Aula 5: Análise de séries temporais multivariadas

Marcus L. Nascimento

21 de novembro de 2025

Introdução

Introdução

- Nos modelos do tipo AR, MA, ARMA, ARCH e GARCH, analisamos séries temporais nas quais a variável aleatória Y_t (ou R_t) é unidimensional.

- Nesta etapa do curso, analisaremos séries temporais nas quais a variável aleatória Y_t é um vetor de dimensão k, $\boldsymbol{Y}_t = (Y_{1t}, Y_{2t}, \ldots, Y_{kt})'$.

- Podemos pensar, por exemplo, em uma aplicação na qual:

  - Y_1: crescimento anual real do Produto Interno Bruto (PIB);

  - Y_2: taxa de inflação anual;

  - Y_3: taxa de desemprego.

Estacionariedade Fraca

- Nos casos em que $\boldsymbol{Y}_t$ é uma variável aleatória de dimensão k, uma série temporal $\{\boldsymbol{Y}_t\}_{t=1}^n$ é dita fracamente estacionária se:

  1. $\mu_{\boldsymbol{Y}}(t) = \mathrm{E}(\boldsymbol{Y}_t) = \boldsymbol{\mu}$, $|\boldsymbol{\mu}| < \infty$;
  2. $\Gamma_{\boldsymbol{Y}}(t) = \mathrm{E}[(\boldsymbol{Y}_t - \boldsymbol{\mu})(\boldsymbol{Y}_t - \boldsymbol{\mu})'] = \Gamma_0$, $\Gamma_0 < \infty$;
  3. $\Gamma_{\boldsymbol{Y}}(t, t-l) = \mathrm{E}[(\boldsymbol{Y}_t - \boldsymbol{\mu})(\boldsymbol{Y}_{t-l} - \boldsymbol{\mu})'] = \Gamma_l \; \forall l$ com $\Gamma_l < \infty$,

  onde

  - $\boldsymbol{\mu}$ é o vetor de médias $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_k)'$;
  - Γ_0 é uma matriz de covariância de dimensão $k \times k$, na qual a entrada (i, i) é a variância de Y_{it} e a entrada (i, j) é a covariância entre Y_{it} e Y_{jt};
  - Γ_l é a matriz de covariância cruzada na defasagem l.

Matriz de Correlação

A partir da matriz de covariância, é possível obter a matriz de correlação.

Defasagem 0:

Seja Γ_0 a matriz de covariância na defasagem 0 e D uma matriz diagonal de dimensão $k \times k$ contendo os desvios padrão de Y_{it}, $i = 1, 2, \ldots, k$, a matriz de correlação de $\boldsymbol{Y}_t$ é definida da seguinte forma:

$$\boldsymbol{\rho}_0 = D^{-1}\Gamma_0 D^{-1},$$

na qual a entrada (i, j) de $\boldsymbol{\rho}_0$ é o coeficiente de correlação entre Y_{it} e Y_{jt} no tempo t:

$$\rho_{ij}(0) = \frac{\text{Cov}(Y_{it}, Y_{jt})}{\sqrt{\sigma^2_{Y_i}(t)\sigma^2_{Y_j}(t)}}.$$

Matriz de Correlação

Defasagem l:

Seja Γ_l a matriz de covariância cruzada de $\boldsymbol{Y}_t$ na defasagem l e D uma matriz diagonal de dimensão $k \times k$ contendo os desvios padrão de Y_{it}, $i = 1, 2, \ldots, k$, a matriz de correlação cruzada na defasagem l é definida da seguinte forma:

$$\boldsymbol{\rho}_l = D^{-1} \Gamma_l D^{-1},$$

na qual a entrada (i, j) de $\boldsymbol{\rho}_l$ é o coeficiente de correlação entre Y_{it} e Y_{jt-l}:

$$\rho_{ij}(l) = \frac{\mathrm{Cov}(Y_{it}, Y_{jt-l})}{\sqrt{\sigma^2_{Y_i}(t)\sigma^2_{Y_j}(t)}}.$$

Ruído Branco

- Assim como no caso univariado, processos do tipo ruído branco também são essenciais na construção de modelos de séries temporais multivariados

- Uma série temporal estacionária $\{\boldsymbol{W}_t\}_{t=1}^n$ é denominada ruído branco se satisfaz:

  - $\mathrm{E}(\boldsymbol{W}_t) = \mathbf{0}$ para todo t;

  - Não correlação serial:

    - $\boldsymbol{W}_t$'s independendentes (sentido forte);

    - $\mathrm{E}(\boldsymbol{W}_t \boldsymbol{W}_s') = \mathbf{0}$ para $s \neq t$ (sentido fraco).

Ruído Branco

- Como $\boldsymbol{W}_t$ é multivariado, além da correlação temporal, há a correlação entre as componentes do vetor $\boldsymbol{W}_t = (W_{1t}, W_{2t}, \ldots, W_{kt})$.
  - Em outras palavras, em cada instante t, as componentes de $\boldsymbol{W}_t$ podem estar correlacionadas.
- No caso de um ruído branco, temos:

$$\tau_{ij}(h) = \mathrm{E}(W_{it} W_{jt+h}) = \tau_{ij}\delta_0(h),$$

onde $\delta_0(h) = 1$ quando $h = 0$ e $\delta_0(h) = 0$ caso contrário.

Vetor Autorregressivo (VAR)

Introdução

- A estrutura do modelo Vetor Autorregressivo é uma generalização do modelo Autorregressivo (AR) univariado.
    - Sistema de modelo de regressão que trata todas as variáveis como endógenas;
    - Permite que cada variável dependa de p defasagens dela própria e das demais variáveis no sistema.
- Amplamente aplicado em macroeconomia e finanças.
- Dentre os objetivos do modelo, destacam-se:
    - Previsão com base em observações passadas;
    - Quantificação da resposta a um choque inesperado na própria variável períodos à frente e também em outra variável relacionada: análise impulso-resposta.

Introdução

del Negro (2011) destaca:

"VARs appear to be straightforward multivariate generalizations of univariate autoregressive models. They turn out to be one of the key empirical tools in modern macroeconomics.

Sims (1980) proposed that VARs should replace large-scale macroeconometric models inherited from the 1960s, because the latter imposed incredible restrictions, which were largely inconsistent with the notion that economic agents take the effect of today's choices on tomorrow's utility into account.

VARs have been used for macroeconomic forecasting and policy analysis to investigate the sources of business-cycle fluctuations and to provide a benchmark against which modern dynamic macroeconomic theories can be evaluated."

Forma Estrutural e Reduzida

Considere o seguinte sistema bivariado ($k = 2$):

$$
\begin{aligned}
y_{1t} &= \upsilon_{10} - \upsilon_{12} y_{2t} + \eta_{11} y_{1t-1} + \eta_{12} y_{2t-1} + \tilde{w}_{1t} \\
y_{2t} &= \upsilon_{20} - \upsilon_{21} y_{1t} + \eta_{21} y_{1t-1} + \eta_{22} y_{2t-1} + \tilde{w}_{2t},
\end{aligned}
$$

onde assume-se que y_{1t} e y_{2t} são estacionários e w_{1t} e w_{2t} são termos de erro não correlacionados com desvio padrão τ_{1t} e τ_{2t}, respectivamente.

Em notação matricial, temos:

$$
\begin{bmatrix} 1 & \upsilon_{12} \\ \upsilon_{21} & 1 \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \upsilon_{10} \\ \upsilon_{20} \end{bmatrix} + \begin{bmatrix} \eta_{11} & \eta_{12} \\ \eta_{21} & \eta_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \tilde{w}_{1t} \\ \tilde{w}_{2t} \end{bmatrix}
$$

Forma Estrutural e Reduzida

Tomando

$$U = \begin{bmatrix} 1 & \upsilon_{12} \\ \upsilon_{21} & 1 \end{bmatrix}, \ \boldsymbol{y}_t \ = \ \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix}, \ Q_0 = \begin{bmatrix} \upsilon_{10} \\ \upsilon_{20} \end{bmatrix}, \ Q_1 = \begin{bmatrix} \eta_{11} & \eta_{12} \\ \eta_{21} & \eta_{22} \end{bmatrix},$$

$$\boldsymbol{y}_{t-1} \ = \ \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} \text{ e } \tilde{\boldsymbol{w}}_t = \begin{bmatrix} \tilde{w}_{1t} \\ \tilde{w}_{2t} \end{bmatrix},$$

obtemos

$$U\boldsymbol{y}_t \ = \ Q_0 + Q_1\boldsymbol{y}_{t-1} + \tilde{\boldsymbol{w}}_{t-1}.$$

Forma Estrutural e Reduzida

- Vetor Autorregressivo Estrutural (SVAR):
  - $\tilde{w}_{1t}$ e $\tilde{w}_{2t}$ são denominados erros estruturais;
  - Em geral, $\text{Cov}(y_{1t}, \tilde{w}_{2t}) \neq 0$ e $\text{Cov}(y_{2t}, \tilde{w}_{1t}) \neq 0$;
  - As variáveis são endógenas - Mínimos Quadrados Ordinários (MQO) não é apropriado.

Multiplicando a Equação 1 por U^{-1}, temos:

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{a}_0 + A_1 \boldsymbol{y}_{t-1} + \boldsymbol{w}_t \\ &= \boldsymbol{a}_0 + A_1 B \boldsymbol{y}_t + \boldsymbol{w}_t, \end{aligned}$$

onde $\boldsymbol{a}_0 = U^{-1} Q_0$, $A_1 = U^{-1} Q_1$, B é o operador de defasagem e $\boldsymbol{w}_t = U^{-1} \tilde{\boldsymbol{w}}_t$. Equivalentemente,

$$\begin{aligned} y_{1t} &= a_{10} + a_{11} y_{1t-1} + a_{12} y_{2t-1} + w_{1t} \\ y_{2t} &= a_{20} + a_{21} y_{1t-1} + a_{22} y_{2t-1} + w_{2t}. \end{aligned}$$

Forma Estrutural e Reduzida

- Vetor Autorregressivo em forma reduzida (VAR):
  - Dependente apenas de defasagens das variáveis endógenas;
  - Pode ser estimado via Mínimos Quadrados Ordinários (MQO).
- Sem que restrições sejam impostas, os parâmetros do VAR estrutural não são identificados, isto é, dados parâmetros da forma reduzida, não há solução única para os parâmetros estruturais.
- Restrições de identificação incluem:
  - Restrições de exclusão nos elementos de U (por exemplo, $\upsilon_{12} = 0$);
  - Restrições lineares em U (por exemplo, $\upsilon_{12} + \upsilon_{21} = 1$).

Vetor Autorregressivo de Ordem p - VAR(p)

- O modelo VAR(p) é descrito pela seguinte equação:

$$\boldsymbol{y}_t = \boldsymbol{a}_0 + A_1 \boldsymbol{y}_{t-1} + A_2 \boldsymbol{y}_{t-2} + \ldots + A_p \boldsymbol{y}_{t-p} + \boldsymbol{w}_t. \tag{1}$$

- Utilizando o operador de defasagem, temos:

$$(I_k - A_1 B - A_2 B^2 - \ldots - A_p B^p)\boldsymbol{y}_t = \boldsymbol{a}_0 + \boldsymbol{w}_t.$$

onde I_k é a matriz identidade de dimensão $k \times k$.

- O processo é estacionário se todos valores de λ que satisfazem $\det(I_k - A_1\lambda - A_2\lambda^2 - \ldots - A_p\lambda^p) = 0$ estão fora do círculo unitário.

Estimação

Considerando $\boldsymbol{y}_i = (y_{i1}, y_{i2}, \ldots, y_{in})'$, $\boldsymbol{w}_i = (w_{i1}, w_{i2}, \ldots, w_{in})'$ e adicionando o vetor de interceptos $\boldsymbol{a}_0'$, o modelo VAR(p) pode ser representado através de um modelo de regressão aparentemente não correlacionado (SUR):

$$\boldsymbol{y} = (I_k \otimes X)\boldsymbol{\beta} + \boldsymbol{w}, \tag{2}$$

onde

- $\boldsymbol{y} = (\boldsymbol{y}_1', \boldsymbol{y}_2', \ldots, \boldsymbol{y}_k')'$ é um vetor coluna de dimensão nk;
- $X = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)'$ é uma matriz $n \times k(p+1)$ - $\boldsymbol{x}_t = (\boldsymbol{a}_0', \boldsymbol{y}_{t-1}', \boldsymbol{y}_{t-2}', \ldots, \boldsymbol{y}_{t-p}')'$ é um vetor coluna de dimensão $k(p+1)$;
- $\boldsymbol{\beta} = \text{vec}(A)$ - $A = (I_k, A_1, A_2, \ldots, A_p)$ é uma matriz $k \times k(p+1)$;
- $\boldsymbol{w} = (\boldsymbol{w}_1', \boldsymbol{w}_2', \ldots, \boldsymbol{w}_k')'$ é um vetor coluna de dimensão nk com média zero e matriz de covariância $(\boldsymbol{\tau} \otimes I_n)$.

Estimação Clássica

A partir da Equação (2), o estimador de $\boldsymbol{\beta}$ pode ser encontrado via Mínimos Quadrados Ordinários:

$$\hat{\boldsymbol{\beta}} = \left[(I_k \otimes X)'(I_k \otimes X)\right]^{-1} (I_k \otimes X)'\boldsymbol{y}. \tag{3}$$

Observações:

- O estimador em (3) será idêntico ao obtido via Mínimos Quadrados Generalizados, caso nenhuma restrição seja imposta ao modelo;
- Considerando $\boldsymbol{w}_t \sim N(\mathbf{0}, \boldsymbol{\tau})$, o estimador de MQO será idêntico ao estimador de máxima verossimilhança:
  - Estimador com distribuição assintótica normal com menor matriz de covariância assintótica.

Estimação Bayesiana

Considerando $\boldsymbol{w}_t \sim N(\mathbf{0}, \boldsymbol{\tau})$, temos:

$$\boldsymbol{y}|\boldsymbol{\beta}, \boldsymbol{\tau} \sim N((I_k \otimes X)\boldsymbol{\beta}, (\boldsymbol{\tau} \otimes I_n)).$$

Diferentes prioris podem ser consideradas para os parâmetros $(\boldsymbol{\beta}, \boldsymbol{\tau})$:

- Normal-Wishart prior: $p(\boldsymbol{\beta}, \boldsymbol{\tau}) = p(\boldsymbol{\beta}|\boldsymbol{\tau})p(\boldsymbol{\tau})$, onde $\boldsymbol{\beta} \sim N(\boldsymbol{\beta}_0, (\boldsymbol{\tau} \otimes \Omega))$ e $\boldsymbol{\tau} \sim IW(\boldsymbol{\tau}_0, \alpha)$;
- Jeffreys' prior: $p(\boldsymbol{\beta}, \boldsymbol{\tau}) \propto |\boldsymbol{\tau}|^{-(k+1)/2}$;
- Normal-Diffuse prior: $p(\boldsymbol{\beta}, \boldsymbol{\tau}) = p(\boldsymbol{\beta})p(\boldsymbol{\tau})$, onde $\boldsymbol{\beta} \sim N(\boldsymbol{\beta}_0, V_0)$ e $p(\boldsymbol{\tau}) \propto |\boldsymbol{\tau}|^{-(k+1)/2}$;

A distribuição a posteriori será, portanto: $p(\boldsymbol{\beta}, \boldsymbol{\tau}) \propto p(\boldsymbol{\beta}, \boldsymbol{\tau})p(\boldsymbol{y}|\boldsymbol{\beta}, \boldsymbol{\tau})$.

Especificação do Modelo

- Ajustar um modelo VAR envolve a seleção da ordem p.
- Assim como vimos no caso dos processos ARMA, a especificação do modelo pode ser realizada através da seleção de p tal que um ou mais critérios de informação sejam minimizados.
- Vale lembrar que critérios de informação são compostos por dois termos:
  - Medida de qualidade do ajuste (taxa de entropia ou erro de predição);
  - Termos de penalização.
- Minimizar critérios de informação permitirá a identificação de modelos com boa aderência aos dados, porém, parcimoniosos.

Especificação do Modelo

Seja $\tilde{\tau}$ a matriz de covariância estimada através dos resíduos, os principais critérios aplicados são descritos a seguir:

- Critério de Informação de Akaike (AIC):

$$\text{(M)AIC} = \ln[\det(\tilde{\tau})] + 2\frac{k^2p + k}{n}.$$

- Critério de Informação Bayesiano (BIC):

$$\text{(M)BIC} = \ln[\det(\tilde{\tau})] + \ln(n)\frac{k^2p + k}{n}.$$

- Critério de Informação de Hannan-Quinn (HQIC):

$$\text{(M)HQIC} = \ln[\det(\tilde{\tau})] + \ln(\ln(n))\frac{k^2p + k}{n}.$$

Especificação do Modelo

OBSERVAÇÕES:

- AIC e BIC são os critérios mais comumente aplicados;
- Entre o AIC e o BIC, este tende a penalizar mais a ordem p;
- HQIC apresenta uma penalização intermediária da ordem p em relação ao AIC e o BIC.

Previsão

- Suponha que estejamos no tempo h (**origem**) e tenhamos interesse em prever Y_{h+l} onde $l \geq 1$ (**horizonte**).
  - $\mathcal{J}_h$ é a informação disponível em h;
  - $y_h(l)$ é a previsão l-passos à frente de Y_t com origem h.
- Podemos obter previsões para um processo VAR(p) empregando a esperança condicional:

$$\boldsymbol{y}_h(l) = \mathrm{E}(\boldsymbol{Y}_{h+l}|\mathcal{J}_h) = \boldsymbol{a}_0 + A_1\boldsymbol{y}_h(l-1) + A_2\boldsymbol{y}_h(l-2) + \ldots + A_{h+l-p}\boldsymbol{y}_h(t-p),$$

onde $\boldsymbol{y}_h(s) = \boldsymbol{y}_{h+s}$ para $s \leq 0$.

Previsão

OBSERVAÇÕES:

- Assim como visto nos processos autorregressivos (AR(p)), as previsões podem ser computadas recursivamente para $h = 1, 2, \ldots$;

- Sob estacionariedade, as previsões convergem para média incondicional;

- Se o processo $\boldsymbol{W}_t$ é independente e identicamente distribuído, $\boldsymbol{y}_h(l)$ é ótimo: previsão em l-passos a frente em h com erro quadrático médio (EQM) mínimo;

- Se o processo $\boldsymbol{W}_t$ é apenas não correlacionado e não possui média condicional zero, $\boldsymbol{y}_h(l)$ é o melhor preditor linear mas pode não ser ótimo (EQM mínimo) em uma classe maior.

Análise impulso-resposta

- Tradicionalmente, a interação entre variáveis econômicas é estudada considerando o efeito de mudanças em uma variável sobre as demais variáveis de interesse.
- Nos modelos VAR, mudanças em uma variável são induzidas por resíduos diferentes de zero.
  - Choques podem ter uma interpretação estrutural se as restrições de identificação estrutural foram especificadas corretamente.
- Para estudar as relações entre as variáveis, os efeitos dos choques são analisados através do sistema.
- Análise impulso-resposta.

Análise impulso-resposta

- Na forma reduzida, vimos que a matriz de covariância $\boldsymbol{\tau}$, geralmente, não é diagonal.
  - As componentes de $\boldsymbol{w}_t$ podem ser contemporaneamente correlacionadas;
  - Na prática, é pouco provável que os choques w_{it}, $i = 1, 2, \ldots, k$, aconteçam isoladamente;
  - A análise pode não refletir o que de fato ocorre no sistema caso um choque ocorra.
- Tipicamente, tenta-se especificar choques estruturais e analisar seus efeitos.
  - Caso uma forma estrutural identificada tenha sido especificada, os resíduos correspondentes são os choques estruturais.

Análise impulso-resposta

- Um processo VAR estrutural de ordem p (SVAR(p)) pode ser escrito da seguinte forma:

$$U_1 \boldsymbol{y}_t = Q_1 \boldsymbol{y}_{t-1} + Q_2 \boldsymbol{y}_{t-2} + \ldots + Q_p \boldsymbol{y}_{t-p} + U_2 \tilde{\boldsymbol{w}}_t.$$

- Considerando um processo estacionário, as respostas ao impulso podem ser obtidas invertendo a representação do VAR estrutural:

$$\begin{aligned} \boldsymbol{y}_t &= (U_1 - Q_1 B - Q_2 B^2 - \ldots - Q_p B^p)^{-1} U_2 \tilde{\boldsymbol{w}}_t \\ &= \sum_{\ell=1}^{\infty} \Phi_\ell U_1^{-1} U_2 \tilde{\boldsymbol{w}}_{t-\ell} \\ &= \sum_{\ell=1}^{\infty} \Psi_\ell \tilde{\boldsymbol{w}}_{t-\ell}, \end{aligned}$$

onde $\Psi_\ell = \Phi_\ell U_1^{-1} U_2$ e $\Phi_\ell = \sum_{i=1}^{\ell} \Phi_{\ell-i} A_i$ com $\Phi_0 = I_k$ e $A_i = \mathbf{0}$ para $i > p$.

Análise impulso-resposta

OBSERVAÇÕES:

- Os elementos de Ψ_ℓ correspondem às respostas às inovações $\tilde{\boldsymbol{w}}_t$.
    - A resposta marginal de $y_{i(t+\ell)}$ a um impulso é dada pelos elementos ψ_{ij} das matrizes Ψ_ℓ.
- Como o processo é estacionário, os efeitos são transitórios.

Cointegração

Introdução

- No modelo Vetor Autorregressivo (VAR), assumimos que as séries temporais $\{Y_{1t}\}$, $\{Y_{2t}\}$, ..., $\{Y_{kt}\}$ são estacionárias.
  - Retornos de ativos;
  - Taxas de crescimento de séries temporais macroeconômicas.
- Em muitos casos, a teoria econômica descreve relações de equilíbrio entre variáveis integradas, por exemplo, $I(1)$.
- Combinações lineares de séries temporais integradas podem ser estacionárias.
  - Variáveis cointegradas: k variáveis cuja combinação linear é estacionária.

Introdução

- Tomar diferenças de um conjunto de variáveis não estacionárias simultaneamente a fim de torná-las estacionárias pode implicar:
    - Perda de informação;
    - Perda na capacidade preditiva;
    - Possível invalidade no processo de inferência.
- Considerando que a combinação linear é estacionária, diferenciar variáveis cointegradas para serem $I(0)$ implica erro de má especificação.

Cointegração

- Uma série temporal $\{\boldsymbol{Y}_t\}$ é dita cointegrada de ordem (d, b), $CI(d, b)$, $0 < b \leq d$, se cada componente Y_{it} é $I(d)$ mas alguma combinação linear $\boldsymbol{\kappa}'\boldsymbol{Y}_t$ é $I(d-b)$ para algum vetor constante $\boldsymbol{\kappa} \neq \mathbf{0}$.
  - $\boldsymbol{\kappa}$: vetor de cointegração;
  - Vetor de cointegração não é único. Por exemplo, para qualquer escalar c, $c\boldsymbol{\kappa}'\boldsymbol{Y}_t = \boldsymbol{\kappa}_*'\boldsymbol{Y}_t \sim I(d-b)$;
  - Alguma normalização é aplicada em $\boldsymbol{\kappa}$. Normalmente, κ_1 (coeficiente de Y_{1t}) é normalizado para 1;
  - Caso mais comum: $d = b = 1$.

Cointegração

- A combinação linear $\boldsymbol{\kappa}'\boldsymbol{Y}_t$ é comumente motivada pela teoria econômica e é denominada relação de equilíbrio de longo prazo.
  - A teoria da renda permanente implica cointegração entre consumo e renda;
  - A equação de Fisher implica cointegração entre taxa de juros nominal e inflação;
  - A hipótese das expectativas na estrutura a termo da taxa de juros implica cointegração na taxa de juros nominal em diferentes vencimentos.
- As relações de equilíbrio são ditas de longo prazo, pois as "forças econômicas" que agem em resposta a desvios do equilíbrio podem tomar um longo período no restabelecimento do equilíbrio.

Cointegração e Tendência Comum

- Se $\boldsymbol{Y}_t$ é cointegrada, pode haver $0 < m < k$ vetores de cointegração linearmente independentes.
  - Por exemplo, $k = 3$ e $m = 2$ vetores de integração:

$$\begin{aligned} \boldsymbol{\kappa}_1 &= (\kappa_{11}, \kappa_{12}, \kappa_{13})' \\ \boldsymbol{\kappa}_2 &= (\kappa_{21}, \kappa_{22}, \kappa_{23})'. \end{aligned}$$

Então,

$$\begin{aligned} \boldsymbol{\kappa}_1' \boldsymbol{Y}_t &= \kappa_{11} Y_{1t} + \kappa_{12} Y_{2t} + \kappa_{13} Y_{1t} \sim I(0) \\ \boldsymbol{\kappa}_2' \boldsymbol{Y}_t &= \kappa_{21} Y_{1t} + \kappa_{22} Y_{2t} + \kappa_{23} Y_{1t} \sim I(0). \end{aligned}$$

A matriz (3×2)

$$\tilde{\boldsymbol{\kappa}}' = \begin{bmatrix} \boldsymbol{\kappa}_1' \\ \boldsymbol{\kappa}_2' \end{bmatrix} = \begin{bmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} \\ \kappa_{31} & \kappa_{32} & \kappa_{33} \end{bmatrix}$$

forma uma base para o espaço de vetores de cointegração.

Cointegração e Tendência Comum

- Se $\boldsymbol{Y}_t$ é cointegrada com $0 < m < k$ vetores de cointegração, então há $k - m$ tendências estocásticas comuns.
  - Por exemplo, tome $\boldsymbol{Y}_t = (Y_{1t}, Y_{2t})' \sim I(1)$ e $\boldsymbol{\varepsilon} = (\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t})' \sim I(0)$ e suponha que $\boldsymbol{Y}_t$ é cointegrada com vetor de cointegração $\boldsymbol{\kappa} = (1, -\kappa_2)'$.
    A relação de cointegração pode ser representada da seguinte forma:
    $$\begin{aligned} y_{1t} &= \kappa_2 \sum_{s=1}^{t} \varepsilon_{1s} + \varepsilon_{3t} \\ y_{2t} &= \sum_{s=1}^{t} \varepsilon_{1s} + \varepsilon_{2t}, \end{aligned}$$
    onde $\sum_{s=1}^{t} \varepsilon_{1s}$ é a tendência comum. Então,
    $$\begin{aligned} \boldsymbol{\kappa}_t' \boldsymbol{Y}_t &= y_{1t} - \kappa_2 y_{2t} = \kappa_2 \sum_{s=1}^{t} \varepsilon_{1s} + \varepsilon_{3t} - \kappa_2 (\sum_{s=1}^{t} \varepsilon_{1s} + \varepsilon_{2t}) \\ &= \varepsilon_{3t} - \kappa_2 \varepsilon_{2t} \sim I(0). \end{aligned}$$

- Considere um processo VAR(p) para o vetor $\boldsymbol{Y}_t$:

$$\boldsymbol{y}_t = A_1\boldsymbol{y}_{t-1} + A_2\boldsymbol{y}_{t-2} + \ldots + A_p\boldsymbol{y}_{t-p} + \boldsymbol{w}_t. \tag{4}$$

  O processo é estacionário se todos os valores de λ que satisfazem $\det(I_k - A_1\lambda - A_2\lambda^2 - \ldots - A_p\lambda^p) = 0$ estão fora do círculo unitário.

- Supondo que $\boldsymbol{Y}_t$ seja $I(1)$ e possivelmente cointegrado, relações de cointegração podem ser analisadas ao representarmos o processo VAR(p) em (4) através de um Modelo Vetor de Correção de Erros (VECM):

$$\Delta\boldsymbol{y}_t = \Pi\boldsymbol{y}_{t-1} + \sum_{i=1}^{p-1} A_i^*\Delta\boldsymbol{y}_{t-i} + \boldsymbol{w}_t, \tag{5}$$

  onde $\Pi = (A_1 + A_2 + \ldots + A_p - I_k)$ e $A_j^* = -\sum_{i=j+1}^{p-1} A_i$, $j = 1, 2, \ldots, p-1$.

Cointegração e VAR

Observação:

No modelo VECM descrito na Equação (5), temos: $\Delta \boldsymbol{y}_t$ e suas defasagens são $I(0) \Rightarrow \Pi \boldsymbol{y}_{t-1}$ deve ser $I(0) \Rightarrow \Pi$ deve informar acerca das relações de cointegração.

Interpretação:

- Se $\text{rank}(\Pi) = 0$ ($\Pi = \mathbf{0}$), então os autovalores da matriz são 0.
  - $\boldsymbol{Y}_t$ é $I(1)$ e não há cointegração;
  - VECM reduz ao caso VAR$(p-1)$.
- Se $\text{rank}(\Pi) = m$, $0 < m < k$, então $\Pi = \boldsymbol{\gamma}\tilde{\boldsymbol{\kappa}}'$, onde $\tilde{\boldsymbol{\kappa}}$ contém m vetores de cointegração e $\boldsymbol{\gamma}$ é uma matriz $m \times k$.
  - $\boldsymbol{Y}_t$ é $I(1)$ com m vetores de cointegração linearmente independentes e $k - m$ tendências estocásticas comuns.

Procedimento de Johansen

- Testar para cointegração no caso multivariado recai sobre determinar o posto da matriz Π.
  - Necessário determinar o número de autovalores em Π diferentes de 0.
- Supondo que os autovalores (normalizados) sejam ordenados de tal forma que $1 > \lambda_1 > \lambda_2 > \ldots > \lambda_k > 0$, a estatística de teste é especificada da seguinte forma:

$$\lambda_{\text{traço}}(m_0) = -n \sum_{i=m_0+1}^{k} \ln(1 - \hat{\lambda}_i).$$

O teste, por sua vez, é especificado como a seguir:

$$\begin{aligned} H_0 &: \quad m = m_0 \\ H_1 &: \quad m > m_0. \end{aligned}$$

Procedimento de Johansen

OBSERVAÇÕES:

- Se $\text{rank}(\Pi) = m_0$, então $\hat{\lambda}_{m_0+1}, \ldots, \hat{\lambda}_m$ devem ser próximos de 0 e $\lambda_{\text{traço}}(m_0)$ deve ser pequeno;
- Se $\text{rank}(\Pi) > m_0$, então $\hat{\lambda}_{m_0+1}, \ldots, \hat{\lambda}_m$ devem ser diferentes de 0 (mas menores que 1) e $\lambda_{\text{traço}}(m_0)$ deve ser grande.

- A distribuição assintótica da estatística de teste não possui representação em forma fechada.
  - Valores críticos podem ser encontrados em Osterwald-Lenum (1992).

Procedimento de Johansen

- Johansen propõe um procedimento de teste sequencial que determina o número de vetores de cointegração:

1. $H_0 : m_0 = 0$ contra $H_1 : m_0 > 0$ é testado. Caso a hipótese nula não seja rejeitada, todos os autovalores de Π são iguais a 0. Caso contrário, há pelo menos um vetor de cointegração e segue-se para o passo seguinte;

2. $H_0 : m_0 = 1$ contra $H_1 : m_0 > 1$ é testado. Caso a hipótese nula não seja rejeitada, há apenas um vetor de cointegração. Caso contrário, há pelo menos dois vetores de cointegração e segue-se para o passo seguinte;

3. O procedimento é repetido até que a hipótese nula não seja rejeitada.

Causalidade de Granger

Definição

- Granger (1969) propôs uma noção de causalidade baseada no quão bem observações passadas de uma série temporal X_t podem prever valores futuros de outra série Y_t.

- Informação disponível em t, $\mathcal{J}_t$, é da forma $(y_t, x_t, y_{t-1}, x_{t-1}, \ldots, y_1, z_1)$.

- Dizemos que X causa (no sentido de Granger) Y com respeito a $\mathcal{J}_t$ se

$$\mathrm{Var}(Y_{t+1} - \mathcal{P}(Y_{t+1}|\mathcal{J}_t)) < \mathrm{Var}(Y_{t+1} - \mathcal{P}(Y_{t+1}|\mathcal{J}_t \backslash x_{\leq t})),$$

  onde $\mathcal{P}(Y_{t+1}|\mathcal{J}_t)$ é um predito linear ótimo condicional em $\mathcal{J}_t$ e $\mathcal{J}_t \backslash x_{\leq t}$ a denota a exclusão dos valores de $x_{\leq t}$ de $\mathcal{J}_t$.

  - X causa (no sentido de Granger) Y_t se $x_{\leq t}$ “ajuda” na previsão de Y.

Definição

- Podemos pensar, por exemplo, no caso do modelo ADL(p,q) - *Autoregressive Distributed Lag*:

$$y_t = \beta_0 + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \beta_j x_{t-j} + w_t. \tag{6}$$

  - X não causa (no sentido de Granger) Y se $\beta_j = 0$ para $j = 1, 2, \ldots, q$.
  - Teste F ($H_0 : \beta_1 = \ldots = \beta_q = 0$ vs. H_1 : pelo menos um coeficiente diferente de 0).

$$\begin{aligned} \text{Modelo restrito} \quad &: \quad y_t = \beta_0 + \sum_{i=1}^{p} \phi_i y_{t-i} + w_t \\ F \quad &= \quad \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-p-q-1)}, \end{aligned}$$

SSR_{ur} e SSR_r são a soma de quadrados dos resíduos dos modelos irrestrito (Equação (6)) e restrito, respectivamente.

Causalidade de Granger e VAR

- Natural pensar em Causalidade de Granger no contexto dos modelos VAR.

- Assume o seguinte processo VAR(p):

$$\begin{bmatrix} \boldsymbol{y}_{1t} \\ \boldsymbol{y}_{2t} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_1^{(11)} & \boldsymbol{A}_1^{(12)} \\ \boldsymbol{A}_1^{(21)} & \boldsymbol{A}_1^{(22)} \end{bmatrix} \begin{bmatrix} \boldsymbol{y}_{1t-1} \\ \boldsymbol{y}_{2t-1} \end{bmatrix} + \ldots + \begin{bmatrix} \boldsymbol{A}_p^{(11)} & \boldsymbol{A}_p^{(12)} \\ \boldsymbol{A}_p^{(21)} & \boldsymbol{A}_p^{(22)} \end{bmatrix} \begin{bmatrix} \boldsymbol{y}_{1t-p} \\ \boldsymbol{y}_{2t-p} \end{bmatrix} + \begin{bmatrix} \boldsymbol{w}_{1t} \\ \boldsymbol{w}_{2t} \end{bmatrix}.$$

- Se $A_j^{(12)} = 0$ para todo $1 \leq j \leq p$, então $\boldsymbol{y}_2$ não causa (no sentido de Granger) $\boldsymbol{y}_1$.

- A principal desvantagem neste tipo de verificação está na dependência da escolha correta do conjunto no qual estamos condicionando.
    - Na prática, não é possível determinar se o conjunto está corretamente definido.

Referências

del Negro, Marco. 2011. "Bayesian Macroeconometrics." In *The Oxford Handbook of Bayesian Econometrics*, Oxford University Press.

Granger, C. W. J. 1969. "Investigating Causal Relations by Econometric Models and Cross-spectral Methods." *Econometrica* 37 (3): 424–438.

Osterwald-Lenum, Michael. 1992. "A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics." *Oxford Bulletin of Economics and Statistics* 54 (3): 461–472.