

EAE1223: ECONOMETRIA III

AULA 6 - MODELOS VETORIAIS AUTORREGRESSIVOS

Luis A. F. Alvarez

22 de abril de 2024

VETORES ALEATÓRIOS E MATRIZ DE VARIÂNCIA-COVARIÂNCIA

- Um vetor (coluna) aleatório \mathbf{X} , com valores em \mathbb{R}^d , é uma função com domínio no espaço de probabilidade $(\Omega, \Sigma, \mathbb{P})$ e contradomínio em \mathbb{R}^d .
 - Posto de outra forma, um vetor aleatório \mathbf{X} é um vetor cujas entradas $\mathbf{X}_j, j = 1 \dots, d$, são variáveis aleatórias com valores reais.
- Para um vetor aleatório \mathbf{X} , a matriz de variância-covariância, $\mathbb{V}[\mathbf{X}]$, é definida como:

$$\mathbb{V}[\mathbf{X}] = \mathbb{E}[\mathbf{X}\mathbf{X}'] - \mathbb{E}[\mathbf{X}]\mathbb{E}[\mathbf{X}']$$

- Matriz de variância-covariância é $d \times d$, simétrica, positiva semidefinida, com entrada (i, j) iguais a:

$$\mathbb{V}[\mathbf{X}]_{i,j} = \text{cov}(\mathbf{X}_i, \mathbf{X}_j) .$$

MATRIZ DE COVARIÂNCIA ENTRE DOIS VETORES ALEATÓRIOS

- Seja \mathbf{X} um vetor aleatório em \mathbb{R}^d , e \mathbf{Y} um vetor aleatório em \mathbb{R}^p , definimos a matriz de covariância entre \mathbf{X} e \mathbf{Y} como:

$$\text{cov}(\mathbf{X}, \mathbf{Y}) = \mathbb{E}[\mathbf{X}\mathbf{Y}'] - \mathbb{E}[\mathbf{X}]\mathbb{E}[\mathbf{Y}'].$$

- Matriz $d \times p$ em que a entrada (i, j) é igual a:

$$\text{cov}(\mathbf{X}, \mathbf{Y})_{i,j} = \text{cov}(\mathbf{X}_i, \mathbf{Y}_j)$$

PROCESSO VETORIAL FRACAMENTE ESTACIONÁRIO

- Um processo estocástico vetorial é uma coleção de vetores aleatórios definidos em \mathbb{R}^d , indexados por um conjunto \mathcal{I} , isto é $\{\mathbf{X}_t : t \in \mathcal{I}\}$, onde cada \mathbf{X}_t é vetor aleatório em \mathbb{R}^d .
 - Cada entrada $j = 1 \dots d$ define um processo estocástico com valores reais $\{\mathbf{X}_{j,t} : t \in \mathcal{I}\}$ descrevendo a evolução da j -ésima entrada ao longo de \mathcal{I} .
- Um processo estocástico vetorial $\{\mathbf{X}_t : t \in \mathcal{T}\}$ indexado no tempo \mathcal{T} é dito fracamente estacionário se:
 1. $\mathbb{E}[\mathbf{X}_t] = \boldsymbol{\mu}$ para todo $t \in \mathcal{T}$.
 2. $\mathbb{V}[\mathbf{X}_t] = \boldsymbol{\Sigma}_0$ para todo $t \in \mathcal{T}$, com $\text{tr}(\boldsymbol{\Sigma}_0) < \infty$.
 3. $\text{cov}(\mathbf{X}_t, \mathbf{X}_{t-h}) = \boldsymbol{\Sigma}_h$, para todo $t \in \mathcal{T}$, $h \in \mathcal{N}$.
- Extensão do conceito de série de tempo estacionária para o caso vetorial.
- Pedimos estabilidade das covariâncias contemporâneas e extemporâneas entre as entradas dos vetores.
- **Obs:** se processo vetorial é fracamente estacionário, cada uma das entradas $\{\mathbf{X}_{j,t} : t \in \mathcal{T}\}$, $j = 1, \dots, d$, é fracamente estacionária.

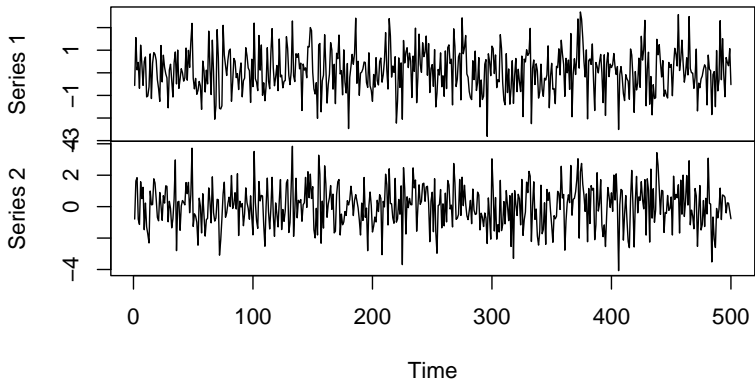
RUÍDO BRANCO VETORIAL

- Um processo estocástico vetorial $\{\mathbf{X}_t : t \in \mathcal{T}\}$ com valores em \mathbb{R}^d é dito um ruído branco vetorial se:
 - $\mathbb{E}[\mathbf{X}_t] = \mathbf{0}_{d \times 1}$, para todo $t \in \mathcal{T}$.
 - $\mathbb{V}[\mathbf{X}_t] = \Sigma_0$ para todo $t \in \mathcal{T}$, com $\text{tr}(\Sigma_0) < \infty$.
 - $\text{cov}(\mathbf{X}_t, \mathbf{X}_l) = \mathbf{0}_{d \times d}$, para todo $t \neq l$.
- No ruído branco vetorial, permitimos associação contemporânea entre as entradas do vetor, mas não há nem autodependência nem dependência cruzada entre as entradas no tempo.

RUÍDO BRANCO VETORIAL GAUSSIANO

$$\begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix} \stackrel{\text{iid}}{\sim} \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \right)$$

rb



MODELOS VETORIAIS AUTORREGRESSIVOS

- Considere d séries de tempo $\{X_{j,t} : t \in \mathbb{Z}\}$, $j = 1, \dots, d$.
- Dizemos que estas séries definem um processo vetorial autorregressivo de ordem p , ou VAR(p), se, para todo $j = 1, \dots, d$ e $t \in \mathbb{Z}$:

$$\begin{aligned} X_{j,t} &= a_{j,0} + \sum_{l=1}^p a_{j,1,l} X_{1,t-l} + \sum_{l=1}^p a_{j,2,l} X_{2,t-l} + \dots + \sum_{l=1}^p a_{j,d,l} X_{d,t-l} + \epsilon_{j,t} \\ &= a_{j,0} + \sum_{k=1}^d \sum_{l=1}^p a_{j,k,l} X_{k,t-l} + \epsilon_{j,t}, \end{aligned}$$

onde $\epsilon_t = (\epsilon_{1,t}, \epsilon_{2,t}, \dots, \epsilon_{d,t})'$ é um ruído branco vetorial.

- No VAR(p), assim como no AR(p), evolução em cada uma das d variáveis depende do que ocorreu nela mesma nos últimos p períodos.
- Mas além disso, trajetória depende do que ocorreu nos últimos p períodos **nas demais variáveis**.
- Série depende também de uma inovação $\epsilon_{j,t}$, imprevisível com base no passado, mas que pode estar contemporaneamente associada às demais inovações nas outras equações (choques comuns).

VAR(p) EM NOTAÇÃO VETORIAL

- Um VAR(p) pode ser escrito, compactamente, em notação vetorial.
- De fato, definindo $\mathbf{X}_t = (X_{1,t}, X_{2,t}, \dots, X_{d,t})'$, podemos escrever o sistema como

$$\mathbf{X}_t = \mathbf{a}_0 + \mathbf{A}_1 \mathbf{X}_{t-1} + \mathbf{A}_2 \mathbf{X}_{t-2} + \dots + \mathbf{A}_p \mathbf{X}_{t-p} + \boldsymbol{\epsilon}_t,$$

onde

$$\mathbf{a}_0 = \begin{bmatrix} a_{1,0} \\ a_{2,0} \\ \vdots \\ a_{d,0} \end{bmatrix} \quad \mathbf{A}_l = \begin{bmatrix} a_{1,1,l} & a_{1,2,l} & \dots & a_{1,d,l} \\ a_{2,1,l} & a_{2,2,l} & \dots & a_{2,d,l} \\ \vdots & \vdots & \dots & \vdots \\ a_{d,1,l} & a_{d,2,l} & \dots & a_{d,d,l} \end{bmatrix}$$

VAR(p) ESTACIONÁRIO

- Um VAR(p) é dito estacionário se o processo vetorial resultante é estacionário.
- Condição para estacionariedade do VAR(p) é que:

$$|z| \leq 1 \implies \det(\mathbb{I}_{d \times d} - \mathbf{A}_1 z - \mathbf{A}_2 z^2 \dots - \mathbf{A}_p z^p) \neq 0$$

- Todas as raízes do polinômio $\phi(z) = \det(\mathbb{I}_{d \times d} - \mathbf{A}_1 z - \mathbf{A}_2 z^2 \dots - \mathbf{A}_p z^p)$ devem se encontrar fora do círculo unitário.
- Nesta aula, focaremos na estimação de VAR(p) estacionários.
 - Portanto, cada uma das séries deverá estar devidamente estacionarizada.

ESTIMAÇÃO DO VAR(p)

- Dado um painel com observações de d séries durante T períodos, $\{\mathbf{X}_{j,t}\}_{t=1}^T$, como estimar os parâmetros de um VAR(p)?
- Maneira mais simples é estimar os parâmetros através de MQO, equação a equação.
- Isto é, estimamos os parâmetros da j -ésima equação resolvendo:

$$\min_{b_{0,j}, \{b_{j,k,l}\}_{k,l}} \sum_{t=p+1}^T \left(\mathbf{x}_{j,t} - b_{0,j} - \sum_{k=1}^d \sum_{l=1}^p b_{j,k,l} \mathbf{x}_{j,t-l} \right)^2$$

- Para cada equação, rodamos uma regressão com T observações e $p \times d + 1$ parâmetros.

SUR

- A estimação dos parâmetros de um VAR por MQO, equação a equação, é potencialmente ineficiente.
- Isso se deve ao fato de que os choques em uma equação j , por serem (potencialmente) contemporaneamente correlacionados com os choques das demais equações, podem conter informação relevante para estimar os parâmetros de outras equações
 - Erro em uma equação é informativo sobre a outra equação.
- Seja $\hat{\Sigma}_0$ um estimador preliminar de $\mathbb{V}(\epsilon_t)$.
 - Por exemplo, estimador da variância-covariância com base nos resíduos do estimador de MQO equação a equação.
- O estimador de *seemingly unrelated regression* (SUR) dos parâmetros de um VAR propõe-se a estimar os parâmetros do sistema **conjuntamente**, minimizando:

$$\min_{b_0, B_1, \dots, B_p} \sum_{t=p+1}^T \left(\mathbf{X}_t - \mathbf{b}_0 - \sum_{l=1}^p \mathbf{B}_l \mathbf{X}_{t-l} \right)' \hat{\Sigma}^{-1} \left(\mathbf{X}_t - \mathbf{b}_0 - \sum_{l=1}^p \mathbf{B}_l \mathbf{X}_{t-l} \right)$$

MÁXIMA VEROSSIMILHANÇA CONDICIONAL

- Sob a hipótese auxiliar:

$$\epsilon_t \stackrel{iid}{\sim} N(\mathbf{0}_{d \times 1}, \Sigma_0)$$

podemos calcular a verossimilhança de $\mathbf{X}_{t+p+1}, \dots, \mathbf{X}_T$, condicional a $\mathbf{X}_1, \dots, \mathbf{X}_p$.

- Estimador de máxima verossimilhança condicional estima **simultaneamente** \mathbf{a}_0 , os \mathbf{A}_l e Σ_0 .

RELAÇÃO ENTRE OS MÉTODOS DE ESTIMAÇÃO

- Se o espaço de parâmetros é irrestrito, no sentido de que os parâmetros \mathbf{a}_0 e os \mathbf{A}_I podem tomar qualquer valor real, os estimadores de MQO equação a equação, SUR, e máxima verossimilhança condicional são numericamente iguais.
- Se impomos restrições nos parâmetros \mathbf{a}_0 e os \mathbf{A}_I , estimador de MQO equação a equação é consistente, embora SUR e máxima verossimilhança condicional sejam mais eficientes.
- Se, além disso, impomos restrições na matriz Σ_0 , os três estimadores são consistentes, mas máxima verossimilhança condicional é o mais eficiente.

REFERÊNCIAS