

Séries de Tempo

José Fajardo

Fundação Getulio Vargas-EBAPE

Setembro 2011

- O uso de modelos univariados é limitado para expressar relações importantes em negócios.
- O vetor auto-regressivo permite que se expressem modelos completos e se estimem os parâmetros desse modelo.
- Os modelos em VAR definem restrições entre as equações do modelo. Estudar essas restrições e usá-las para identificar os parâmetros estruturais do VAR constitui um objetivo fundamental da metodologia.

FORMA ESTRUTURAL

- Pode-se expressar um modelo auto-regressivo de ordem p por um vetor com n variáveis endógenas, X_t , conectadas entre si por meio de uma matriz A :

$$AX_t = B_0 + \sum_{i=1}^p B_i X_{t-i} + B\varepsilon_t, \quad (1)$$

em que

A é uma matriz $n \times n$ que define as restrições contemporâneas entre as variáveis que constituem o vetor $n \times 1$, X_t ;

B_0 é um vetor de constantes $n \times 1$;

B_i são matrizes $n \times n$;

B é uma matriz diagonal $n \times n$ de desvios padrão;

ε_t é um vetor $n \times 1$ de perturbações aleatórias não correlacionadas entre si contemporânea ou temporalmente, isto é:

$$\varepsilon_t \sim i.i.d. (\mathbf{0}; I_n).$$

FORMA REDUZIDA

- A equação (1) expressa as relações entre as variáveis endógenas, a partir de um modelo econômico teoricamente estruturado: **forma estrutural**.
- Os choques ε_t são os choques estruturais porque afetam individualmente cada uma das variáveis endógenas.
- Os choques estruturais são considerados independentes entre si.
- Esse modelo é normalmente estimado em sua **forma reduzida**:

$$\begin{aligned}X_t &= A^{-1}B_0 + \sum_{i=1}^p A^{-1}B_i X_{t-i} + A^{-1}B\varepsilon_t = \\&= \Phi_0 + \sum_{i=1}^p \Phi_i X_{t-i} + e_t,\end{aligned}$$

em que

$$\Phi_i \equiv A^{-1}B_i, \quad i = 0, 1, \dots, p$$

$$B\varepsilon_t \equiv Ae_t.$$

EXEMPLO

- Seja um modelo bivariado de ordem 1:

$$y_t = b_{10} - a_{12}z_t + b_{11}y_{t-1} + b_{12}z_{t-1} + \sigma_y \varepsilon_{yt};$$

$$z_t = b_{20} - a_{21}y_t + b_{21}y_{t-1} + b_{22}z_{t-1} + \sigma_z \varepsilon_{zt}.$$

- Não pode ser estimado diretamente, já que as variáveis contemporâneas z_t e y_t são individualmente correlacionadas com os erros ε_{yt} ou ε_{zt} .
- O objetivo do VAR é desenvolver técnicas para evitar esse problema, visando-se a encontrar a trajetória da variável de interesse ante um choque estrutural.
- Hipóteses:
 - a) y_t e z_t são ambos estacionários;
 - b) $\varepsilon_{yt} \sim RB(0, 1)$ e $\varepsilon_{zt} \sim RB(0, 1)$;
 - c) $\varepsilon_{yt} \perp \varepsilon_{zt} \implies Cov(\varepsilon_{yt}, \varepsilon_{zt}) = 0$.

EXEMPLO

Em matrizes:

$$\underbrace{\begin{bmatrix} 1 & a_{12} \\ a_{21} & 1 \end{bmatrix}}_{\equiv A} \underbrace{\begin{bmatrix} y_t \\ z_t \end{bmatrix}}_{\equiv X_t} = \underbrace{\begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix}}_{\equiv B_0} +$$
$$\underbrace{\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}}_{\equiv B_1} \underbrace{\begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix}}_{\equiv X_{t-1}} + \underbrace{\begin{bmatrix} \sigma_y & 0 \\ 0 & \sigma_z \end{bmatrix}}_{\equiv B} \underbrace{\begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}}_{\equiv \varepsilon_t} \Rightarrow$$

$$AX_t = B_0 + B_1X_{t-1} + B\varepsilon_t.$$

A forma reduzida desse modelo simplificado é:

$$\begin{aligned} X_t &= \Phi_0 + \Phi_1 X_{t-1} + e_t; \\ \Phi_0 &\equiv A^{-1}B_0; \\ \Phi_1 &\equiv A^{-1}B_1; \\ Ae_t &\equiv B\varepsilon_t. \end{aligned} \tag{2}$$

EXEMPLO

- Condição de estabilidade: os autovalores de $(I - \Phi_1 L)$ fora do círculo unitário.
- Observe que:

$$\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} \equiv A^{-1} B \varepsilon_t = \begin{bmatrix} \frac{\sigma_y \varepsilon_{yt} - a_{12} \sigma_z \varepsilon_{zt}}{1 - a_{12} a_{21}} \\ \frac{\sigma_z \varepsilon_{zt} - a_{21} \sigma_y \varepsilon_{yt}}{1 - a_{12} a_{21}} \end{bmatrix}.$$

- Logo:

$$E(e_t) = \mathbf{0};$$

$$\text{Cov}(e_t) \equiv \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \frac{\sigma_y^2 + a_{12}^2 \sigma_z^2}{(1 - a_{12} a_{21})^2} & -\frac{a_{21} \sigma_y^2 + a_{12} \sigma_z^2}{(1 - a_{12} a_{21})^2} \\ -\frac{a_{21} \sigma_y^2 + a_{12} \sigma_z^2}{(1 - a_{12} a_{21})^2} & \frac{\sigma_z^2 + a_{21}^2 \sigma_y^2}{(1 - a_{12} a_{21})^2} \end{bmatrix}.$$

Os erros não são autocorrelacionados, pois:

$$\text{Cov}(e_{it}, e_{i(t-j)}) = E \left[\left(\frac{\sigma_i \varepsilon_{it} - a \sigma_{\sim i} \varepsilon_{\sim it}}{1 - a_{12} a_{21}} \right) \left(\frac{\sigma_i \varepsilon_{i(t-j)} - a \sigma_{\sim i} \varepsilon_{\sim i(t-j)}}{1 - a_{12} a_{21}} \right) \right]$$

$i = y, z; j \neq 0; a = a_{12}, a_{21}; \sim$ representam negação.

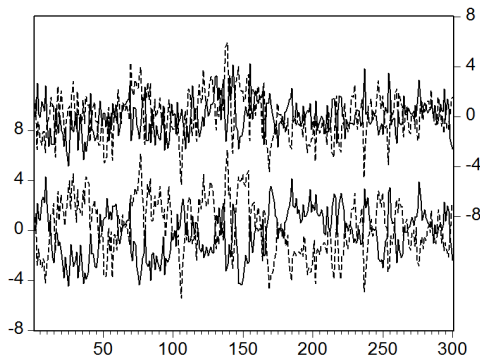


Figura: VAR (1): séries superiores - eixo à direita: $\Phi_0 = \mathbf{0}$; $\phi_{11} = \phi_{22} = 0,6$ e $\phi_{12} = \phi_{21} = 0,2$; séries inferiores: $\phi_{12} = \phi_{21} = -0,2$. A série pontilhada: z_t .

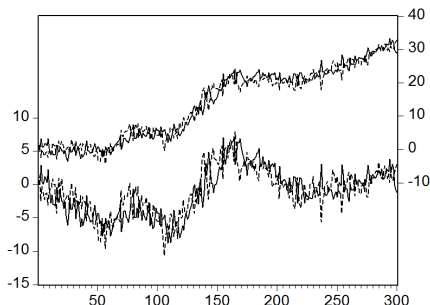


Figura: VAR (1): séries inferiores - eixo à esquerda: $\Phi = \mathbf{0}$; $\phi_{11} = \phi_{22} = 0,7$ e $\phi_{12} = \phi_{21} = 0,3$; séries superiores: $\phi_{10} = 0,5$. A série pontilhada: z_t .

Note que basta uma das constantes ser diferente de zero para que ambas as séries tenham um *drift*.

- As séries y_t e z_t movimentam-se conjuntamente, mesmo na presença de raiz unitária.
- Concentramo-nos na discussão sobre o caso de variáveis estacionárias, porém que Sims (1980) e Sims, Stock e Watson (1990) admitem a mistura de variáveis estacionárias e não estacionárias num modelo VAR. O VAR é uma metodologia interessada nas inter-relações entre as variáveis.
- Generalização:

$$X_t = \Phi_0 + \sum_{i=1}^p \Phi_i X_{t-i} + GZ_t + e_t,$$

em que

X_t é um vetor $n \times 1$ de variáveis endógenas, como anteriormente;

G é uma matriz de coeficientes $n \times g$;

Z_t é um vetor $g \times 1$ de variáveis exógenas que pode incluir variáveis determinísticas.

- Diferentemente dos modelos univariados, o VAR busca responder qual a trajetória da série, dado um **choque estrutural**.
- Por trajetória, deseja-SE conhecer o tempo que um choque afeta uma série, se ela muda de patamar ou não, para que patamar vai, entre outras informações.
- O VAR resulta na estimação de uma infinidade de coeficientes. Um $VAR(p)$, por exemplo, com n variáveis endógenas teria $n + n^2 p$ coeficientes a estimar, já que as matrizes Φ_i têm dimensão $n \times n$ e as n primeiras variáveis referem-se à constante, sem contar ainda os coeficientes de possíveis variáveis exógenas.
- Muitas vezes os coeficientes estimados serão estatisticamente insignificantes, até porque algumas variáveis são normalmente colineares, entretanto deve-se evitar impor restrições sobre os coeficientes, sob pena de perder informações relevantes, a menos que sejam restrições econômicas bem fundamentadas.

- Como selecionar a ordem p de um modelo VAR ? Que critérios podem ser utilizados nessa tarefa?
- Usar tantas defasagens quantas forem necessárias para obter "resíduos brancos" em todas as variáveis endógenas.
- Critério de informação: considere um $VAR(m)$, em que $m = 0, 1, 2, \dots, p_{\max}$. O problema é escolher a ordem p que minimiza a seguinte fórmula geral do critério de informação:

$$Cr(m) = \ln \left| \hat{\Gamma}_0 \right| + c_T \varphi(m),$$

em que

$$\hat{\Gamma}_0 = \frac{\sum_{t=1}^T \hat{e}_t \hat{e}_t'}{T};$$

c_T é uma seqüência que depende do tamanho da amostra;

$\varphi(m)$ é uma função que penaliza VAR de grandes ordens.

- O tamanho amostral tem de ser mantido constante para tornar o critério de informação comparável. Logo, o tamanho da amostra, comum a todas as ordens, será $T - p_{\max}$.
- A versão multivariada dos critérios AIC , BIC e HQ é:

$$AIC(m) = \ln \left| \hat{\Gamma}_0(m) \right| + \frac{2}{T} mn^2;$$

$$BIC(m) = \ln \left| \hat{\Gamma}_0(m) \right| + \frac{\ln T}{T} mn^2;$$

$$HQ(m) = \ln \left| \hat{\Gamma}_0(m) \right| + \frac{\ln \ln T}{T} 2mn^2,$$

em que mn^2 é o número total de parâmetros estimados em todas as equações.

- Outra maneira de escolher a ordem de defasagem, é aplicar testes seqüenciais para definir a ordem p do modelo *VAR*. Estabeleça o p_{\max} e considere $H_0 : \Phi_{p_{\max}} = \mathbf{0} \times H_1 : \Phi_{p_{\max}} \neq \mathbf{0}$. Se o teste não for rejeitado, repete-se o procedimento considerando $p_{\max} - 1$. Quando a nula for rejeitada, ter-se-á encontrada a ordem de defasagem do modelo.
- Problema: estabelecer p_{\max} . Se p_{\max} é muito pequeno, os resíduos estimados não serão um ruído branco. Contudo, se p_{\max} é muito grande, o impacto sobre a probabilidade de erros como um todo poderá ser severamente afetado, de modo que é difícil confiar nos intervalos de confiança gerados.

EXEMPLO

Seja $Z_t = [y_t, e_t, \pi_t, i_t]$, em que y_t representa o hiato do produto, e_t a variação cambial, π_t a taxa de inflação e i_t a taxa de juros (Selic), todas as variáveis em base mensal. Os critérios de Akaike e Schwarz são:

Defasagens	Akaike	Schwarz
1	-22,088	- 21,451
2	- 22,498	-21,342
3	-22,026	-20,343
4	-21,366	-19,146
5	-20,943	-18,179
6	-20,788	-17,470

EXEMPLO

Considere o exemplo de endogeneidade estudado no capítulo sobre GMM. Por conveniência, lembre-se de que o modelo foi simulado da seguinte forma:

$$\begin{aligned}y_t &= \phi y_{t-1} + \varepsilon_t; \\x_t &= \delta y_t + v_t,\end{aligned}$$

em que ε_t e v_t são ambos ruídos brancos independentes. Pode-se reescrever esse modelo na forma matricial:

$$\begin{bmatrix} 1 & 0 \\ -\delta & 1 \end{bmatrix} \begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} \phi & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ v_t \end{bmatrix}.$$

Reescrevendo o modelo na forma reduzida, tem-se:

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} \phi & 0 \\ \delta\phi & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} e_{yt} \\ e_{xt} \end{bmatrix}.$$

EXEMPLO

Ao estimar o modelo, a matriz Φ_1 conterá quatro coeficientes estimados. Espera-se que os coeficientes da coluna 2 sejam insignificantes e que os coeficientes da 1 satisfaçam aproximadamente a seguinte restrição:

$$\begin{aligned}\phi_{11,1} &= \phi; \\ \phi_{21,1} &= \delta\phi,\end{aligned}$$

em que δ e ϕ são valores usados na simulação. No caso da simulação, definiram-se $\delta = 0,1$ e $\phi = 0,8$.

Estimando-se o VAR (1), obtêm-se:

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} 0,794 & 0,097 \\ (0,035) & (0,063) \\ 0,077 & -0,002 \\ (0,032) & (0,058) \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \hat{e}_{yt} \\ \hat{e}_{xt} \end{bmatrix}.$$

$$\phi_{11,1} \simeq 0,8;$$

$$\phi_{21,1} \simeq 0,1 \times 0,8 = 0,08.$$

TESTANDO HIPÓTESES

- Testar hipóteses em modelos multivariados é semelhante ao caso univariado.
- Diferença: em vez de calcular a soma dos quadrados dos resíduos, calcula-se o determinante da matriz de covariância dos resíduos do modelo restrito e do não restrito.

Seja

$$X_t = \Phi_0 + \sum_{i=1}^p \Phi_i X_{t-i} + GZ_t + e_t.$$

Definition

O modelo anterior é estacionário se os autovalores da polinomial $\sum_{i=1}^p \Phi_i L^i$ estiverem dentro do círculo unitário.

TESTANDO HIPÓTESES

- 1 Estima-se o modelo sem restrição e calcula-se a matriz de covariância dos resíduos, $\hat{\Sigma}_u$;
- 2 Estima-se o modelo com restrição, excluindo $k \leq g$ variáveis exógenas e/ou m defasagens, e calcula-se $\hat{\Sigma}_r$;

- 3 Calcula-se a razão de verossimilhança da seguinte forma:

$$LR = (T - c) \left(\log |\hat{\Sigma}_r| - \log |\hat{\Sigma}_u| \right) \rightarrow \chi_r^2,$$

em que

T é o número de observações utilizadas na regressão;

$c = 1 + g + np$ é o número de parâmetros estimados em *cada* equação do sistema não restrito, incluindo a constante e as variáveis exógenas ;

$r = mn^2 + kn$ é o número de restrições no sistema.

- A inferência sobre os coeficientes estimados em modelos multivariados tem algumas particularidades bastante diferentes dos modelos univariados.
- Pode-se proceder a inferências estatísticas sobre os coeficientes individual ou coletivamente, mesmo ante a existência de variáveis endógenas não estacionárias de ordem 1, sempre que o modelo puder ser reescrito de maneira que os coeficientes sob inspeção passem a multiplicar variáveis estacionárias (veja Sims, Stock e Watson, 1990).

EXEMPLO

Sejam $y_t \sim I(1)$ e $z_t \sim I(1)$ da primeira linha de um VAR (2) bivariado:

$$y_t = \phi_{11,1}y_{t-1} + \phi_{12,1}z_{t-1} + \phi_{11,2}y_{t-2} + \phi_{12,2}z_{t-2} + \varepsilon_{yt}.$$

É possível fazer inferências sobre os coeficientes $\phi_{11,2}$ e $\phi_{12,2}$ desse modelo. Para isso, adicione e subtraia $\phi_{11,2}y_{t-1}$ e $\phi_{12,2}z_{t-1}$:

$$\begin{aligned} y_t &= \phi_{11,1}y_{t-1} + \phi_{11,2}y_{t-1} + \phi_{12,1}z_{t-1} + \phi_{12,2}z_{t-1} + \\ &+ \phi_{11,2}y_{t-2} - \phi_{11,2}y_{t-1} + \phi_{12,2}z_{t-2} - \phi_{12,2}z_{t-1} + \varepsilon_{zt} = \\ &= (\phi_{11,1} + \phi_{11,2})y_{t-1} + (\phi_{12,1} + \phi_{12,2})z_{t-1} + \\ &- \phi_{11,2}\Delta y_{t-1} - \phi_{12,2}\Delta z_{t-1} + \varepsilon_{yt}. \end{aligned}$$

Como Δy_{t-1} e Δz_{t-1} são variáveis estacionárias, pode-se testar conjuntamente por F – *Snedecor* a hipótese $\phi_{11,2} = \phi_{12,2} = 0$. Pouco se pode dizer a respeito dos coeficientes que multiplicam as variáveis não estacionárias y_{t-1} e z_{t-1} depois da transformação.

EXEMPLO

Sob condições convencionais, a matriz de coeficientes estimados

$\hat{\Phi} = [\hat{\Phi}_1 : \hat{\Phi}_2 : \dots : \hat{\Phi}_p]$ satisfaz:

$$\sqrt{T} \text{vec}(\hat{\Phi} - \Phi) \xrightarrow{d} N(\mathbf{0}, \Sigma_{\hat{\Phi}}).$$

- Por que se podem testar normalmente os coeficientes individuais de um modelo multivariado não estacionário?
- A não-estacionaridade emerge quando um dos autovalores da matriz de coeficientes está sobre ou fora do círculo unitário.
- Exemplo: Considere o seguinte VAR (1)

$$X_t = \begin{bmatrix} 0,7 & 0,3 \\ 0,3 & 0,7 \end{bmatrix} X_{t-1} + \varepsilon_t.$$

- Os autovalores de Φ_1 são 1 e 0,4. Trata-se de um modelo não estacionário. Esses autovalores estão associados aos coeficientes da matriz Φ_1 indiretamente. Logo, a estatística t permanece válida para a inferência individual dos coeficientes.

VERIFICAÇÃO: TESTE DE LJUNG-BOX

- Generalização dos testes de resíduos dos modelos univariados

$$H_0 : E(e_t e'_{t-j}) = \mathbf{0}, \text{ para todo } j = 1, 2, \dots, J > p.$$

versus

$$H_1 : E(e_t e'_{t-j}) \neq \mathbf{0}, \text{ para algum } j.$$

- A estatística do teste é parecida com a de Ljung-Box:

$$Q = T \sum_{j=1}^J \text{tr} \left(\hat{\Gamma}_j' \hat{\Gamma}_0^{-1} \hat{\Gamma}_j \hat{\Gamma}_0^{-1} \right) \xrightarrow{d} \chi_{n^2(J-p)}^2,$$

em que $\hat{\Gamma}_j = \frac{\sum_{t=j+1}^T \hat{e}_t \hat{e}'_{t-j}}{T}$ é a autocovariância na defasagem j .

- Alternativamente, usa-se a estatística de Ljung-Box ajustada:

$$Q = T^2 \sum_{j=1}^J \frac{1}{T-j} \text{tr} \left(\hat{\Gamma}_j' \hat{\Gamma}_0^{-1} \hat{\Gamma}_j \hat{\Gamma}_0^{-1} \right) \xrightarrow{d} \chi_{n^2(J-p)}^2.$$

EXEMPLO:

TESTE DE LJUNG-BOX

Considere o modelo não estacionário sem *drift* já simulado. O Eviews permite calcular os valores do testes automaticamente:

Defasagens	Q-Stat	Prob.	Adj Q-Stat	Prob.	df
1	0,196	NA	0,197	NA	NA
3	7,057	0,531	7,121	0,524	8
9	22,362	0,897	22,751	0,886	32
12	32,360	0,903	33,139	0,884	44

- df representa o número de graus de liberdade para a distribuição χ^2 .
- Embora o modelo seja não estacionário, as autocorrelações caem rapidamente.
- Se J é muito pequeno, a aproximação à distribuição será bastante pobre, ao passo que um valor alto para J pode resultar em perda de poder.

VERIFICAÇÃO:

TESTE DE BREUSCH-GODFREY

- O objetivo é testar se existe autocorrelação de resíduos no modelo:

$$\hat{e}_t = \Theta_1 \hat{e}_{t-1} + \Theta_2 \hat{e}_{t-2} + \cdots + \Theta_h \hat{e}_{t-h} + u_t$$

e verificar se:

$$H_0 : \Theta_1 = \Theta_2 = \cdots = \Theta_h = \mathbf{0} \times H_1 : \Theta_1 \neq \mathbf{0} \vee \Theta_2 \neq \mathbf{0} \vee \cdots \vee \Theta_h \neq \mathbf{0}$$

- Utiliza-se a regressão auxiliar:

$$\begin{aligned} \hat{e}_t = & \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \cdots + \Phi_p X_{t-p} + \\ & + \Theta_1 \hat{e}_{t-1} + \Theta_2 \hat{e}_{t-2} + \cdots + \Theta_h \hat{e}_{t-h} + u_t. \end{aligned}$$

- Trata-se de um teste LM. Para executá-lo, é necessário estimar \hat{e}_t contra o gradiente da função de verossimilhança. Aqui, esse gradiente é dado pela matriz $[X_{t-1} \ X_{t-2} \ \cdots \ X_{t-p} \ \hat{e}_{t-1} \ \hat{e}_{t-2} \ \cdots \ \hat{e}_{t-h}]$.
- Johnston e Dinardo comentam que uma importante característica desse teste é funcionar simultaneamente contra a hipótese alternativa de um processo $MA(q)$ para os resíduos.

VERIFICAÇÃO:

TESTE DE BREUSCH-GODFREY

- No primeiro estágio, estima-se o modelo completo por MQO, de forma que os \hat{e}_t s em que $t < 0$ são substituídos por zero. E calcula-se:

$$\hat{\Sigma}_u = \frac{\sum_{t=1}^T \hat{u}_t \hat{u}_t'}{T}.$$

- Em seguida, estima-se o modelo impondo a hipótese nula para obter os resíduos restritos, \hat{u}_t^r , e calcula-se:

$$\hat{\Sigma}_r = \frac{\sum_{t=1}^T \hat{u}_t^r \hat{u}_t^{r'}}{T}.$$

- O teste LM é definido como:

$$LM_h = T \left[n - \text{tr} \left(\hat{\Sigma}_u \hat{\Sigma}_r^{-1} \right) \right] \xrightarrow{d} \chi_{hn^2}^2.$$

EXEMPLO:

TESTE DE BREUSCH-GODFREY

- Continuando com o exemplo anterior. Estimado o modelo, verificam-se os resíduos usando o teste LM:

Defasagens	Q-Stat	Prob.	df
1	1,678	0,795	4
3	6,523	0,887	12
9	4,208	1,000	36
12	4,282	1,000	48

- O Eviews reporta o resultado das probabilidades com n^2 graus de liberdade. Por isso, é preciso olhar a tabela correta. A tabela anterior contém as probabilidades corretas, mas, se fossem usados n^2 , a hipótese nula não seria rejeitada. No caso do exemplo, como não se rejeita a nula com menos graus de liberdade do que seria correto, com mais razão não se rejeitaria se o número de graus de liberdade fosse ainda maior.

VERIFICAÇÃO:

TESTE DE NORMALIDADE

- A assimetria e a curtose têm distribuição normal. Defina os momentos 3 e 4 conforme segue:

$$m_3 = (m_{31}, m_{32}, \dots, m_{3n})', \text{ com } m_{3i} = \frac{\sum_{t=1}^T \varepsilon_{it}^3}{T};$$

$$m_4 = (m_{41}, m_{42}, \dots, m_{4n})', \text{ com } m_{4i} = \frac{\sum_{t=1}^T \varepsilon_{it}^4}{T}.$$

- Essas medidas têm a seguinte distribuição:

$$\sqrt{T} \begin{pmatrix} m_3 \\ m_4 - 3_n \end{pmatrix} \sim N \left[\mathbf{0}; \begin{pmatrix} 6I_n & \mathbf{0} \\ \mathbf{0} & 24I_n \end{pmatrix} \right],$$

em que $3_n = (3, 3, \dots, 3)'$ é um vetor $n \times 1$ de 3s.

VERIFICAÇÃO:

TESTE DE NORMALIDADE

- Na prática, primeiro deve-se obter a matriz de covariância dos resíduos estimados:

$$\hat{\Sigma}_e = \frac{\sum_{t=1}^T \left(\hat{e}_t - \overline{\hat{e}_t} \right) \left(\hat{e}_t - \overline{\hat{e}_t} \right)'}{T}.$$

- Calcule a matriz raiz quadrada: $\hat{\Sigma}_e^{\frac{1}{2}}$, para padronizar os resíduos:

$$\hat{e}_t^s = \hat{\Sigma}_e^{-\frac{1}{2}} \left(\hat{e}_t - \overline{\hat{e}_t} \right).$$

- Calcule a assimetria, m_3 , e a curtose, m_4 , dos n resíduos:

$$\hat{m}_3 = \left(\hat{m}_{31}, \hat{m}_{32}, \dots, \hat{m}_{3n} \right)', \text{ com } \hat{m}_{3i} = \frac{\sum_{t=1}^T (\hat{e}_{it}^s)^3}{T};$$

$$\hat{m}_4 = \left(\hat{m}_{41}, \hat{m}_{42}, \dots, \hat{m}_{4n} \right)', \text{ com } \hat{m}_{4i} = \frac{\sum_{t=1}^T (\hat{e}_{it}^s)^4}{T}.$$

VERIFICAÇÃO:

TESTE DE NORMALIDADE

- As estatísticas são dadas por:

$$s_3^2 = T \frac{\hat{m}_3' \hat{m}_3}{6}; \quad s_4^2 = T \frac{(\hat{m}_4 - 3_n)' (\hat{m}_4 - 3_n)}{24}.$$

- Ambas as estatísticas têm distribuição χ_n^2 sob a nula: $s_3^2 = s_4^2 = 0$.
- Alternativamente, pode-se usar a distribuição conjunta de ambos os testes, $JB_{2n} = s_3^2 + s_4^2$:

$$JB_{2n} = s_3^2 + s_4^2 \xrightarrow{d} \chi_{2n}^2.$$

- Análoga aos processos univariados.
- Quando se conhece o processo gerador de dados, a previsão h passos à frente é dada por:

$$E(X_{t+h}|I_t) \equiv X_{t+h|t} = \Phi_1 X_{t+h-1|t} + \Phi_2 X_{t+h-2|t} + \cdots + \Phi_p X_{t+h-p|t},$$

em que $X_{t+j|t} = X_{t+j}$ para $j \leq 0$.

- Transformando X_t num modelo de médias móveis infinito:

$$X_{t+h} = \left(I - \sum_{j=1}^p \Phi_j L^j \right)^{-1} e_{t+h} = e_{t+h} + \Psi_1 e_{t+h-1} + \Psi_p e_{t+h-2} + \cdots$$

- Conseqüentemente, a previsão correspondente é dada por:

$$X_{t+h|t} = \sum_{j=h}^{\infty} \Psi_j e_{t+h-j}.$$

- Desse modo, o erro de previsão será obtido extraíndo-se de X_{t+h} o termo $X_{t+h|t}$:

$$X_{t+h} - X_{t+h|t} = \sum_{j=0}^{h-1} \Psi_j e_{t+h-j},$$

em que $\Psi_0 = I_n$.

- A expectativa de previsão dos erros é zero. O EQM é dado por:

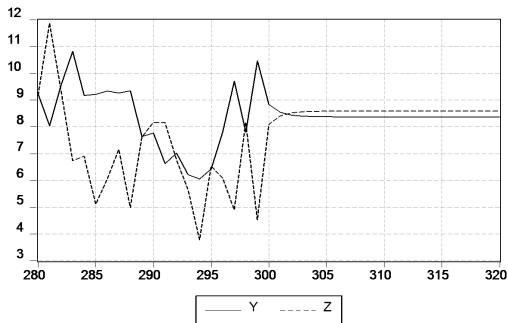
$$\Sigma_x(h) = E(X_{t+h} - X_{t+h|t})(X_{t+h} - X_{t+h|t})' = \sum_{j=0}^{h-1} \Psi_j \Sigma \Psi_j'.$$

- No processo é estacionário, a incerteza da previsão é limitada. Os processos integrados têm erro de previsão indeterminado em horizontes longos, mas isso não exclui a possibilidade de que a previsão de alguns componentes de variáveis integradas tenha o erro de previsão limitado também.
- Na presença de variáveis exógenas, entre as quais algumas determinísticas, pode-se estender a fórmula anterior facilmente.

Considere o VAR (1) estimado a partir do modelo simulado sem *drift*:

$$X_t = \begin{bmatrix} 0,694 & 0,332 \\ 0,298 & 0,676 \end{bmatrix} X_{t-1}.$$

Prevendo 20 passos à frente na Figura 3:

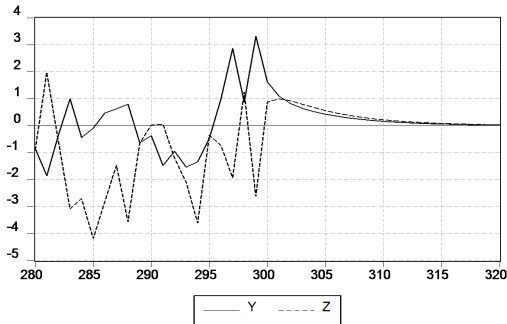


PREVISÃO

Em contraposição, considere o seguinte modelo $VAR(1)$, estimado a partir do modelo simulado estacionário:

$$X_t = \begin{bmatrix} 0,578 & 0,186 \\ 0,284 & 0,611 \end{bmatrix} X_{t-1}.$$

Prevendo 20 passos à frente, a Figura 4 mostra que a convergência para a média de longo prazo.



FUNÇÃO RESPOSTA AO IMPULSO

- O modelo VAR não permite identificar todos os parâmetros da forma estrutural, a menos que se imponham restrições adicionais.
- Para ver isso, observe que no sistema restrito dado pela equação (2) conseguem-se estimar os seis parâmetros na equação da média, $Var(e_1)$, $Var(e_2)$ e $Cov(e_1, e_2)$. No sistema primitivo há dez parâmetros a calcular: além dos oito coeficientes estruturais, há ainda as variâncias de cada um dos choques.
- Sims (1980) sugere um sistema recursivo para identificar o modelo. Trata-se de impor que alguns coeficientes sejam iguais a zero. Geralmente, usam-se argumentos econômicos para definir quais deles são iguais a zero.
- Sims impõe que o efeito *feedback* seja limitado. No caso mais simples, de um modelo bivariado, impor-se-ia, por exemplo, que $a_{12} = 0$:

$$y_t = b_{10} + b_{11}y_{t-1} + b_{12}z_{t-1} + \varepsilon_{yt};$$

$$z_t = b_{20} - a_{21}y_t + b_{21}y_{t-1} + b_{22}z_{t-1} + \varepsilon_{zt}.$$

FUNÇÃO RESPOSTA AO IMPULSO

Essa restrição é importante porque torna os parâmetros estruturais restantes identificáveis, conforme se observa no exemplo bivariado:

$$\begin{aligned} A^{-1} &= \begin{bmatrix} 1 & 0 \\ -a_{21} & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -a_{21} & 1 \end{bmatrix} \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \\ &\begin{bmatrix} 1 & 0 \\ -a_{21} & 1 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \\ &+ \begin{bmatrix} 1 & 0 \\ -a_{21} & 1 \end{bmatrix} \begin{bmatrix} \sigma_y & 0 \\ 0 & \sigma_z \end{bmatrix} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}. \end{aligned}$$

Sabendo-se ainda que $a_{12} = 0$, então os erros reduzidos ficam

$$\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \begin{bmatrix} \sigma_y \varepsilon_{yt} \\ \sigma_z \varepsilon_{zt} - a_{21} \sigma_y \varepsilon_{yt} \end{bmatrix},$$

de modo que

$$\begin{aligned} \text{var}(e_1) &= \sigma_y^2; \\ \text{cov}(e_1, e_2) &= -a_{21} \sigma_y^2; \text{ e } \text{var}(e_2) = \sigma_z^2 + a_{21}^2 \sigma_y^2. \end{aligned}$$

FUNÇÃO RESPOSTA AO IMPULSO

- Essas três equações combinam-se com as demais estimativas para identificar o modelo:

$$\begin{aligned}\phi_{10} &= b_{10}; & \phi_{20} &= b_{20} - b_{10}a_{21}; \\ \phi_{11} &= b_{11}; & \phi_{12} &= b_{12}; \\ \phi_{21} &= -a_{21}b_{11} + b_{21}; & \phi_{22} &= -a_{21}b_{12} + b_{22}.\end{aligned}$$

- A metodologia proposta por Sims pode ser generalizada para um vetor com n variáveis endógenas. Trata-se de uma maneira triangular de decompor os resíduos, chamada de decomposição de **Choleski**.
- No caso de n variáveis endógenas, a matriz de covariância é de dimensão $n \times n$. As condições de identificação requerem a imposição de $\frac{n^2-n}{2}$ restrições. O problema dessa imposição é definir a ordenação das variáveis, que é arbitrária, ainda que atribuída a razões econômicas. A ordenação das variáveis define a forma das restrições, de modo que diferentes ordenações geram diferentes restrições.

FUNÇÃO RESPOSTA AO IMPULSO

- Se os autovalores da polinomial $(I - \sum_{i=1}^p \Phi_i L^i)$ estiverem fora do círculo unitário, pode-se escrever um VAR (p) como um vetor de médias móveis infinito VMA (∞):

$$X_t = \bar{X} + \sum_{i=0}^{\infty} \Phi_1^i e_{t-i} = \bar{X} + \sum_{i=0}^{\infty} \frac{\Phi_1^i}{1 - a_{12}a_{21}} \begin{bmatrix} 1 & -a_{12} \\ -a_{21} & 1 \end{bmatrix} \begin{bmatrix} \sigma_y \varepsilon_{yt-i} \\ \sigma_z \varepsilon_{zt-i} \end{bmatrix}$$

em que $\bar{X} \equiv (I - \Phi_1)^{-1} \Phi_0$ é a média de longo prazo.

- Seja:

$$\Psi_i = \frac{\Phi_1^i}{1 - a_{12}a_{21}} \begin{bmatrix} 1 & -a_{12} \\ -a_{21} & 1 \end{bmatrix}.$$

- Desse modo:

$$\begin{aligned} X_t &= \bar{X} + \sum_{i=0}^{\infty} \Psi_i B \varepsilon_{t-i} = \\ &= \bar{X} + \sum_{i=0}^{\infty} \begin{bmatrix} \psi_{i,11} & \psi_{i,12} \\ \psi_{i,21} & \psi_{i,22} \end{bmatrix} \begin{bmatrix} \sigma_y \varepsilon_{yt-i} \\ \sigma_z \varepsilon_{zt-i} \end{bmatrix}. \end{aligned}$$

FUNÇÃO RESPOSTA AO IMPULSO

- Os elementos da matriz Ψ_i são os multiplicadores de impacto de um choque sobre as variáveis endógenas. Assim, o impacto total de um choque de ε_{yt} sobre y_{t+h} é dado pela soma dos coeficientes $\psi_{i,11}$, $i = 0, 1, 2, \dots, h$.
- Veja os gráficos da função resposta ao impulso. Inicialmente, impõe-se à simulação que $a_{12} = 0$ e $a_{21} = 0,8$. Em seguida, estima-se o modelo simulado e gera-se a função resposta ao impulso usando a decomposição de Choleski que assume $a_{12} = 0$.

INTERVALO DE CONFIANÇA

- A função resposta ao impulso é calculada mediante coeficientes estimados. Logo, é claro que há um intervalo de confiança a ser considerado nessas estimativas. Esse intervalo pode ser calculado de forma analítica ou por métodos de experimentos de Monte Carlo.
- O método analítico torna-se bem complicado quando se imagina um problema multivariado, em razão das covariâncias cruzadas.

INTERVALO DE CONFIANÇA

- 1 Estime o modelo multivariado e armazene os resíduos estimados, $\{\hat{e}_t\}$;
- 2 Sorteie os resíduos armazenados com reposição e simule uma nova série usando as matrizes Φ estimadas no passo anterior. Por exemplo, no caso de um VAR (1):

$$\hat{X}_t = \hat{\Phi}_0 + \hat{\Phi}_1 \hat{X}_{t-1} + \hat{e}_t;$$

- 3 Reestime o modelo e a nova função resposta ao impulso;
- 4 Repita o processo milhares de vezes;
- 5 Para construir um intervalo com 95% de confiança, exclua 2,5% das menores e maiores respostas.

DECOMPOSIÇÃO DA VARIÂNCIA

Considere o exemplo das seções anteriores, VAR (1) com duas variáveis endógenas, y e z :

$$X_{t+h} = \bar{X} + \sum_{i=0}^{\infty} \Psi_i \varepsilon_{t+h-i}.$$

Calcule o erro de previsão:

$$X_{t+h} - E_t(X_{t+h}) = \sum_{i=0}^{h-1} \Psi_i \varepsilon_{t+h-i}.$$

Esmiuçando apenas y_{t+h} :

$$\begin{aligned} y_{t+h} - E_t(y_{t+h}) &= \psi_{0,11} \varepsilon_{yt+h} + \psi_{1,11} \varepsilon_{yt+h-1} + \cdots + \psi_{h-1,11} \varepsilon_{yt+1} + \\ &\quad + \psi_{0,12} \varepsilon_{zt+h} + \psi_{1,12} \varepsilon_{zt+h-1} + \cdots + \psi_{h-1,12} \varepsilon_{zt+1}. \end{aligned}$$

Logo:

$$\begin{aligned} \sigma_y^2(h) &= \sigma_y^2(\psi_{0,11}^2 + \psi_{1,11}^2 + \cdots + \psi_{h-1,11}^2) + \\ &\quad + \sigma_z^2(\psi_{0,12}^2 + \psi_{1,12}^2 + \cdots + \psi_{h-1,12}^2). \end{aligned}$$

DECOMPOSIÇÃO DA VARIÂNCIA

Agora, pode-se decompor a variância do erro de previsão em seus diversos elementos. Dividindo-se ambos os lados por $\sigma_y^2(h)$:

$$1 = \frac{\sigma_y^2 (\psi_{0,11}^2 + \psi_{1,11}^2 + \cdots + \psi_{h-1,11}^2)}{\sigma_y^2(h)} + \frac{\sigma_z^2 (\psi_{0,12}^2 + \psi_{1,12}^2 + \cdots + \psi_{h-1,12}^2)}{\sigma_y^2(h)}.$$

Decomposição da variância de $y_{t+h} - E_t(y_{t+h})$

h	$\sigma_y(h)$	$\frac{\sigma_y^2(\psi_{0,11}^2 + \psi_{1,11}^2 + \cdots + \psi_{h-1,11}^2)}{\sigma_y^2(h)}$	$\frac{\sigma_z^2(\psi_{0,12}^2 + \psi_{1,12}^2 + \cdots + \psi_{h-1,12}^2)}{\sigma_y^2(h)}$
1	0,956	100,00	0,00
2	1,063	94,53	5,47
3	1,116	88,09	11,91
4	1,152	83,24	16,76
5	1,176	80,03	19,97

DECOMPOSIÇÃO DA VARIÂNCIA

O mesmo é feito com a variável z_{t+h} :

Decomposição da variância de $z_{t+h} - E_t(z_{t+h})$

h	$\sigma_z(h)$	$\frac{\sigma_y^2(\psi_{0,21}^2 + \psi_{1,21}^2 + \dots + \psi_{h-1,21}^2)}{\sigma_z^2(h)}$	$\frac{\sigma_z^2(\psi_{0,22}^2 + \psi_{1,22}^2 + \dots + \psi_{h-1,22}^2)}{\sigma_z^2(h)}$
1	1,244	32,96	67,04
2	1,404	25,51	74,49
3	1,462	29,04	70,96
4	1,490	27,07	72,93
5	1,506	26,06	73,94

TESTE DE GRANGER-CAUSALIDADE

- uma variável é capaz de prever outra? em que condições? Em outras palavras, a questão fundamental é saber se o escalar y ajuda a **prever** o escalar z . Se isso não acontece, então diz-se que y **não-Granger-cause** z .
- O teste tem o sentido de previsão, não de causalidade econômica, apesar do nome.
- A forma de responder a essa pergunta é usando um teste F convencional, válido quando os coeficientes de interesse puderem ser escritos de modo a multiplicar variáveis estacionárias.

TESTE DE GRANGER-CAUSALIDADE

- O teste é feito da seguinte maneira:
- ① Estime $z_t = \phi_{20} + \sum_{i=1}^p \phi_{i,21} y_{t-i} + \sum_{i=1}^p \phi_{i,22} z_{t-i} + e_{2t}$;
- ② Teste se y não-Granger-causa z usando o teste de F , sob:

$$H_0 : \phi_{1,21} = \phi_{2,21} = \dots = \phi_{p,21} = 0 \times H_1 : \phi_{i,21} \neq 0, i = 1, 2, \dots, p,$$

em que a estatística do teste é dada por:

$$S_1 = \frac{\frac{(e_r^2 - e_u^2)}{p}}{\frac{e_u^2}{T-2p-1}} \xrightarrow{d} F(p, T-2p-1),$$

em que r representa restrito e u , não restrito. Se $S_1 > F^{5\%}$, rejeita-se a hipótese nula de que y não-Granger-causa z .

- ③ Um teste equivalente é:

$$S_2 = \frac{T(e_r^2 - e_u^2)}{e_u^2} \xrightarrow{d} \chi_p^2.$$

Rejeita-se a nula se $S_2 > \chi_p^2(5\%)$.

TESTE DE GRANGER-CAUSALIDADE

- Teste de causalidade de Granger não é a mesma coisa que teste de exogeneidade. Para que z_t seja exógeno a y_t , é preciso que z_t não seja afetado contemporaneamente por y_t . A forma reduzida do VAR não permite que se faça esse tipo de teste. O teste de causalidade de Granger inclui, pois, valores correntes e passados de y_t sobre z_t .
- Pode-se fazer o mesmo teste em contextos de mais variáveis, cujo nome é teste de bloco-exogeneidade. Estima-se o modelo com e sem restrição e utiliza-se o teste F , como visto anteriormente.
- É importante observar que, em sistemas com $n > 2$, o teste de causalidade é mais complicado de ser feito, e a interpretação necessita de maiores cuidados. O problema que pode ocorrer ao não rejeitar a nula é não perceber a dinâmica mais complicada do modelo em que uma variável, apesar de não causar diretamente outra (por exemplo, y_{2t} não Granger-causa y_{1t}), pode causá-la indiretamente. Isso ocorrerá quando y_{2t} causar y_{3t} , que, por sua vez, causa y_{1t} . O teste de Granger-causalidade não foi desenvolvido para esse tipo de caso.

EXEMPLO

Se o Banco Central, de fato, determina a taxa Selic, então alterações na taxa efetiva praticada no mercado financeiro deveriam responder às alterações na meta definida pelo Banco Central. Por outro lado, não há razões para acreditar que alterações nas taxas de mercado afetem a meta definida pelo Banco Central.

Teste Ng-Perron

	MZa	MZt	MSB	MPT
Meta - Nível	-5,706	-1,440	0,252	5,030
Meta - Primeira Diferença	-943,032	-21,714	0,023	0,025
Efetiva - Nível	-5,153	-1,345	0,261	5,420
Efetiva - Primeira Diferença	-14,426	-2,685	0,186	1,698
1%	-13,800	-2,580	0,174	1,780
5%	-8,100	-1,980	0,233	3,170
10%	-5,700	-1,620	0,275	4,450

Como os dados são diários e apenas os para dias úteis, utilizam-se 20 defasagens de cada variável. As variáveis estão em primeira diferença, então os resultados informam se a variação de uma variável ajuda a explicar a variação da outra defasagem.

Teste de Causalidade de Granger			
Hipótese Nula	Obs	Estatística F	Probabilidade
META não-Granger-causa EFETIVA	1858	4,503	0,000
EFETIVA não-Granger-causa META		0,395	0,992

- Existem outras formas de definir restrições sobre a matriz A , de modo a identificar os parâmetros estruturais. Usa-se a teoria econômica para definir as restrições da matriz A completamente, com risco de sobreidentificação.
- Para entender o procedimento, considere um VAR (1) com n variáveis:

$$AX_t = B_0 + B_1X_{t-1} + B\varepsilon_t.$$

- Quando se estima a forma reduzida, obtêm-se os resíduos \hat{e}_t , tal que:

$$B\hat{e}_t = A\hat{e}_t.$$

- O problema é restringir o sistema, de forma a recuperar ε_t conforme a hipótese de que cada resíduo do sistema estrutural é independente um do outro.

- Considere a matriz de covariância dos erros da forma reduzida:

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix},$$
$$\hat{\sigma}_{ij} = \frac{1}{T} \sum_{t=1}^T \hat{e}_{it} \hat{e}_{jt}.$$

- Σ é simétrica $\rightarrow \frac{n(n+1)}{2}$ elementos diferentes. Elementos da diagonal de A são unitários $\rightarrow (n^2 - n)$ elementos desconhecidos. Desconhecem-se as n variâncias de ε . Logo, há n^2 elementos desconhecidos e $\frac{n(n+1)}{2}$ conhecidos.
- Restrições a impor: $n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}$, embora insuficiente para identificação exata, pois existem não-linearidades no sistema de equações.

- Logo:

$$BB' = \text{Var}(Ae_t) = A \text{Var}(e_t) A'.$$

- Suponha o seguinte sistema de equações, em que se admite conhecer a matriz de covariância de e_t :

$$\begin{aligned} \begin{bmatrix} \sigma_{\varepsilon_1}^2 & 0 \\ 0 & \sigma_{\varepsilon_2}^2 \end{bmatrix} &= \begin{bmatrix} 1 & a_{12} \\ a_{21} & 1 \end{bmatrix} \begin{bmatrix} 0,5 & 0,4 \\ 0,4 & 0,6 \end{bmatrix} \begin{bmatrix} 1 & a_{21} \\ a_{12} & 1 \end{bmatrix} = \\ &= \begin{bmatrix} 0,5 + 0,4a_{12} & 0,4 + 0,6a_{12} \\ 0,5a_{21} + 0,4 & 0,4a_{21} + 0,6 \end{bmatrix} \begin{bmatrix} 1 & a_{21} \\ a_{12} & 1 \end{bmatrix} = \\ \begin{bmatrix} 0,5 + 0,8a_{12} + 0,6a_{12}^2 & 0,5a_{21} + 0,4a_{12}a_{21} + 0,4 + 0,6a_{12} \\ 0,5a_{21} + 0,4 + 0,4a_{12}a_{21} + 0,6a_{12} & 0,5a_{21}^2 + 0,8a_{21} + 0,6 \end{bmatrix} \end{aligned}$$

Exemplo

- Há três equações independentes com quatro incógnitas $(a_{12}, a_{21}, \sigma_{\varepsilon_1}^2, \sigma_{\varepsilon_2}^2)$:

$$\sigma_{\varepsilon_1}^2 = 0,5 + 0,8a_{12} + 0,6a_{12}^2;$$

$$0 = 0,5a_{21} + 0,4 + 0,4a_{12}a_{21} + 0,6a_{12};$$

$$\sigma_{\varepsilon_2}^2 = 0,5a_{21}^2 + 0,8a_{21} + 0,6.$$

- Impondo a restrição de Choleski, $a_{12} = 0$, calculam-se:

$$\sigma_{\varepsilon_1}^2 = 0,5;$$

$$0 = 0,5a_{21} + 0,4 \implies a_{21} = -0,8$$

$$\sigma_{\varepsilon_2}^2 = 0,5a_{21}^2 + 0,8a_{21} + 0,6 \implies \sigma_{\varepsilon_2}^2 = 0,28.$$

- A função resposta ao impulso. No caso de um VAR (1) bivariado, essa função é:

$$\Psi_i = \Phi_1^i \begin{bmatrix} 1 & 0 \\ -a_{21} & 1 \end{bmatrix}.$$

Exemplo

Outra restrição: uma inovação em ε_{2t} pode ter um efeito unitário em y_{1t}
 $\rightarrow a_{12} = 1$:

$$\sigma_{\varepsilon_1}^2 = 1,9;$$

$$a_{21} = -\frac{10}{9};$$

$$\sigma_{\varepsilon_2}^2 = 0,5\frac{100}{81} - 0,8\frac{10}{9} + 0,6 = 0,328.$$

Exemplo

Há casos em que as restrições podem não gerar resultados únicos. Por exemplo: a correlação cruzada no modelo bivariado estrutural é única, $a_{12} = a_{21} = a$:

$$0 = 1,1a + 0,4 + 0,4a^2 \implies a = \begin{cases} -0.173 \\ -0.928 \end{cases} ;$$

$$\sigma_{\varepsilon_1}^2 = 0,5 + 0,8a + 0,6a^2 \implies \sigma_{\varepsilon_1}^2 = \begin{cases} 0.414 \\ 0.460 \end{cases} ;$$

$$\sigma_{\varepsilon_2}^2 = 0,5a^2 + 0,8a + 0,6 \implies \sigma_{\varepsilon_2}^2 = \begin{cases} 0.477 \\ 0.288 \end{cases} .$$

DECOMPOSIÇÃO DE BLANCHARD E QUAH

- Forma de identificação com base em restrições determinadas pela teoria econômica.
- Imposição de restrições a respeito do comportamento de longo prazo de uma variável a partir do choque estrutural em uma outra variável.
- Suponha um vetor $X_t = (\Delta y_t, u_t)'$, em que y_t representa o logaritmo do produto, e u_t , a taxa de desemprego.
- As variáveis respondem a um vetor de choques dado por $\varepsilon_t = (\varepsilon_{t,d}, \varepsilon_{t,s})$.
- O produto é afetado por choques de demanda, e o desemprego, por choques de oferta.
- Os choques entre oferta e demanda não são correlacionados.
- Os choques de demanda têm efeitos temporários; os choques de oferta têm efeitos permanentes.

DECOMPOSIÇÃO DE BLANCHARD E QUAH

Normalizando $B = I$, considere

$$AX_t = \sum_{j=1}^p B_j X_{t-j} + \varepsilon_t,$$

em que $\text{var}(\varepsilon_t) = I$.

O modelo na forma reduzida fica:

$$X_t = \sum_{j=1}^p A^{-1} B_j X_{t-j} + A^{-1} \varepsilon_t = \sum_{j=1}^p \Phi_j X_{t-j} + e_t,$$

em que

$$A^{-1} B_j \equiv \Phi_j;$$

$$A e_t = \varepsilon_t.$$

Há duas representações equivalentes de X_t , dependendo de o modelo partir da forma reduzida ou estrutural.

DECOMPOSIÇÃO DE BLANCHARD E QUAH

Partindo da forma reduzida, pode-se escrevê-lo como:

$$\begin{aligned}\left(I - \sum_{j=1}^p \Phi_j L^j\right) X_t &= e_t \implies X_t = \left(I - \sum_{j=1}^p \Phi_j L^j\right)^{-1} e_t \implies \\ X_t &= \left(\sum_{j=0}^{\infty} C_j L^j\right) e_t,\end{aligned}\tag{3}$$

em que

$$C_0 \equiv I;$$

$$\left(I - \sum_{j=1}^p \Phi_j L^j\right)^{-1} \equiv \sum_{j=0}^{\infty} C_j L^j;$$

$$\text{var}(e_t) = \Sigma.$$

DECOMPOSIÇÃO DE BLANCHARD E QUAH

Partindo da forma estrutural, obtém-se:

$$\begin{aligned}\left(A - \sum_{j=1}^p B_j L^j\right) X_t &= \varepsilon_t \implies X_t = \left(A - \sum_{j=1}^p B_j L^j\right)^{-1} \varepsilon_t \implies \\ X_t &= \left(\sum_{j=0}^{\infty} A_j L^j\right) \varepsilon_t,\end{aligned}\tag{4}$$

em que $\left(A - \sum_{j=1}^p B_j L^j\right)^{-1} \equiv \sum_{j=0}^{\infty} A_j L^j$.

DECOMPOSIÇÃO DE BLANCHARD E QUAH

Dado que $C_0 = I$, das equações (4) e (3) resultam as seguintes relações:

$$\begin{aligned}A_0 \varepsilon_t &= e_t \implies A_0 \varepsilon_{t-j} = e_{t-j}; \\A_j \varepsilon_{t-j} &= C_j e_{t-j} \implies A_j \varepsilon_{t-j} = C_j A_0 \varepsilon_{t-j} \therefore \\A_j &= C_j A_0, \quad j > 0.\end{aligned}$$

DECOMPOSIÇÃO DE BLANCHARD E QUAH

Qual a relação entre a matriz A_0 e a matriz A ?

$$\begin{aligned} A \left(I - \sum_{j=1}^p A^{-1} B_j L^j \right) X_t &= \varepsilon_t \implies X_t = \left(I - \sum_{j=1}^p A^{-1} B_j L^j \right)^{-1} A^{-1} \varepsilon_t = \\ &= \left(\sum_{j=0}^{\infty} C_j L^j \right) A^{-1} \varepsilon_t = \sum_{j=0}^{\infty} C_j A^{-1} \varepsilon_{t-j}. \end{aligned}$$

Comparando essa equação com a equação (4), conclui-se que:

$$\begin{aligned} A_j &= C_j A^{-1} \implies \\ C_j A_0 &= C_j A^{-1} \therefore \\ A_0 &= A^{-1}. \end{aligned}$$

Essa relação é importante porque vão ser impostas restrições à matriz A_0 .

DECOMPOSIÇÃO DE BLANCHARD E QUAH

Do termo $A_0\varepsilon_t$, verifica-se que:

$$\begin{aligned}A_0\varepsilon_t &= \varepsilon_t \implies \\A_0A_0' &= \Sigma.\end{aligned}$$

Esse sistema de equações resulta, como já foi visto, em $\frac{n(n+1)}{2}$ equações. No caso bivariado, são explicitamente dadas por:

$$\begin{aligned}a_{11,0}^2 + a_{12,0}^2 &= \text{var}(e_1); \\a_{11,0}a_{21,0} + a_{12,0}a_{22,0} &= \text{cov}(e_1, e_2); \\a_{21,0}^2 + a_{22,0}^2 &= \text{var}(e_2).\end{aligned}$$

DECOMPOSIÇÃO DE BLANCHARD E QUAH

Aqui introduzem-se as restrições de longo prazo. Considere a equação (4) no modelo bivariado:

$$X_t = \sum_{j=0}^{\infty} \begin{bmatrix} a_{11,j} & a_{12,j} \\ a_{21,j} & a_{22,j} \end{bmatrix} \varepsilon_{t-j}.$$

Para que o choque de demanda não tenha efeitos de longo prazo sobre o produto, impõe-se:

$$\sum_{j=0}^{\infty} a_{11,j} = 0.$$

Tendo-se A_0 e C_j , este obtido da estimação do modelo reduzido, e valendo a relação $A_j = C_j A_0$, pode-se obter A_j .

DECOMPOSIÇÃO DE BLANCHARD E QUAH

Para operacionalizar o procedimento:

- 1 Estime o modelo $VAR(p)$ e obtenha as matrizes $\hat{\Phi}_j$;
- 2 Obtenha as matrizes C_j resolvendo:

$$\left(\sum_{j=0}^{\infty} \hat{C}_j L^j \right) = \left(I - \sum_{j=1}^p \hat{\Phi}_j L^j \right)^{-1};$$

- 3 Sabendo que $A_j = C_j A_0$, imponha a restrição de longo prazo;
- 4 Resolva o sistema de equações e encontre os coeficientes de A_0 .

EXEMPLO

$$\begin{bmatrix} a_{11,j} & a_{12,j} \\ a_{21,j} & a_{22,j} \end{bmatrix} = \begin{bmatrix} c_{11,j} & c_{12,j} \\ c_{21,j} & c_{22,j} \end{bmatrix} \begin{bmatrix} a_{11,0} & a_{12,0} \\ a_{21,0} & a_{22,0} \end{bmatrix} = \\ = \begin{bmatrix} a_{11,0}c_{11,j} + a_{21,0}c_{12,j} & a_{12,0}c_{11,j} + a_{22,0}c_{12,j} \\ a_{11,0}c_{21,j} + a_{21,0}c_{22,j} & a_{12,0}c_{21,j} + a_{22,0}c_{22,j} \end{bmatrix}$$

Impondo a restrição de longo prazo:

$$\sum_{j=0}^{\infty} a_{11,j} = 0 = a_{11,0} \sum_{j=0}^{\infty} c_{11,j} + a_{21,0} \sum_{j=0}^{\infty} c_{12,j}.$$

Logo, as equações a serem resolvidas para obter os coeficientes de A_0 são:

$$\begin{aligned} a_{11,0}^2 + a_{12,0}^2 &= \text{var}(e_1); \\ a_{11,0}a_{21,0} + a_{12,0}a_{22,0} &= \text{cov}(e_1, e_2); \\ a_{21,0}^2 + a_{22,0}^2 &= \text{var}(e_2); \\ a_{11,0} \sum_{j=0}^{\infty} c_{11,j} + a_{21,0} \sum_{j=0}^{\infty} c_{12,j} &= 0. \end{aligned}$$

DECOMPOSIÇÃO DE BLANCHARD E QUAH

- Uma vez obtida a matriz A_0 , podem-se recuperar os erros estruturais da seguinte relação:

$$\varepsilon_t = A_0^{-1} e_t,$$

e, assim, encontrar a função resposta ao impulso e a decomposição da variância.

- Não se devem associar choques de demanda ao componente transitório do ciclo de negócios, e choques de oferta, ao componente permanente de tendência. O componente transitório pode ocorrer devido a uma combinação de choques na oferta e demanda, assim como choques de oferta afetarão simultaneamente os componentes cíclico e permanente da economia.

ESTIMAÇÃO DO MODELO ESTRUTURAL*

Considere mais uma vez o modelo estrutural:

$$\mathbf{A}\mathbf{X}_t = \sum_{i=1}^p B_i \mathbf{X}_{t-i} + B\boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim i.i.d.N(\mathbf{0}, I_n).$$

Empilham-se as matrizes \mathbf{X}_i s, de modo que

$$\mathbf{Y}_t = \begin{bmatrix} \mathbf{X}'_{t-1} & \mathbf{X}'_{t-2} & \cdots & \mathbf{X}'_{t-p} \end{bmatrix}', \quad \mathbf{A} = A^{-1} [B_1, B_2, \dots, B_p]:$$

$$\mathbf{X}_t = \mathbf{A}\mathbf{Y}_t + A^{-1}B\boldsymbol{\varepsilon}_t,$$

em que B é a matriz de desvio padrão de dimensão $n \times n$.

A função de verossimilhança será dada por:

$$\begin{aligned} \ln(\mathbf{A}, A, B) &= -\frac{n(T-p)}{2} \ln 2\pi - \frac{(T-p)}{2} \ln |A^{-1}BB'A^{-1}| - \\ &\quad - \frac{1}{2} \sum_{t=p+1}^T (\mathbf{X}_t - \mathbf{A}\mathbf{Y}_t)' (A^{-1}BB'A^{-1})^{-1} (\mathbf{X}_t - \mathbf{A}\mathbf{Y}_t). \end{aligned}$$

ESTIMAÇÃO DO MODELO ESTRUTURAL*

$$\begin{aligned} & \sum_{t=p+1}^T (X_t - \mathbf{A}Y_t)' (A^{-1}BB'A^{-1'})^{-1} (X_t - \mathbf{A}Y_t) = \\ &= (T-p) \text{tr} \left[(X - \mathbf{A}Z)' (A^{-1}BB'A^{-1'})^{-1} (X - \mathbf{A}Z) \right] \end{aligned}$$

em que $X = [X'_{p+1}, X'_{p+2}, \dots, X'_T]'$ e $Z = [Y'_{p+1}, Y'_{p+2}, \dots, Y'_T]'$. Logo

$$\begin{aligned} \ln(\mathbf{A}, A, B) &= \text{constante} + \frac{(T-p)}{2} \left(\ln |A|^2 - \ln |B|^2 \right) - \\ &\quad - \frac{(T-p)}{2} \text{tr} \left[A' B'^{-1} B^{-1} A (X - \mathbf{A}Z)' (X - \mathbf{A}Z) \right]. \end{aligned}$$

ESTIMAÇÃO DO MODELO ESTRUTURAL*

Se não houver restrições sobre as matrizes A e B , pode-se estimar $\hat{\Sigma} = \frac{(X - \hat{A}Z)(X - \hat{A}Z)'}{T}$, em que $(X - \hat{A}Z)$ são os resíduos do VAR reduzido, $\hat{A} = XZ'(ZZ')^{-1}$. Daí, é possível concentrar a função anterior para maximizar:

$$\ln_c(A, B) = \text{constante} + \frac{(T-p)}{2} \left(\ln |A|^2 - \ln |B|^2 \right) - \frac{(T-p)}{2} \text{tr} \left[A' B'^{-1} B^{-1} A \hat{\Sigma} \right].$$

TESTE LR PARA SOBREIDENTIFICAÇÃO

- ① Estimada a matriz de covariância não restrita, $\hat{\Sigma}$, e como se sabe que $BB' = A\hat{\Sigma}A'$, reescrever $\hat{\Sigma}$:

$$\hat{\Sigma} = A^{-1}BB'A'^{-1}.$$

- ② Imponha as restrições sobre A e B e maximize :
 $-\frac{T}{2} \ln |A^{-1}BB'A'^{-1}| - \frac{1}{2} \sum_{t=1}^T \hat{e}_t' (A^{-1}BB'A'^{-1})^{-1} \hat{e}_t.$
- ③ Calcule a nova matriz de covariância restrita: $\hat{\Sigma}_r = \hat{A}^{-1}\hat{B}\hat{B}'\hat{A}'^{-1}.$
- ④ Proceda ao teste estatístico de razão de verossimilhança dado por:

$$LR = T \left(\ln |\hat{\Sigma}_r| - \ln |\hat{\Sigma}| \right) \xrightarrow{d} \chi^2_q,$$

em que q é o número de restrições acima de $\frac{n(n-1)}{2}$.

- ⑤ Se há dois conjuntos de restrições tais que $q_1 > q_2$, pode-se testar a hipótese de as restrições serem equivalentes pelo teste:

$$LR = T \left(\ln |\hat{\Sigma}_{r_1}| - \ln |\hat{\Sigma}_{r_2}| \right) \xrightarrow{d} \chi^2_{q_1 - q_2}.$$