$\begin{aligned}
Z_t &= b + B_1 Z_{t-1} + B_2 Z_{t-2} + \cdots + B_q Z_{t-q} + u_t, \\
&= b + B_1 L(Z_t) + B_2 L^2(Z_t) + \cdots + B_q L^q(Z_t) + u_t, \\
\Rightarrow \ B(L) Z_t &= b + u_t, \quad \text{where } B(L) = I - B_1 L - B_2 L^2 - \cdots - B_q L^q. \\
\Rightarrow \ Z_t &= B(L)^{-1} b + B(L)^{-1} u_t, \\
&= \mu + \Psi(L) u_t \qquad \Psi(L) = B(L)^{-1} \Leftrightarrow B(L)\Psi(L) = I. \\
\Psi(L) &= B(L)^{-1} = \Psi_0 + \Psi_1 L + \Psi_2 L^2 + \Psi_3 L^3 + \cdots \quad \text{refer to Hamilton, 1994, p.35} \\
I &= (I - B_1 L - B_2 L^2 - \cdots - B_q L^q)(\Psi_0 + \Psi_1 L + \Psi_2 L^2 + \cdots) \\
&= \Psi_0 + (\Psi_1 - \Psi_0 B_1) L + (\Psi_2 - \Psi_1 B_1 - \Psi_0 B_2) L^2 + \cdots + (\Psi_i - \sum_{j=1}^{i} \Psi_{i-j} B_j) L^i + \cdots
\end{aligned}$$

$$\Leftrightarrow \left.\begin{aligned}
I &= \Psi_0, \\
0 &= \Psi - \Psi_0 B_1, \\
0 &= \Psi_2 - \Psi_1 B_1 - \Psi_0 B_2, \\
&\ \vdots \\
0 &= \Psi_i - \sum_{j=1}^{i} \Psi_{i-j} B_j, \\
&\ \vdots
\end{aligned}\right\} \xRightarrow{\text{where } B_j = 0 \text{ for } j > p} \begin{cases} \Psi_0 = I, \\ \Psi_i = \sum_{j=1}^{i} \Psi_{i-j} B_j, \quad \text{for } i = 1, 2, \ldots \end{cases}$$

$$\mu = \Psi(1) b = B(1)^{-1} b = (I - B_1 - \cdots - B_q)^{-1} b \quad \leftarrow \text{the mean of } Z_t.$$

The error forecast of the s period ahead forecast is

$$\begin{aligned}
Z_t - \mathbb{E}_0 Z_t &= u_t + \Psi_1 u_{t-1} + \Psi_2 u_{t-2} + \cdots + \Psi_{t-1} u_1, \\
Z_{t+s} - \mathbb{E}_t Z_{t+s} &= u_{t+s} + \Psi_1 u_{t+s-1} + \Psi_2 u_{t+s-2} + \cdots + \Psi_{s-1} u_{t+1}, \\
\operatorname{cov}(Z_{t+s} - \mathbb{E}_t Z_{t+s}) &= \mathbb{E}[(Z_{t+s} - \mathbb{E}_t Z_{t+s})(Z_{t+s} - \mathbb{E}_t Z_{t+s})'] = \Omega + \Psi_1 \Omega \Psi_1' + \cdots + \Psi_{s-1} \Omega \Psi_{s-1}'.
\end{aligned}$$

Update the Teaching schedule

Lec1: 4 Methods to Solve Linear Defference Equations

1.1 Solving DEs with Constant Coefficients and Constant Terms (Chiang, ch.17; Enders, ch.1)
1.2 Solving DEs with Constant Coefficients and Variable Terms (Enders, ch.1; Hamilton, ch.2, ch.1)

Lec2: Covariance-Stationary ARMA Models

2.1 Stationary Restrictions for ARMA(p, q) (Enders, ch.2; Hamilton, ch.3; Cochrane, ch.6)
2.2 The Autocorrelation Function (Enders, ch.2; Cochrane, ch.4)
2.3 ACF+PACF+AIC+SBC→Identification/Specification→Estimation→Diagnostic Check→Forecasting

Lec3: Covariance-Stationary Vector Processes

3.1 VAR(p)→VAR(1) (Cochrane, ch.4.5)
3.2 Stationary Restrictions for Vector Processes (Hamilton, ch.10)

Lec4: Forecasts Based on Conditional Expectation

4.1 Predicting ARMA (Cochrane, ch.5)
4.2 Forecasts from VAR (Cochrane, ch.5)

Lec5: Forecasts Based on Linear Projection

5.1 Linear Projection vs. Conditional Expectation (Hamilton 1994, ch.4)
5.2 Linear Projection vs. OLS Regression (Hamilton 1994, ch.4)
5.3 Wold Decomposition Theorem (Cochrane, ch.6)

Lec6: Calibration and Simulation

6.1 Parameter Calibration
6.2 Impulse Response Simulation

Lec7: Specification and Estimation

7.1 ARMA
7.2 VAR

Lec8: Autocovariance-Generating Functions and Spectral Analysis

8.1 The Autocovariance-Generating Function for ARMA Models (Hamilton, ch.3)
8.2 The Autocovariance-Generating Function for Vector Processes (Hamilton, ch.10, ch.6)
⋮