Vetor Auto-Regressivo - VAR

Profa. Rosângela Ballini

Bibliografia:

- Enders, W. *Applied Econometrics*, 3a. Edição, Wiley, 2010. Cap. 5: 5-9.
- Bueno, R. L. S..*Econometria de Séries Temporais*, 2a. Edição, Cengage Learning, 2011. Cap. 6: 6.1-6.8.

Vetor Autoregressivo

Modelo VAR irrestrito ou padrão:

- Proposto por Sims (1980);
- Todas as variáveis são tratadas como sendo, a priori, endógenas;
- Busca responder qual a trajetória da série, dado um choque estrutural

VAR Irrestrito:

É um modelo auto-regressivo multivariado em que cada variável é expressa como função de suas defasagens e das defasagens das demais variáveis do modelo.

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:

$$AX_t = B_0 + \sum_{i=1}^{p} B_i X_{t-i} + B\varepsilon_t \tag{1}$$

sendo:

- A uma matriz $n \times n$ que define as restrições contemporâneas entre as variáveis que constituem o vetor X_t, $n \times 1$;
- B_0 um vetor de constantes $n \times 1$; B_i matrizes $n \times n$;
- B uma matriz diagonal $n \times n$ de desvios padrão;
- ε_t 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

O modelo (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

O modelo (1) é normalmente estimado na **forma reduzida**:

$$X_t = A^{-1}B_0 + \sum_{i=1}^{p} A^{-1}B_i X_{t-i} + A^{-1}B\varepsilon_t \qquad (2)$$

Ou,

$$X_t = \Phi_0 + \sum_{i=1}^{p} \Phi_i X_{t-i} + e_t \qquad (3)$$

em que $\Phi_i \equiv A^{-1}B_i$ para $i = 0, 1, \ldots, p$ e $e_t \equiv A^{-1}B\varepsilon_t$

Exemplo:

Vamos considerar um modelo bivariado:

$$y_t = b_{10} - a_{12} z_t + b_{11} y_{t-1} + b_{12} z_{t-1} + \sigma_y \varepsilon_{yt} \quad (4)$$

$$z_t = b_{20} - a_{21} y_t + b_{21} y_{t-1} + b_{22} z_{t-1} + \sigma_z \varepsilon_{zt} \quad (5)$$

sendo que:

1. y_t e z_t são estacionárias;
2. $\varepsilon_{yt} \sim RB(0, 1)$ e $\varepsilon_{zt} \sim RB(0, 1)$;
3. $\varepsilon_{yt} \perp \varepsilon_{zt} \Longrightarrow cov(\varepsilon_{yt}, \varepsilon_{zt}) = 0$.

As equações (4) e (5) constituem um VAR(1), sendo denominado de VAR estrutural.

Exemplo:

Representação matricial da forma estrutural:

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

Ou,

$$AX_t = B_0 + B_1 X_{t-1} + B\varepsilon_t$$

Forma reduzida do modelo:

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

Exemplo:

Ou seja, na forma reduzida:

$$y_t = \phi_{10} + \phi_{11} y_{t-1} + \phi_{12} z_{t-1} + e_{yt} \tag{6}$$

$$z_t = \phi_{20} + \phi_{21} y_{t-1} + \phi_{22} z_{t-1} + e_{zt} \tag{7}$$

em que:

$$E(e_{it}) = 0$$

$$Var(e_{it}) = 1$$

$$cov(e_{it}, e_{it-j}) = 0, j \neq 0$$

para $i = \{y, z\}$. Devemos notar que, em geral,

$$cov(e_{yt}, e_{zt}) \neq 0$$

Simulação

Gerar séries temporais a partir de simulações de modelos VAR na forma reduzida sem e com constante.

Modelo VAR(p)

Um modelo VAR(p) na forma reduzida pode ser representado por:

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

em que G é um matriz de coeficientes $n \times g$ e Z representa um vetor, $g \times 1$ de variáveis exógenas que pode incluir termos determinísticos.

Diferentemente dos modelos univariados, no modelo VAR analisar a trajetória da série, dado um **choque estrutural**

Por trajetória entende-se conhecer o tempo que um choque afeta uma série, se ela muda de patamar ou não, por exemplo.

Identificação

Modelo VAR(p):

$$X_t = \Phi_0 + \sum_{i=1}^{p} \Phi_i X_{t-i} + e_t \tag{8}$$

em que:

X_t: vetor $n \times 1$ contendo cada uma das n variáveis do modelo;

Φ_0: vetor $n \times 1$ de termos constantes;

Φ_i: matriz $n \times n$ dos coeficientes

e_t: vetor $n \times 1$ de termos de erros;

- $c = n + n^2 p$ coeficientes que devem ser estimados
- Muitas vezes os coeficientes estimados são estatisticamente insignificantes

Identificação

Questão:

Como selecionar a ordem p do modelo VAR?

Considere um VAR(p), $p = 0, 1, 2, \ldots, p_{max}$.

Problema:

Escolher a ordem p que minimiza a fórmula geral do critério de informação:

$$Cr(p) = \ln(\det(\hat{\Sigma})) + c_T\varphi(p)$$

em que $\hat{\Sigma} = T^{-1}\,(\hat{e}_t\hat{e}_t')$ é a matriz de covariância para um modelo de ordem p; c_T é uma sequência que depende do tamanho da amostra T; $\varphi(p)$ é uma função que penaliza VAR de grandes ordens.

Identificação – Critérios de Informação

$$AIC(p) = \ln(\det(\hat{\Sigma})) + \frac{2}{T}c \tag{9}$$

$$BIC(p) = \ln(\det(\hat{\Sigma})) + \frac{\ln T}{T}c \tag{10}$$

$$HQ(p) = \ln(\det(\hat{\Sigma})) + \frac{\ln \ln T}{T}2c \tag{11}$$

em que $\hat{\Sigma}$: matriz das variâncias covariâncias do modelo; $c = pn^2$: número total de parâmetros estimados em todas as equações do modelo; T: tamanho da amostra.

- Tamanho amostral deve ser mantido constante para tornar o critério comparável. Logo, o tamanho da amostra é $(T - p_{max})$.

Identificação – Final Prediction Error - FPE

O critério FPE fornece uma medida da qualidade do modelo por meio de ajustes de modelos com diferentes ordens. O modelo mais preciso tem o menor FPE de acordo com o critério:

$$FPE(p) = \left(\det(\hat{\Sigma})\right)\left(\frac{T + np + 1}{T - np - 1}\right)^{n} \tag{12}$$

- Se usarmos o mesmo conjunto de dados tanto para estimação como para validação do modelo, o ajuste sempre melhora à medida que aumenta a ordem do modelo.

Identificação – Teste de Razão de Verossimilhança

- Outra maneira de escolher a ordem de defasagem é aplicar testes sequenciais para definir a ordem p do modelo VAR
- Estabeleça o p_{max} e considere:

$$
\begin{aligned}
H_0 &: \quad \Phi_i = 0, \quad i = p_{max}, p_{max} - 1, \ldots, 1 \\
H_1 &: \quad \Phi_i \neq 0
\end{aligned}
$$

- Quando a hipótese nula for rejeitada, a ordem do modelo é encontrada.

- Dificuldade: estabelecer o p_{max}, pois se p_max for muito pequeno os resíduos estimados não serão um ruído branco. Se p_max muito grande, o impacto sobre o erro tipo I (rejeitar H_0) poderá afetar os resultados e os intervalos de confiança gerados não serão confiáveis.

Identificação – Teste de Razão de Verossimilhança

Considere o modelo irrestrito e estabeleça quais as restrições a serem impostas.

1. Estime o modelo irrestrito e calcule $\hat{\Sigma}_u$;
2. Estime o modelo restrito e calcule $\hat{\Sigma}_r$;
3. Calcule a razão de verossimilhança:

$$LR = (T - c)(\ln(\det(\hat{\Sigma}_r)) - \ln(\det(\hat{\Sigma}_u))) \rightarrow \chi^2_r \qquad (13)$$

em que T é o número de observações; c é o número de parâmetros estimado em cada equação; r é o número de restrições do sistema.

Estimação

- Mínimos Quadrados Ordinários aplicados a cada equação do modelo: número de variáveis em todas as equações deve ser igual

- Propriedades dos estimadores:
    1. Consistência;
    2. Eficiência assintótica;
    3. Normalidade assintótica.

Exemplo 6.11 Adaptado (Bueno,2011)

O Banco Central do Brasil (BC) utiliza em suas projeções para inflação quatro modelos VAR, descritos no Relatório de Inflação 2T04.

Com base no modelo VAR1, vamos verificar o impacto que variáveis como a variação da taxa de câmbio nominal e a taxa de juros Selic real têm sobre os preços livres.

Período: Setembro de 1994 a Maio de 2007.

Exemplo 6.11 (Bueno,2011)

Variáveis endógenas utilizadas:

1. Variação da taxa de câmbio nominal;
2. Variação da taxa Selic real;
3. Inflação dos preços livres;
4. Inflação dos preços administrados.

Variáveis exógenas utilizadas:

1. 11 dummies sazonais;
2. Uma dummy de tendência para o período de desinflação (1994:9 a 1997:7)

Verificação do Modelo

1. Estabilidade do modelo;
2. Correlograma;
3. Testes de Autocorrelação: Teste de Portmanteau;
4. Teste de Normalidade: ortogonalização da matriz $\hat{\Sigma}$.

Problema:

Diagnóstico de resíduos dos modelos

Estabilidade do modelo

Seja o modelo VAR(p) com n variáveis, na forma reduzida:

$$X_t = \Phi_0 + \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \ldots + \Phi_p X_{t-p} + e_t$$

Usando operadores de defasagens:

$$X_t = \Phi_0 + \left(\Phi_1 L + \Phi_2 L^2 + \ldots + \Phi_p L^p\right) X_t + e_t$$

Ou,

$$\left(\mathbf{I} - \Phi_1 L - \Phi_2 L^2 - \ldots - \Phi_p L^p\right) X_t = \Phi_0 + e_t$$

que pode ser escrito na forma resumida como:

$$\Phi(L) X_t = \Phi_0 + e_t$$

- Condição de estabilidade em modelos AR(p): $|\phi_i| < 1, i = 1, \ldots, p$
- Analogia com modelos VAR

Estabilidade do modelo

Supondo um modelo VAR(1):

$$X_t = \Phi_0 + \Phi_1 X_{t-1} + e_t$$

e resolvendo recursivamente:

$$X_t = \mu + \sum_{i=0}^{\infty} \Phi_1^i e_{t-i}$$

Condição de estabilidade $\Phi_1^n \rightarrow \infty$:

1. Os autovalores da matriz Φ_1 devem ser menores que 1 em valor absoluto;

Estabilidade do modelo

Condição de estabilidade impõe que as raízes de $\Phi(L) = 0$.

Fazendo $z = L$:

$$\mathbf{I} - \Phi_1 z - \Phi_2 z^2 - \ldots - \Phi_p z^p = 0$$

esta condição impõe que as raízes estejam **fora** do círculo unitário.

Ou, escrevendo $z = \dfrac{1}{\lambda}$, é o mesmo que dizer que as raízes de

$$\lambda^p - \Phi_1 \lambda^{p-1} - \Phi_2 \lambda^{p-2} - \ldots - \Phi_p = 0$$

estejam **dentro** do círculo unitário.

Estabilidade Estrutural da Regressão

Uma das forma de análise da estabilidade estrutural é a partir do Gráficos de Controle de Soma Acumulada (Cumulative Sum Control Charts - CUSUM).

O gráfico de controle CUSUM é uma ferramenta estatística que acumula informações das amostras de um processo ponderando-as igualmente, isto é, as amostras têm o mesmo peso.

No R, podemos usar a função `stability()`, a qual tem os seguintes argumentos:

```
stability(x, type = c(''OLS-CUSUM'', ''Rec-CUSUM'',
''Rec-MOSUM'',''OLS-MOSUM''))
```

Correlograma

- Cálculo da correlação cruzada entre pares:

$$\rho_{ij}(k) = \frac{cov(e_{it}, e_{jt-k})}{std(e_{it})std(e_{jt})}, i = 1, 2; j = 1, 2$$

em que $std(e_{it})$ representa o desvio padrão para $i = 1, 2$ e

$$\rho_k \to N\left(0, \frac{1}{T}\right)$$

- Tomando um intervalo com 95% de confiança em torno de zero, os limites inferiores e superiores são dados por:

$$\left[-\frac{2}{\sqrt{T}}, \frac{2}{\sqrt{T}}\right]$$

sendo T o número de observações.

Teste de Autocorrelação dos Resíduos

Teste de Portmanteau

Objetivo: Verificar se as autocorrelações multivariadas são nulas.

- Generalização dos testes de resíduos dos modelos univariados:

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

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

- A estatística é usada para testar que as autocorrelações e correlações cruzadas não são significativas, para um dado nível de significância.

Teste de Autocorrelação dos Resíduos

Estatística do teste (Teste de Box-Pierce):

$$Q = T\sum_{j=1}^{J} tr(\hat{\Sigma}_j'\hat{\Sigma}_0^{-1}\hat{\Sigma}_j'\hat{\Sigma}_0^{-1}) \rightarrow \chi^2_{n^2(J-p)}$$

Estatística do teste (Teste de Ljung-Box):

$$Adj\_Q = T^2\sum_{j=1}^{J} \frac{1}{T-j} tr(\hat{\Sigma}_j'\hat{\Sigma}_0^{-1}\hat{\Sigma}_j'\hat{\Sigma}_0^{-1}) \rightarrow \chi^2_{n^2(J-p)}$$

sendo $\hat{\Sigma}_j = \frac{\sum_{t=j+1}^{T} \hat{e}_t\hat{e}_{t-j}}{T}$ a autocovariância na defasagem j; n o número de variáveis no modelo; p a ordem do VAR; J a defasagem; e $tr(\cdot)$ é o traço da matriz.

Autocorrelação dos Resíduos – Teste LM

Objetivo:

Testar se existe autocorrelação de resíduos no modelo:

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

$$H_0 : \Theta_j = 0, \text{ para todo } j = 1, 2, \ldots, h$$

$$H_1 : \Theta_j \neq 0, \text{ para algum } j$$

Utiliza-se a regressão auxiliar:

$$\hat{e}_t = \Phi_1 x_{t-1} + \ldots + \Phi_p x_{t-p} + \Theta_1 \hat{e}_{t-1} + \ldots + \Theta_h \hat{e}_{t-h} + u_t$$

Autocorrelação dos Resíduos – Teste LM

O teste é executado em dois estágios:

1. O modelo completo é estimado por MQO, de forma que os $\hat{e}_t$ são substituídos por zero, para $t < 0$. Calcula-se:

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

2. Estima-se o modelo impondo H_0 para obter os resíduos $\hat{e}_t^r$ e calcula-se:

$$\hat{\Sigma}_r = \frac{\sum_{t=1}^{T} \hat{e}_t^r (\hat{e}_t^r)'}{T}$$

Estatística do teste LM:

$$LM_h = T\left[n - tr\left(\hat{\Sigma}_u \hat{\Sigma}_r^{-1}\right)\right] \rightarrow \chi^2_{hn^2}$$

Teste ARCH-LM - Heterocedasticidade

Objetivo:

Análise de heterocedasticidade condicional.

Regredir a equação:

$$\hat{\epsilon}_t^2 = \beta_1 \hat{\epsilon}_{t-1}^2 + \beta_2 \hat{\epsilon}_{t-2}^2 + \ldots + \beta_h \hat{\epsilon}_{t-h}^2 + u_t$$

Sob as hipóteses

$$H_0 : \beta_1 = \beta_2 = \ldots = \beta_h = 0$$

$$H_1 : \beta_1 \neq 0, \text{ou} \beta_2 \neq 0, \text{ou} \ldots \beta_h \neq 0$$

Estatística do teste:

$$ARCH - LM = T \cdot R^2 \xrightarrow{d} \chi_h^2$$

Não rejeitar H_0 significa ausência de heterocedasticidade.

Teste de Normalidade

- Teste de Jarque-Bera Multivariado
- Devemos escolher uma fatoração dos *n* resíduos que são ortogonais entre si.

Etapas do teste:

1. Obter a matriz de covariância dos resíduos

$$\hat{\Sigma}_e = \frac{\left(\hat{e}_t - \bar{\hat{e}}_t\right)\left(\hat{e}_t - \bar{\hat{e}}_t\right)'}{T}$$

Teste de Normalidade

2. Calcular a raiz quadrada da matriz: $\hat{\Sigma}_e^{1/2}$. Para tanto, são obtidos os autovalores $\lambda_1, \ldots, \lambda_n$ da matriz $\hat{\Sigma}_e$ e a correspondente matriz Q composta pelos autovetores ortonormais, tal que $\hat{\Sigma}_e = Q\Lambda Q'$ com $\Lambda = diag(\lambda_1, \ldots, \lambda_n)$. Assim,

$$\hat{\Sigma}_e^{1/2} = Q\ diag(\lambda_1^{1/2}, \ldots, \lambda_n^{1/2})\ Q'$$

Teste de Normalidade

O teste para não normalidade é baseado na assimetria e curtose dos resíduos padronizados $\hat{e}_t^s = (\hat{e}_{1t}^s, \ldots, \hat{e}_{nt}^s)'$:

$$\hat{e}_t^s = \hat{\Sigma}_e^{-1/2}(\hat{e}_t - \bar{\hat{e}})$$

Assimetria:

$$\hat{m}_3 = (\hat{m}_{31}, \hat{m}_{32}, \ldots, \hat{m}_{3n})' \quad \text{com} \quad \hat{m}_{3i} = \frac{\sum_{t=1}^{T} (e_{it}^s)^3}{T}$$

Curtose:

$$\hat{m}_4 = (\hat{m}_{41}, \hat{m}_{42}, \ldots, \hat{m}_{4n})' \quad \text{com} \quad \hat{m}_{4i} = \frac{\sum_{t=1}^{T} (e_{it}^s)^4}{T}$$

Teste de Normalidade

As estatísticas são dadas por:

$$s_3^2 = T\frac{\hat{m}_3'\hat{m}_3}{6} \quad s_4^2 = T\frac{(\hat{m}_4 - 3_n)'(m_4 \hat{-} 3_n)}{24}$$

sendo $3_n = (3, 3, \ldots, 3)'$ é um vetor $n \times 1$ de $3s$. Ambas as estatísticas tem distribuição χ_n^2 sob a hipótese nula $s_3^2 = s_4^2 = 0$..

Alternativamente, pode-se usar a a distribuição conjunta de ambos os testes:

$$JB_{2n} = s_3^2 + s_4^2 \to \chi_{2n}^2$$

Teste de Normalidade

- A hipótese nula do teste é de que os resíduos são normalmente distribuídos;
- Devemos ressaltar que a rejeição da hipótese nula não impede a interpretação e análise dos resultados, apenas de sugerir cautela;
- A não normalidade dos resíduos em análises de séries macroeconômicas é comum nos estudos que realizam o teste de Jarque-Bera