

# Macroeconometria II

## Mestrado em Econometria Aplicada e Previsão

Jorge Caiado

CEMAPRE/ISEG, Universidade de Lisboa

Email: [jcaiado@iseg.utl.pt](mailto:jcaiado@iseg.utl.pt)

Web: <http://pascal.iseg.utl.pt/~jcaiado/>



# INTRODUÇÃO AOS MODELOS VAR

## Conceitos e definições

Limitações de outros modelos:

- Modelos de equações simultâneas
- Modelos de função transferência de Box-Jenkins

Modelos VAR:

- Christopher A. Sims – Prémio Nobel da Economia em 2011, “Macroeconomics and Reality” (*Econometrica* 48, 1-49, 1980)
- Permitem avaliar o impacto dinâmico das perturbações aleatórias sobre o sistema de variáveis.
- Trata todas as variáveis simetricamente sem impor qualquer restrição quanto à independência e dependência entre elas

# INTRODUÇÃO AOS MODELOS VAR

## Conceitos e definições

- A expressão matemática do modelo VAR de ordem  $p$  ou, simplesmente, VAR( $p$ ) é dada por:

$$Y_t = A_0 + A_1 Y_{t-1} + \dots + A_p Y_{t-p} + \varepsilon_t,$$

$Y_t = (Y_{1t}, \dots, Y_{kt})'$  é um vector de  $k$  variáveis endógenas,

$A_0$  é um vector de termos independentes,

$A_1, \dots, A_p$  são matrizes de coeficientes e

$\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{kt})'$  é um vector de perturbações aleatórias não correlacionadas com os seus próprios valores passados e não correlacionadas com nenhuma das variáveis do segundo membro, embora possam estar contemporaneamente correlacionadas, e com matriz de covariâncias  $\Omega$  não singular.

# MODELOS VAR

## Condições de estabilidade/estacionaridade

Condições de estabilidade/estacionaridade:

Considere-se um modelo VAR(1) bivariado na sua forma matricial,

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix},$$

ou, de forma abreviada,

$$Y_t = A_0 + A_1 Y_{t-1} + \varepsilon_t,$$

# MODELOS VAR

## Condições de estabilidade/estacionaridade

Resolvendo o sistema através de sucessivas iterações, obtém-se

$$Y_t = A_0 + A_1(A_0 + A_1 Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t$$

$$Y_t = A_0 + A_1 A_0 + A_1^2 Y_{t-2} + A_1 \varepsilon_{t-1} + \varepsilon_t$$

$$Y_t = (I + A_1)A_0 + A_1^2 Y_{t-2} + A_1 \varepsilon_{t-1} + \varepsilon_t$$

...

$$Y_t = (I + A_1 + \dots + A_1^n)A_0 + A_1^{n+1}Y_{t-n-1} + \sum_{i=0}^n A_1^i \varepsilon_{t-i}$$

Pode verificar-se que  $A_1^{n+1}$  converge para zero à medida que  $n \rightarrow \infty$ . Se todos os valores próprios da matriz  $A_1$  tiverem módulo inferior a 1, então

$$(I + A_1 + \dots + A_1^n)A_0 \xrightarrow{n \rightarrow \infty} (I - A_1)^{-1}A_0$$

*Demonstração:* Ver Lütkepohl (2005)

# MODELOS VAR

## Condições de estabilidade/estacionaridade

Deste modo, pode apresentar-se a solução de  $Y_t$  através de um processo estocástico da forma

$$Y_t = \mu + \sum_{i=0}^{\infty} A_1^i \varepsilon_{t-i}$$

O processo VAR(1) diz-se então um processo estável se todos os valores próprios da matriz  $A_1$  têm módulo inferior a 1. Esta condição é equivalente a pedir que as raízes da equação

$$\det(I - A_1 L) = 0, \text{ para } |L| \leq 1.$$

se encontrem fora do círculo unitário. Para processos VAR( $p$ ) de ordem  $p > 1$ , a condição de estabilidade é as raízes da equação

$$\det(I - A_1 L - \dots - A_p L^p) = 0.$$

se encontrem fora do círculo unitário.

# MODELOS VAR

## Condições de estabilidade/estacionaridade

**Exemplo:** Lütkepohl (2005). Considere-se um processo VAR(1) tri-dimensional:

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \\ Y_{3t} \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \\ a_{30} \end{bmatrix} + \begin{bmatrix} 0.5 & 0 & 0 \\ 0.1 & 0.1 & 0.3 \\ 0 & 0.2 & 0.3 \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \\ Y_{3,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}$$

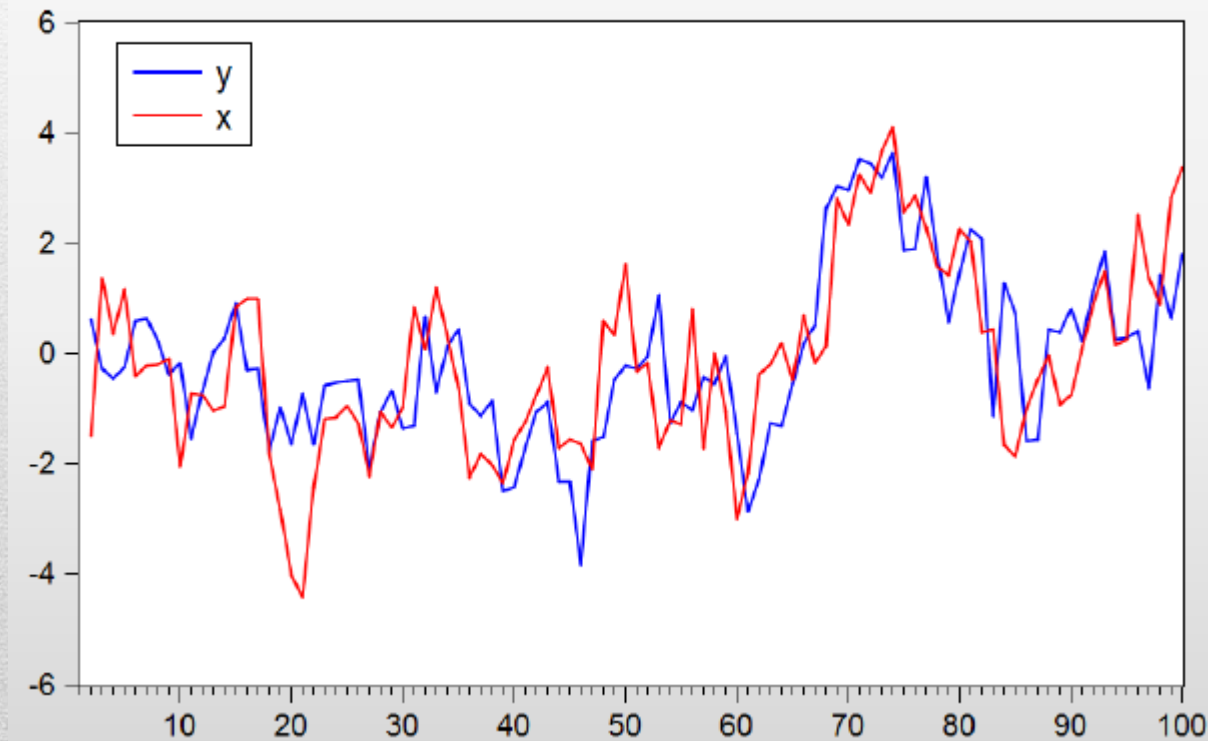
Calculemos as raízes do polinómio  $\det(I - A_1L)$ .

# MODELOS VAR

## Condições de estabilidade/estacionaridade

$$a_{10} = a_{20} = 0, a_{11} = a_{22} = 0.6 \text{ e } a_{12} = a_{21} = 0.3$$

(a) Modelo VAR(1) estavel



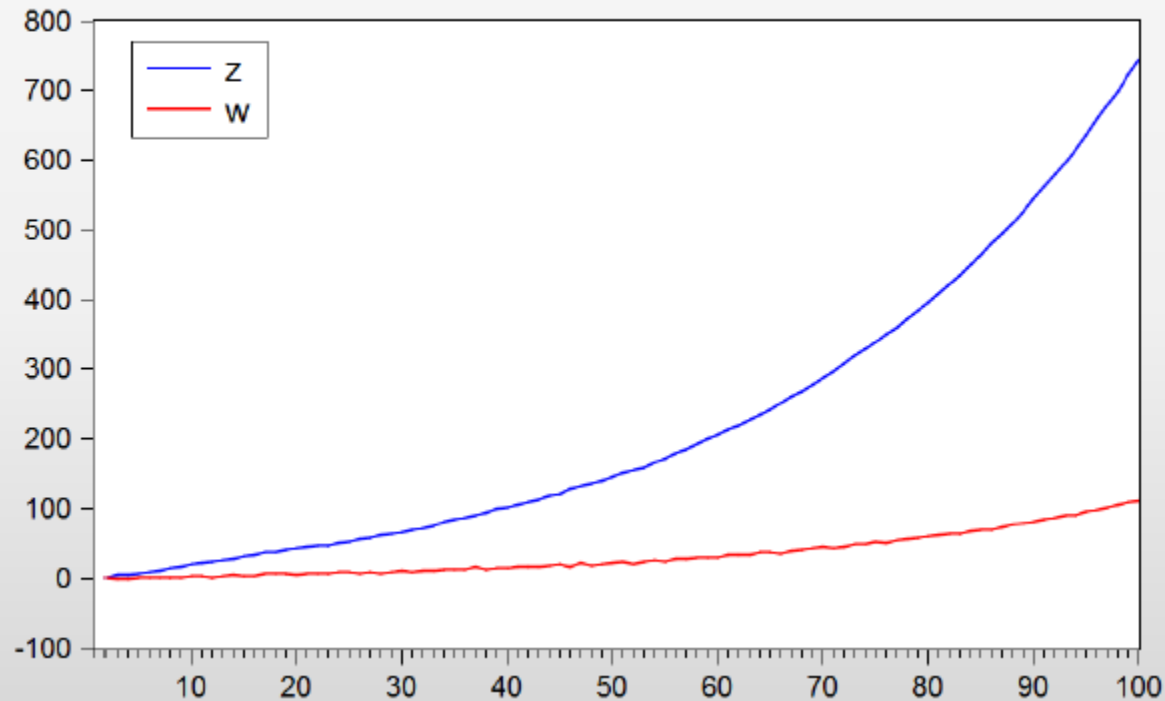


# MODELOS VAR

## Condições de estabilidade/estacionaridade

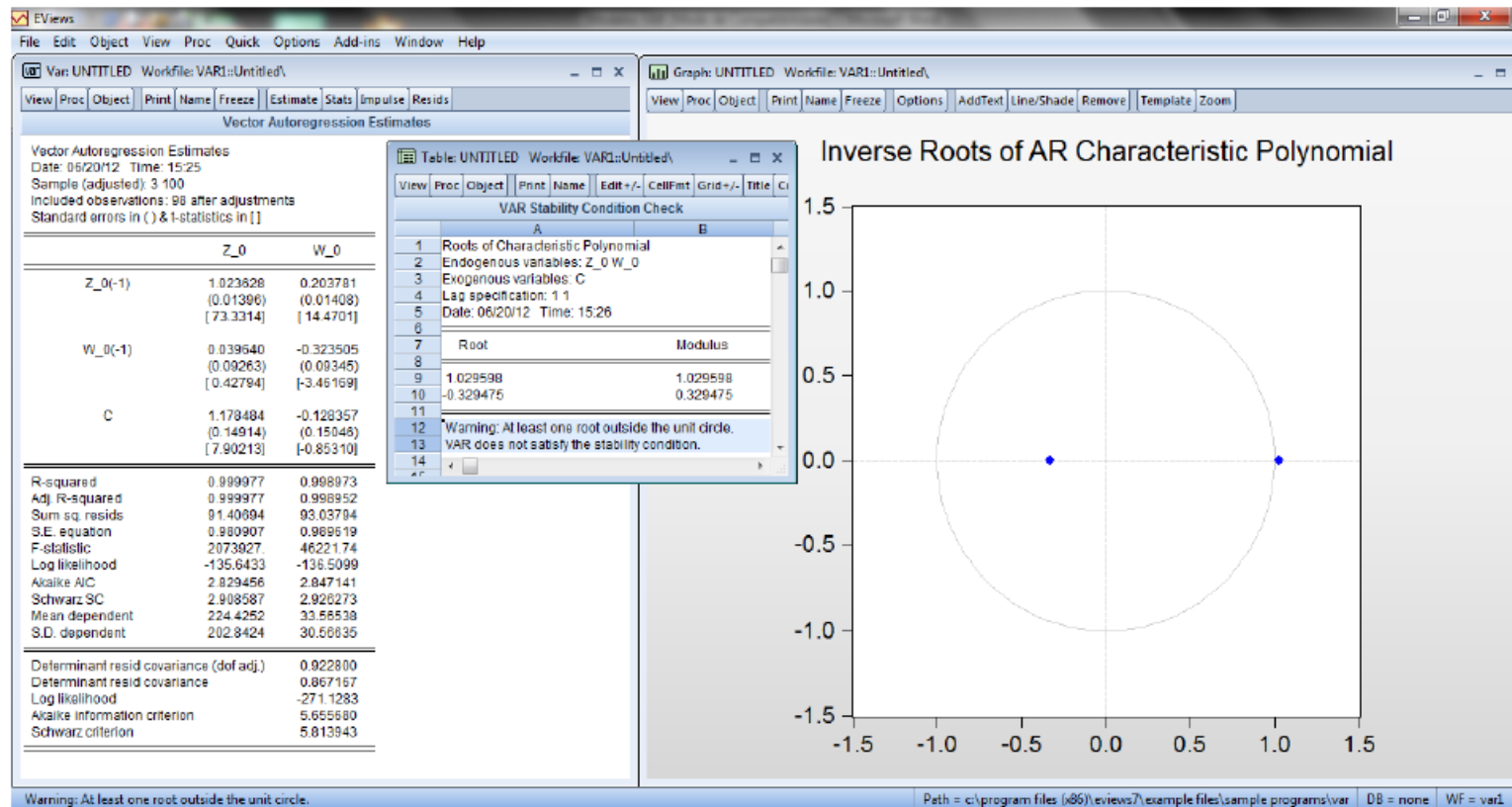
$$a_{10} = 1, a_{20} = 0, a_{11} = 1, a_{22} = -0.3, a_{12} = 0.2, a_{21} = 0.2$$

(b) Modelo VAR(1) instavel



# MODELOS VAR

## Condições de estabilidade/estacionaridade



Estimação do VAR(1) instável simulado e raízes do polinómio

# MODELOS VAR

## Representação médias móveis

Sob as condições de invertibilidade do VAR, o processo  $Y_t$  tem a seguinte representação médias móveis (MA),

$$\begin{aligned} Y_t &= \mu + (I - A_1L - \dots - A_pL^p)^{-1} \varepsilon_t = \mu + (I + \Psi_1L + \Psi_2L^2 + \dots)\varepsilon_t, \\ &= \mu + \sum_{i=0}^{\infty} \Psi_i \varepsilon_{t-i}, \end{aligned}$$

Os coeficientes  $\psi$  na representação médias móveis são obtidos através de um processo recursivo a partir da relação de identidade,

$$(I - A_1L - \dots - A_pL^p)(I + \Psi_1L + \Psi_2L^2 + \dots) = I.$$

**Exemplo:** Retome o exemplo do processo VAR(1) tri-dimensional (Lütkepohl, 2005) e calcule os pesos MA

# MODELOS VAR

## Estimação

Problemas que se colocam na estimação (para além da escolha das variáveis e da ordem de defasamento):

- Elevado número de parâmetros (sobreparametrização) do modelo ( $k=n+pn^2$  parâmetros)
- Colinearidade dos regressores (que pode invalidar os testes-t)

Como os segundos membros das equações do sistema VAR contêm apenas os valores defasados das variáveis endógenas e os erros  $\varepsilon_t$  são não correlacionados com os seus valores passados, cada equação do sistema pode ser estimada pelo método dos mínimos quadrados (OLS).

Um outro aspecto importante na estimação do modelo VAR prende-se com a questão das variáveis serem ou não estacionárias. Alguns autores, como Sims (1980) e Sims, Stock e Watson (1990), recomendam a estimação do VAR em níveis mesmo que as variáveis contenham uma raiz unitária.

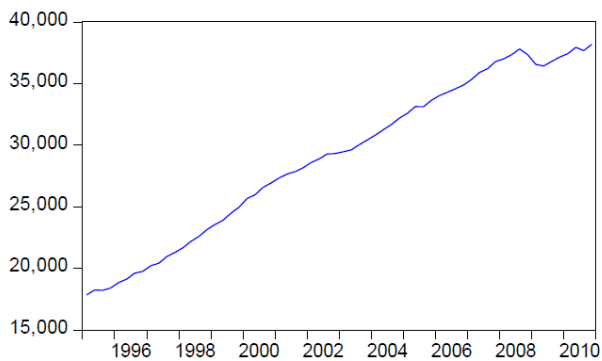
# MODELOS VAR

## Estimação

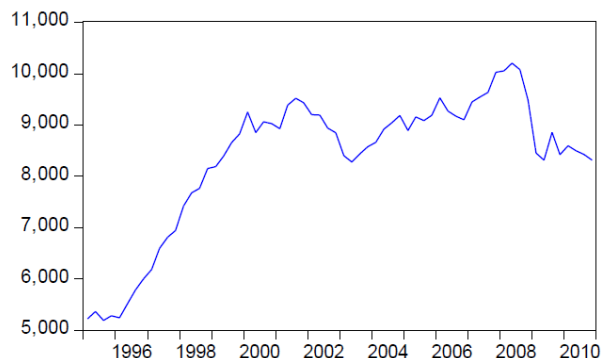
### Exemplo de aplicação:

Para ilustrar a estimação do modelo VAR, foram recolhidos os dados das contas nacionais trimestrais das variáveis Consumo (CONS), Investimento (INVES) e Rendimento Disponível (REND), no período de 1995:1ºT a 2010:4ºT (64 observações trimestrais).

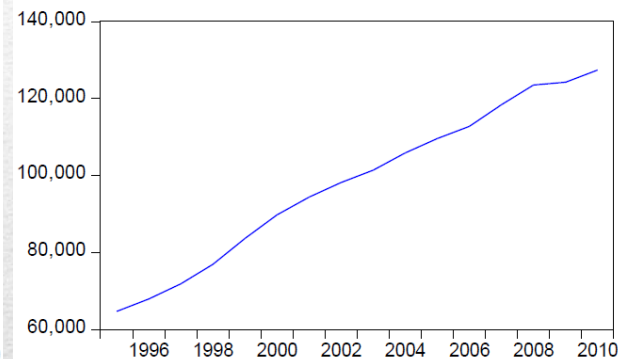
Consumo



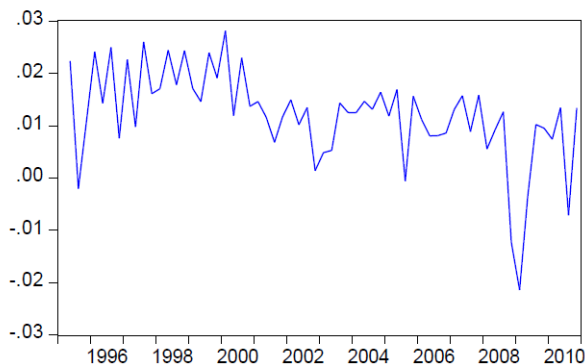
Investimento



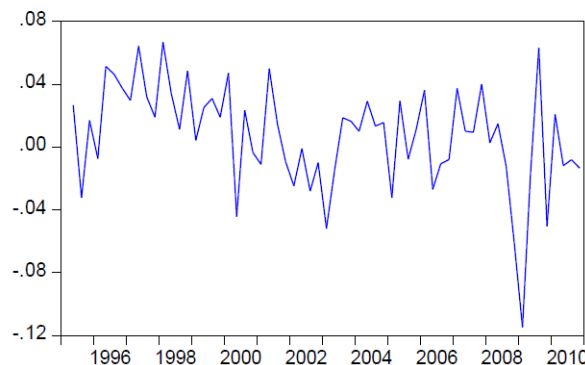
rendimento disponível



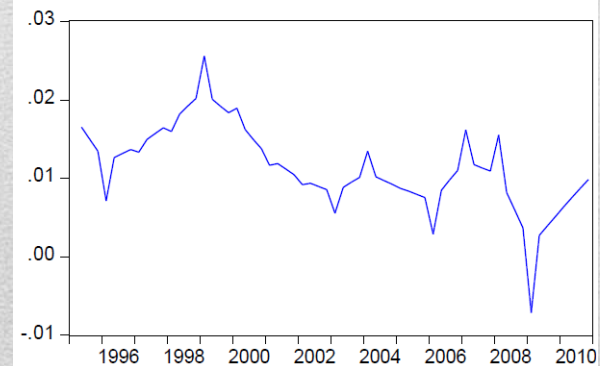
Log Differenced Consumo



Log Differenced Investimento



Log Differenced rendimento disponível



# MODELOS VAR

## Estimação

No output EViews, podemos retirar a seguinte informação:

- Estimativas, desvios padrão e rácios-*t* (para cada uma das equações do VAR);
- Coeficiente de determinação e coeficiente de determinação ajustado;
- Soma de quadrados dos resíduos;
- Erro padrão da regressão;
- Valor observado da estatística F;
- Critérios de informação AIC e BIC (selecção de modelos);
- Média da variável dependente e respectivo desvio padrão;
- Determinante da matriz de variâncias e covariâncias dos resíduos (ajustado dos graus de liberdade);
- Determinante da matriz de variâncias e covariâncias dos resíduos;
- Função de verosimilhança (assumindo uma distribuição normal multivariada);
- Critérios de informação AIC e BIC (selecção da ordem do VAR),

Vector Autoregression Estimates			
Date: 06/21/12 Time: 15:44			
Sample (adjusted): 1995Q4 2010Q4			
Included observations: 61 after adjustments			
Standard errors in ( ) & t-statistics in [ ]			
	DLOG(CONS)	DLOG(INVEST)	DLOG(REND)
DLOG(CONS(-1))	-0.111180 (0.16078) [-0.69151]	0.639770 (0.71060) [ 0.90033]	0.108127 (0.06900) [ 1.56699]
DLOG(CONS(-2))	0.042851 (0.16799) [ 0.25508]	-0.517545 (0.74248) [-0.69705]	0.007353 (0.07210) [ 0.10199]
DLOG(INVEST(-1))	0.070000 (0.03795) [ 1.84445]	-0.068725 (0.16773) [-0.40972]	0.006717 (0.01629) [ 0.41241]
DLOG(INVEST(-2))	-0.024471 (0.03862) [-0.63370]	0.056318 (0.17067) [ 0.32998]	-0.001293 (0.01657) [-0.07801]
DLOG(REND(-1))	0.786827 (0.35408) [ 2.22220]	3.513150 (1.56490) [ 2.24497]	0.552204 (0.15196) [ 3.63386]
DLOG(REND(-2))	0.151804 (0.31765) [ 0.47790]	-1.437060 (1.40390) [-1.02362]	0.149534 (0.13633) [ 1.09688]
C	0.001997 (0.00277) [ 0.72163]	-0.016686 (0.01223) [-1.36432]	0.001848 (0.00119) [ 1.55571]
R-squared	0.411149	0.191011	0.704329
Adj. R-squared	0.345722	0.101123	0.671477
Sum sq. resids	0.002722	0.053166	0.000501
S.E. equation	0.007100	0.031378	0.003047
F-statistic	6.284013	2.124999	21.43925
Log likelihood	218.9735	128.3235	270.5729
Akaike AIC	-6.949951	-3.977819	-8.641733
Schwarz SC	-6.707720	-3.735587	-8.399502
Mean dependent	0.012142	0.007706	0.011203
S.D. dependent	0.008777	0.033096	0.005316
Determinant resid covariance (dof adj.)	2.33E-13		
Determinant resid covariance	1.62E-13		
Log likelihood	638.6107		
Akaike information criterion	-20.24953		
Schwarz criterion	-19.52284		

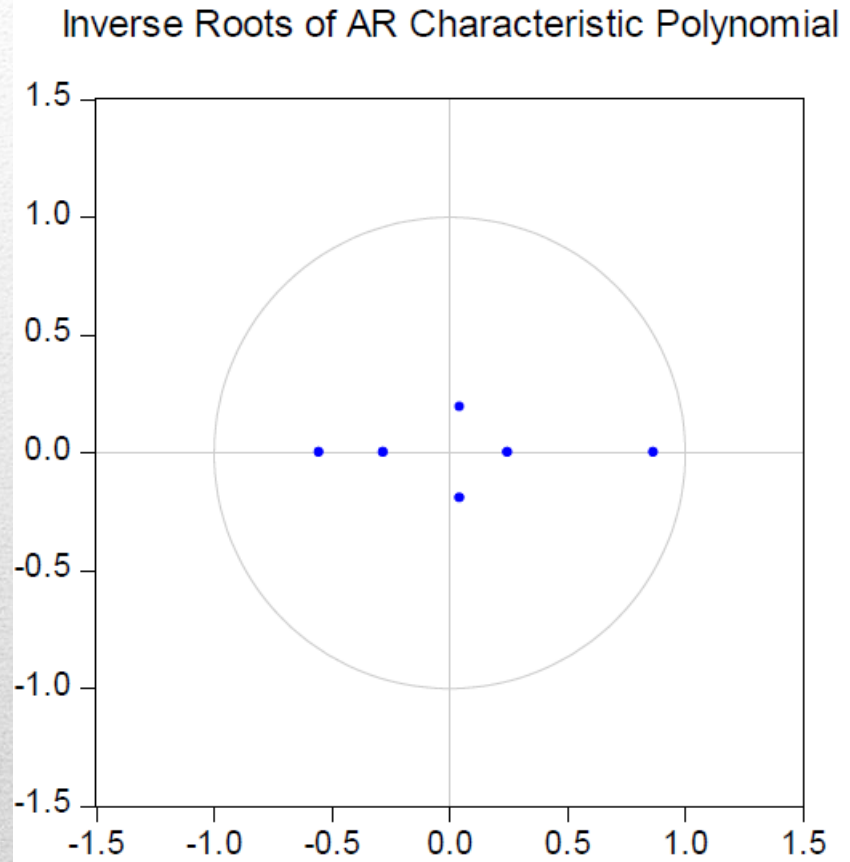
# MODELOS VAR

## Estimação

Roots of Characteristic Polynomial  
Endogenous variables: DLOG(CONS) DLOG(INVEST)  
DLOG(REND)  
Exogenous variables: C  
Lag specification: 1 2  
Date: 06/21/12 Time: 15:45

Root	Modulus
0.867983	0.867983
-0.551871	0.551871
-0.279844	0.279844
0.247912	0.247912
0.044060 - 0.194002i	0.198943
0.044060 + 0.194002i	0.198943

No root lies outside the unit circle.  
VAR satisfies the stability condition.



# MODELOS VAR

## Previsão

Suponha que estimou um modelo VAR( $p$ ) bivariado da forma

$$Y_t = A_0 + A_1 Y_{t-1} + \dots + A_p Y_{t-p} + \varepsilon_t, \text{ com } Y_t = (Y_{1t}, Y_{2t})'$$

e pretende prever os valores futuros das variáveis  $Y_{1t}$  e  $Y_{2t}$ .

Supondo que  $\Omega_t = \{Y_s, s \leq t\}$  contem toda a informação do sistema até ao momento  $t$  (presente e passada), e seja  $Y_t(h)$  o preditor a  $h$  passos (onde  $t$  é a origem da previsão e  $h$  o horizonte da previsão). Então, o preditor com menor **erro quadrático médio** (EQM) de previsão com origem em  $t$  e horizonte temporal de  $h$  períodos é o valor esperado condicional

$$E_t(Y_{t+h}) = E(Y_{t+h} | \Omega_t) = E(Y_{t+h} | \{Y_s, s \leq t\}).$$

O EQM do preditor  $Y_t(h)$  é definido por

$$EQM[Y_t(h)] = E[(Y_{t+h} - Y_t(h))(Y_{t+h} - Y_t(h))'].$$



# MODELOS VAR

## Previsão

Suponha que estimou um modelo VAR( $p$ ) bivariado da forma

$$Y_t = A_0 + A_1 Y_{t-1} + \dots + A_p Y_{t-p} + \varepsilon_t, \text{ com } Y_t = (Y_{1t}, Y_{2t})'$$

e pretende prever os valores futuros das variáveis  $Y_{1t}$  e  $Y_{2t}$ .

Supondo que  $\Omega_t = \{Y_s, s \leq t\}$  contem toda a informação do sistema até ao momento  $t$  (presente e passada), e seja  $Y_t(h)$  o preditor a  $h$  passos (onde  $t$  é a origem da previsão e  $h$  o horizonte da previsão). Então, o preditor com menor **erro quadrático médio** (EQM) de previsão com origem em  $t$  e horizonte temporal de  $h$  períodos é o valor esperado condicional

$$E_t(Y_{t+h}) = E(Y_{t+h} | \Omega_t) = E(Y_{t+h} | \{Y_s, s \leq t\}).$$

O EQM do preditor  $Y_t(h)$  é definido por

$$EQM[Y_t(h)] = E[(Y_{t+h} - Y_t(h))(Y_{t+h} - Y_t(h))'].$$

# MODELOS VAR

## Previsão

A optimalidade do valor esperado condicionado implica que

$$E_t(Y_{t+h}) = A_0 + A_1 E_t(Y_{t+h-1}) + \dots + A_p E_t(Y_{t+h-p}),$$

onde  $\varepsilon_t$  é assumido como sendo um ruído branco independente,  $E_t(\varepsilon_{t+h}) = 0$ .

Se  $\varepsilon_t$  e  $\varepsilon_s$  não forem independentes mas apenas não correlacionados,  $E_t(\varepsilon_{t+h})$ , em geral, não é igual a zero (para mais detalhes, veja-se Lutkepohl, 2005).

Para  $h = 1$ , tem-se:

$$E_t(Y_{t+1}) = A_0 + A_1 Y_t + \dots + A_p Y_{t-p+1}$$

Para  $h = 2$ , tem-se:

$$E_t(Y_{t+2}) = A_0 + A_1 E_t(Y_{t+1}) + A_2 Y_t + \dots + A_p Y_{t-p+2}$$

Procedendo assim recursivamente, obtém-se para um modelo VAR(1)

$$E_t(Y_{t+h}) = (I + A_1 + \dots + A_1^{h-1})A_0 + A_1^h Y_t$$

# MODELOS VAR

## Previsão

O valor esperado condicional tem as seguintes propriedades:

$$(i) \quad E[Y_{t+h} - E_t(Y_{t+h})] = 0$$

$$(ii) \quad EQM[E_t(Y_{t+h})] = EQM[E_t(Y_{t+h}) | Y_t, Y_{t-1}, \dots]$$

**Exemplo:** Retomando o modelo VAR(1) tri-dimensional (Lütkepohl, 2005):

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \\ Y_{3t} \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \\ a_{30} \end{bmatrix} + \begin{bmatrix} 0.5 & 0 & 0 \\ 0.1 & 0.1 & 0.3 \\ 0 & 0.2 & 0.3 \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \\ Y_{3,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}$$

Assumido que  $Y_t = (Y_{1t}, Y_{2t}, Y_{3t})' = (-6, 3, 5)'$  e  $A_0 = (a_{10}, a_{20}, a_{30})' = (0, 2, 1)'$ , obtenha previsões a 1 e a 2 passos para este processo VAR.

# MODELOS VAR

## Previsão

A determinação dos **intervalos de previsão** assenta nos seguintes pressupostos:

- Os erros  $\varepsilon_t$  têm distribuição normal multivariada,  $\varepsilon_t \sim N(0, \Sigma_\varepsilon)$
- Os termos  $\varepsilon_t$  e  $\varepsilon_s$  são independente para  $s \neq t$ .

Sob estas condições, os erros de previsão são normalmente distribuídos como transformações lineares de vectores normais,

$$Y_{t+h} - Y_t(h) = \sum_{i=0}^{h-1} \Psi_i \varepsilon_{t+h-i} \sim N(0, \Sigma_Y(h)),$$

Daqui sai a distribuição dos erros de previsão:

$$\frac{Y_{j,t+h} - Y_{j,t}(h)}{\sigma_j(h)} \sim N(0,1),$$

# MODELOS VAR

## Previsão

E assim, tem-se:

$$P\{Y_{j,t}(h) - z_{\alpha/2}\sigma_j(h) \leq Y_{j,t+h} \leq Y_{j,t}(h) + z_{\alpha/2}\sigma_j(h)\} = 1 - \alpha$$

E sai o seguinte intervalo de previsão a  $100(1 - \alpha)\%$ :

$$[Y_{j,t}(h) - z_{\alpha/2}\sigma_j(h), Y_{j,t}(h) + z_{\alpha/2}\sigma_j(h)]$$

**Exemplo:** Retomando os dados do exemplo anterior, e assumindo que:

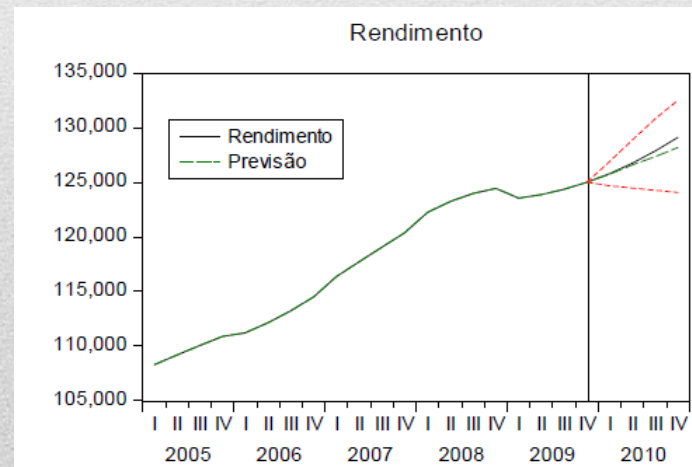
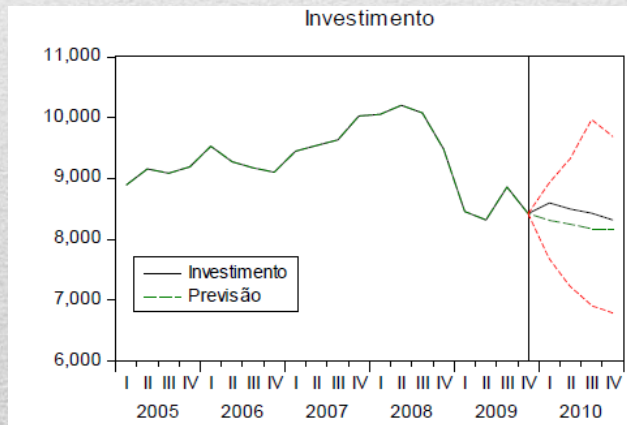
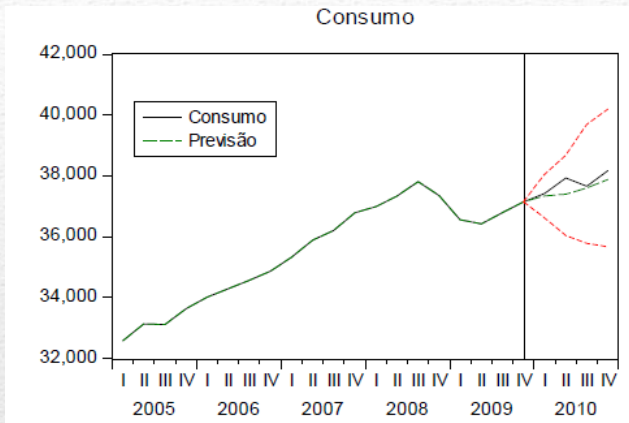
$$\Sigma_{\varepsilon} = \begin{bmatrix} 2.25 & 0 & 0 \\ 0 & 1.0 & 0.5 \\ 0 & 0.5 & 0.74 \end{bmatrix}$$

# MODELOS VAR

## Previsão

### Exemplo de aplicação:

Para exemplificar a aplicação da metodologia VAR em previsão, utilizaram-se mais uma vez os dados das variáveis macroeconómicas Consumo (CONS), Investimento (INVES) e Rendimento Disponível (REND). Para avaliar a qualidade preditiva, truncaram-se as últimas 4 observações trimestrais de 2010:1ºT a 2010:4ºT (“post-sample period”) e utilizaram-se as observações de 1995:1ºT a 2009:4ºT (“training sample”) para o ajustamento do modelo VAR(2) em níveis e em diferenças de logaritmo.



# MODELOS VAR

## Previsão

Erros de previsão (REQM e EPAM) dos modelos VAR(2) estimados

Erros de previsão	Consumo	Investimento	Rendimento
VAR(2) em níveis			
REQM	<b>214.220</b>	246.452	1012.501
EPAM	<b>0.48%</b>	<b>2.29%</b>	0.67%
VAR(2) em diferenças			
REQM	308.151	<b>241.802</b>	<b>540.431</b>
EPAM	0.64%	2.86%	<b>0.33%</b>

**Exercício:** Compare estas previsões com as previsões obtidas com base no modelo ARIMA(2,1,0) univariado de Box-Jenkins.

# MODELOS VAR

## Causalidade à Granger

A definição de causalidade entre duas variáveis foi inicialmente introduzida por Granger (1969) – Pémio Nobel da Economia em 2003, tendo posteriormente sido popularizada por Sims (1972). Granger (1969) introduziu o conceito de causalidade, tendo por base duas condições fundamentais:

- (i) O futuro não pode causar (ou prever) o passado, ou seja, a causalidade estrita pode ocorrer apenas com os valores passados a causar os valores presentes ou futuros;
- (ii) O conceito de causalidade deve ser definido apenas para processos estocásticos (não sendo relevante a sua análise para processos determinísticos)

A ideia base do conceito de **causalidade à Granger**, que a seguir se apresenta, é a de que  $X_t$  causa  $Y_t$  se a informação passada da variável  $X_t$  permite melhorar as previsões da variável  $Y_t$ , isto é, se  $Y_t$  for melhor previsto com base nos valores passados de  $X_t$  e  $Y_t$  juntos do que apenas com os valores passados de  $Y_t$ .



# MODELOS VAR

## Causalidade à Granger

Formalizando, supondo que  $\Omega_t$  contém toda a informação do universo até ao momento  $t$  (presente e passada) e  $\Omega_{t-1}$  toda a informação do universo até ao momento  $t-1$  (passada),  $X$  contém toda a informação presente e passada da variável  $X_t$ , isto é,  $X = \{X_s, s \leq t\}$  e  $X^*$  apenas a informação passada de  $X_t$ , isto é,  $X^* = \{X_s, s < t\}$ , assim como  $Y = \{Y_s, s \leq t\}$  e  $Y^* = \{Y_s, s < t\}$ , e sendo  $Y_t(1)$  e  $X_t(1)$  os preditores a um passo das séries  $Y_t$  e  $X_t$ , respectivamente.

(a) **Causalidade.** Diz-se que  $X_t$  causa  $Y_t$  se

$$EQM(Y_t(1)|\Omega_{t-1}) < EQM(Y_t(1)|\Omega_{t-1} - X^*)$$

(b) **Causalidade instantânea.** Diz-se que  $X_t$  causa instantaneamente  $Y_t$  se

$$EQM(Y_t(1)|\Omega_t) < EQM(Y_t(1)|\Omega_t - X)$$

# MODELOS VAR

## Causalidade à Granger

(c) **Feedback.** Diz-se que existe causalidade bidireccional (ou feedback) entre  $X_t$  e  $Y_t$  se

$$EQM(Y_t(1)|\Omega_{t-1}) < EQM(Y_t(1)|\Omega_{t-1} - X^*)$$

$$EQM(X_t(1)|\Omega_{t-1}) < EQM(X_t(1)|\Omega_{t-1} - Y^*)$$

(d) **Feedback instantâneo.** Diz-se que existe causalidade bidireccional instantânea (ou feedback instantâneo) entre  $X_t$  e  $Y_t$  se

$$EQM(Y_t(1)|\Omega_t) < EQM(Y_t(1)|\Omega_t - X)$$

$$EQM(X_t(1)|\Omega_t) < EQM(X_t(1)|\Omega_t - Y)$$

# MODELOS VAR

## Causalidade à Granger

### Testes de causalidade bivariada

Considere-se um processo VAR( $p$ ) bivariado estacionário da forma,

$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11,1} & a_{12,1} \\ a_{21,1} & a_{22,1} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \dots + \begin{bmatrix} a_{11,p} & a_{12,p} \\ a_{21,p} & a_{22,p} \end{bmatrix} \begin{bmatrix} X_{t-p} \\ Y_{t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}, \quad (1)$$

onde  $(\varepsilon_{1t}, \varepsilon_{2t})$  é um processo ruído branco bivariado com valor esperado igual a zero e matriz de variâncias e covariâncias:

$$\Sigma = \begin{bmatrix} \text{Var}(\varepsilon_{1t}) & \text{Cov}(\varepsilon_{1t}, \varepsilon_{2t}) \\ \text{Cov}(\varepsilon_{2t}, \varepsilon_{1t}) & \text{Var}(\varepsilon_{2t}) \end{bmatrix}.$$

No sentido da definição de causalidade à Granger, pode dizer-se que  $Y_t$  não causa  $X_t$  se na primeira equação do sistema (1) não se encontrarem valores desfasados da série  $Y_t$ ,

$$a_{12,1} = a_{12,2} = \dots = a_{12,p} = 0. \quad (2)$$

Da mesma forma, diz-se que  $X_t$  não causa  $Y_t$  se na segunda equação do processo (1) não aparecem valores desfasados da série  $X_t$ ,

$$a_{21,1} = a_{21,2} = \dots = a_{21,p} = 0. \quad (3)$$

# MODELOS VAR

## Causalidade à Granger

- (i) **Teste directo.** Consiste em estimar pelo método dos mínimos quadrados a primeira equação do sistema (1) sem e com as restrições impostas em (2), respectivamente, e testar a nulidade conjunta dos parâmetros associados aos valores desfasados da variável  $Y_t$  da relação sem restrições, através de um teste- $F$ , usando a estatística

$$F = \frac{(SQR_r - SQR)/p}{SQR/(T - 2p - 1)}, \quad (4)$$

que segue uma distribuição  $F$  com  $p$  e  $T - 2p - 1$  graus de liberdade.

- (ii) **Teste Sims.** Consiste na regressão  $Y_t$  sobre os valores passados, presentes e futuros de  $X_t$ ,

$$Y_t = c + \sum_{i=1}^m h_i Y_{t-i} + \sum_{j=0}^n b_j X_{t-j} + \sum_{l=1}^k d_l X_{t+l} + \eta_t, \quad (5)$$

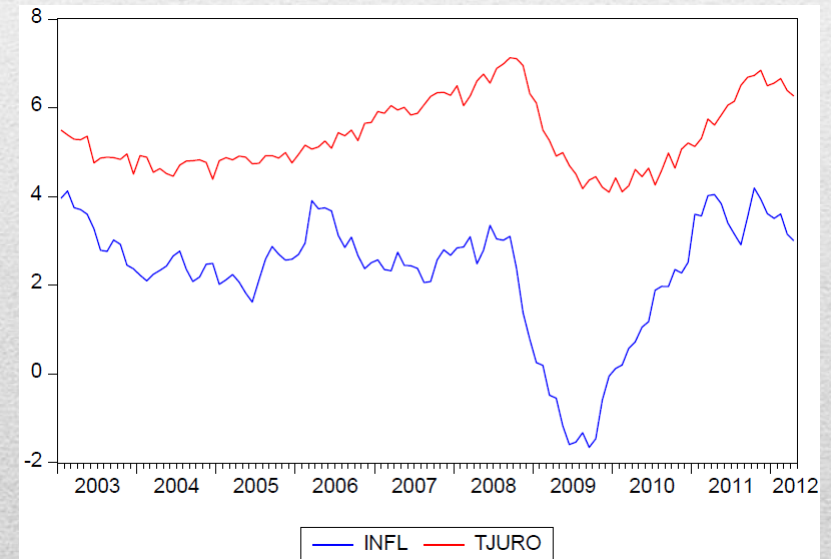
onde os erros  $\eta_t$  são não correlacionados com  $X_s$  para todo o  $t$  e  $s$ . A hipótese de que  $Y_t$  não causa  $X_t$  pode ser testada através de um teste  $F$  à nulidade conjunta dos coeficientes associados aos valores futuros da variável  $X_t$ , isto é, à hipótese nula  $H_0 : d_1 = d_2 = \dots = d_k = 0$ .

# MODELOS VAR

## Causalidade à Granger

(iii) **Teste de GMD** (Geweke, Meese e Dent, 1983) propuseram um método não paramétrico para a estimação da regressão (5), assumindo que os erros  $\eta_t$  podem ser caracterizados por um processo autoregressivo da forma  $\eta_t = \phi^{-1}(L)v_t$ .

**Exemplo de aplicação:** Para ilustrar o teste bivariado de causalidade à Granger, recolheram-se dados das séries da inflação homóloga e taxa de juro de empréstimos a sociedades não financeiras em Portugal (Fonte: Banco de Portugal) entre Janeiro de 2003 e Abril de 2012, num total de 88 observações mensais, como se pode observar na figura seguinte.



# MODELOS VAR

## Causalidade à Granger

Pairwise Granger Causality Tests			
Date: 03/07/12 Time: 15:28			
Sample: 2003M01 2012M04			
Lags: 4			
Null Hypothesis:	Obs	F-Statistic	Prob.
TJURO does not Granger Cause INFL	108	6.20547	0.0002
INFL does not Granger Cause TJURO		2.89645	0.0258

Pairwise Granger Causality Tests			
Date: 03/07/12 Time: 15:29			
Sample: 2003M01 2012M04			
Lags: 8			
Null Hypothesis:	Obs	F-Statistic	Prob.
TJURO does not Granger Cause INFL	104	3.44964	0.0017
INFL does not Granger Cause TJURO		1.57422	0.1443

Teste de causalidade à Granger (teste directo) entre inflação e taxa de juro em Portugal,  $p=4$  e  $p=8$

# MODELOS VAR

## Causalidade à Granger

### Pairwise Granger Causality Tests

Date: 09/13/13 Time: 12:23

Sample: 2003M01 2012M04

Lags: 4

Null Hypothesis:	Obs	F-Statistic	Prob.
DTJURO does not Granger Cause DINFL	107	3.59041	0.0089
DINFL does not Granger Cause DTJURO		2.95113	0.0238

### Pairwise Granger Causality Tests

Date: 09/13/13 Time: 12:24

Sample: 2003M01 2012M04

Lags: 8

Null Hypothesis:	Obs	F-Statistic	Prob.
DTJURO does not Granger Cause DINFL	103	3.04717	0.0046
DINFL does not Granger Cause DTJURO		2.00368	0.0554

# MODELOS VAR

## Causalidade à Granger

### Teste de Wald para a causalidade à Granger

No contexto multivariado de um modelo VAR(p), a relação de causalidade à Granger pode ser definida através da representação médias móveis (MA) do VAR,

$$Y_t = \mu + \sum_{i=0}^{\infty} \Psi_i \varepsilon_{t-i} = \mu + \Psi(L) \varepsilon_t, \text{ com } \Psi(L) = (I + \Psi_1 L + \Psi_2 L^2 + \dots) \quad (6)$$

Diz-se que a variável  $Y_{it}$  não causa (no sentido de Granger) a variável  $Y_{jt}$  se todos os coeficientes do polinómio  $\Psi_{ij}(L)$  forem iguais a zero, isto é, se  $\psi_{ij,k} = 0, k = 1, 2, \dots$

**Exemplo de aplicação:** Retomando os modelos VAR estimados para as variáveis consumo, investimento rendimento disponível, a aplicação do teste de causalidade à Granger/exogeneidade conduziu aos seguintes resultados:



# MODELOS VAR

## Causalidade à Granger

VAR Granger Causality/Block Exogeneity Wald Tests			
Date: 09/12/14 Time: 17:29			
Sample: 1995Q1 2010Q4			
Included observations: 61			
Dependent variable: DLOG(CONS)			
Excluded	Chi-sq	df	Prob.
DLOG(INVEST)	4.357592	2	0.1132
DLOG(REND)	10.20305	2	0.0061
All	17.31045	4	0.0017
Dependent variable: DLOG(INVEST)			
Excluded	Chi-sq	df	Prob.
DLOG(CONS)	1.712176	2	0.4248
DLOG(REND)	5.219635	2	0.0735
All	7.802276	4	0.0991
Dependent variable: DLOG(REND)			
Excluded	Chi-sq	df	Prob.
DLOG(CONS)	2.542580	2	0.2805
DLOG(INVEST)	0.193758	2	0.9077
All	4.653887	4	0.3247

Testes de causalidade à Granger para as variáveis consumo, investimento e rendimento disponível em Portugal através do VAR(2) em diferenças do logaritmo

# MODELOS VAR

## Causalidade à Granger

### Testes com variáveis não estacionárias

Como a maioria das séries económicas são não estacionárias ou  $I(1)$ , na estimação do VAR devemos ter atenção aos seguintes aspectos:

- Na estimação de um VAR com todas as variáveis  $I(0)$ , os testes baseados nas estatística  $t$  e  $F$  são apropriados;
- Se as variáveis forem todas  $I(1)$  e não cointegradas, é possível estimar o VAR em primeiras diferenças e aplicar os habituais testes de causalidade;
- Se as variáveis forem todas  $I(1)$  e cointegradas, não é possível escrever o VAR em primeiras diferenças e os testes de causalidade não podem ser usados com base nas estatística  $t$  e  $F$ .

# MODELOS VAR

## Funções de resposta a impulsos

A **função de resposta ao impulso** define o efeito do choque exógeno de uma perturbação aleatória sobre os valores presentes e passados das variáveis endógenas..

Considerando um modelo simples VAR(1) bivariado:

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}, \quad (1)$$

ou, usando a representação MA,  $Y_t = \mu + \sum_{i=0}^{\infty} \Psi_i \varepsilon_{t-i}$ ,

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} \bar{Y}_1 \\ \bar{Y}_2 \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \psi_{11}(i) & \psi_{12}(i) \\ \psi_{21}(i) & \psi_{22}(i) \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-i} \\ \varepsilon_{2,t-i} \end{bmatrix}, \quad (2)$$

onde  $\Psi_i = A_i^1$  é a  $i$ -ésima matriz de coeficientes da representação MA de um processo VAR(1) e pode ser usado para determinar os efeitos dos choques de  $\varepsilon_{1t}$  e  $\varepsilon_{2t}$  sobre as variáveis  $Y_{1t}$  e  $Y_{2t}$  do sistema. Em particular,  $\psi_{11}(i)$ ,  $\psi_{12}(i)$ ,  $\psi_{21}(i)$  e  $\psi_{22}(i)$  são designados por **funções de resposta a impulsos**.

# MODELOS VAR

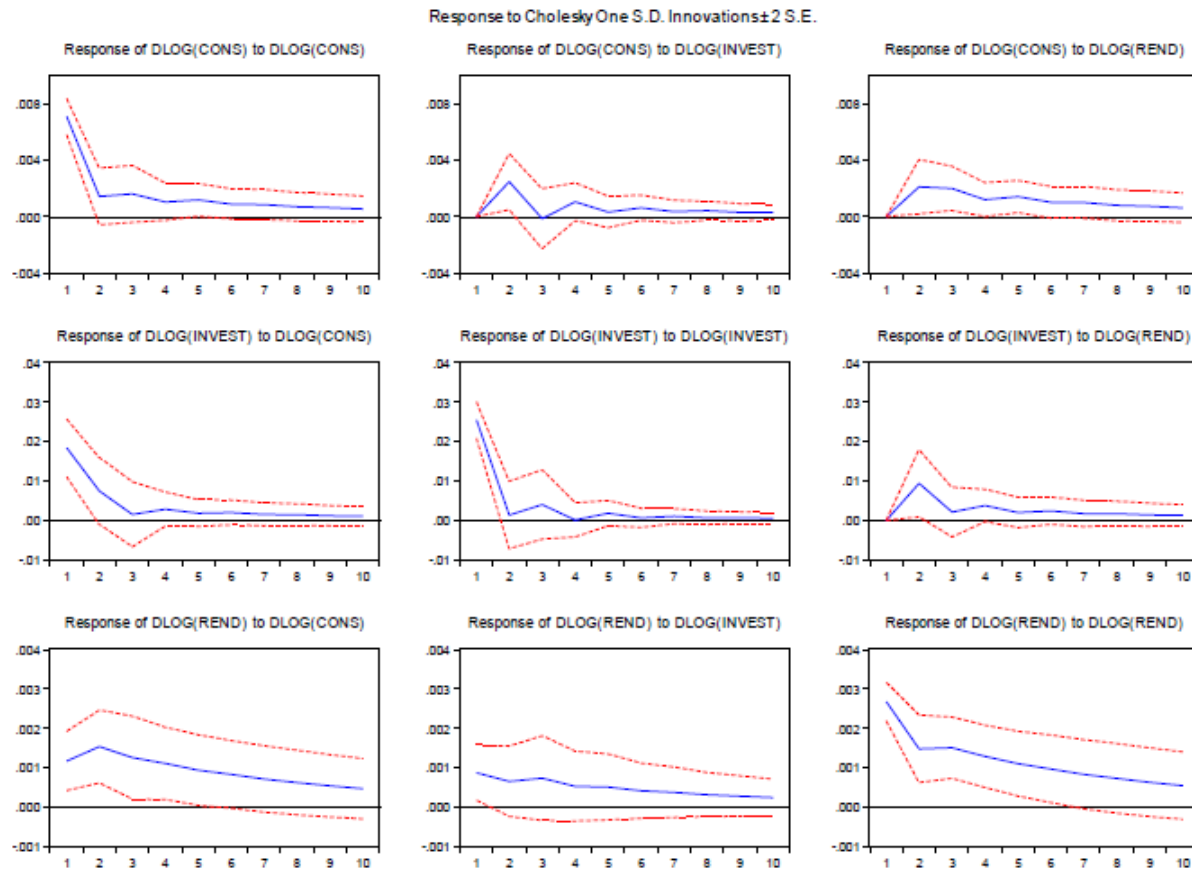
## Funções de resposta a impulsos

Na relação (1), a componente comum aos termos  $\varepsilon_{1t}$  e  $\varepsilon_{2t}$  é totalmente atribuída a  $\varepsilon_{1t}$  porque  $\varepsilon_{1t}$  precede  $\varepsilon_{2t}$ . Deste modo,  $\varepsilon_{1t}$  é o termo de perturbação de  $Y_1$  e  $\varepsilon_{2t}$  vai ser transformado para remover a componente comum a  $\varepsilon_{1t}$  através do método de factorização de Choleski, que consiste em aplicar uma transformação  $P$  às perturbações  $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$  para que estas venham não correlacionadas, isto é,  $u_t = P\varepsilon_t \sim (0, D)$ , onde  $u_t = (u_{1t}, u_{2t})'$  é um vector de perturbações ortogonais com matriz de covariâncias diagonal.

Uma vez que a decomposição de Choleski é um método arbitrário de atribuição de efeitos comuns, a alteração da ordem das equações no VAR pode alterar dramaticamente as respostas aos impulsos, como referem Hamilton (1994) e Lütkepohl (1991). Como alternativa, Pesaran e Shin (1998) propuseram uma função de resposta a impulsos generalizados que não depende da ordem pela qual as variáveis se encontram no VAR, que consiste na aplicação de um factor de Choleski específico para cada variável no sistema VAR.

# MODELOS VAR

## Funções de resposta a impulsos



Funções de resposta a impulsos do modelo VAR(2) para consumo, investimento e rendimento em diferenças de logaritmo

# MODELOS VAR

## Decomposição da variância

Enquanto as funções de resposta ao impulso traçam os efeitos de um choque de uma variável endógena sobre as restantes variáveis no VAR para descrever a dinâmica do sistema, a decomposição da variância atribui a variação de uma variável endógena em termos das perturbações ortogonais no sistema VAR. O método de decomposição da variância mede assim a importância relativa de cada perturbação aleatória para as variáveis do sistema VAR.

Num modelo VAR(1) bivariado da forma  $Y_t = A_0 + A_1 Y_{t-1} + \varepsilon_t$ , o valor esperado condicionado da previsão com origem em  $t$  e horizonte de  $h$  períodos, é dado por

$$E_t(Y_{t+h}) = Y_t(h) = (I + A_1 + \dots + A_1^{h-1})A_0 + A_1^h Y_t,$$

e o respectivo erro de previsão por

$$Y_{t+h} - Y_t(h) = \varepsilon_{t+h} + A_1 \varepsilon_{t+h-1} + A_1^2 \varepsilon_{t+h-2} + \dots + A_1^{h-1} \varepsilon_{t+1}.$$

# MODELOS VAR

## Decomposição da variância

ou, através da representação MA,

$$Y_{t+h} - Y_t(h) = \sum_{i=0}^{h-1} \Psi_i \varepsilon_{t+h-i} = \varepsilon_{t+h} + \Psi_1 \varepsilon_{t+h-1} + \dots + \Psi_{h-1} \varepsilon_{t+1}$$

A variância do erro de previsão de  $Y_{j,t}(h)$  vem dada por:

$$\begin{aligned} EQM[Y_{j,t}(h)] &= E(Y_{j,t+h} - Y_{j,t}(h))^2 = \psi_{11}^2(0) + \psi_{11}^2(1) + \dots + \psi_{11}^2(h-1) \\ &\quad + \psi_{12}^2(0) + \psi_{12}^2(1) + \dots + \psi_{12}^2(h-1) \end{aligned}$$

A partir desta expressão, podemos decompor a variância do erro de previsão em proporções/percentagens devidas a cada uma das perturbações aleatórias.

# MODELOS VAR

## Decomposição da variância

Variance Decomposition of DLOG(CONS):				
Period	S.E.	DLOG(CONS)	DLOG(INVEST)	DLOG(REND)
1	0.007100	100.0000	0.000000	0.000000
2	0.007933	83.27839	9.686854	7.034754
3	0.008336	79.11290	8.808612	12.07848
4	0.008547	76.69364	9.875227	13.43114

Variance Decomposition of DLOG(INVEST):				
Period	S.E.	DLOG(CONS)	DLOG(INVEST)	DLOG(REND)
1	0.031378	34.22963	65.77037	0.000000
2	0.033603	34.68274	57.50109	7.816170
3	0.033930	34.20743	57.76279	8.029780
4	0.034251	34.24077	56.68558	9.073652

Variance Decomposition of DLOG(REND):				
Period	S.E.	DLOG(CONS)	DLOG(INVEST)	DLOG(REND)
1	0.003047	14.74404	8.229891	77.02607
2	0.003775	26.17680	8.357387	65.46582
3	0.004317	28.47211	9.287687	62.24020
4	0.004668	29.97014	9.228669	60.80119

Cholesky Ordering: DLOG(CONS) DLOG(INVEST) DLOG(REND)				

Decomposição da variância do erro de previsão do modelo VAR(2)  
para consumo, investimento e rendimento em diferenças de logaritmo



# MODELOS VAR

## Seleccção da ordem

Um dos principais testes estatísticos para seleccionar a ordem do modelo VAR baseia-se no **rácio de verosimilhanças (LR)**. De acordo com Sims (1980), o teste de LR consiste em estimar um modelo VAR de ordem  $p$  (sem restrições) e um modelo de ordem  $p-r$  (com restrições) e testar a nulidade conjunta dos coeficientes associados ao modelo sem restrições.

Em geral, deve começar-se por estimar um VAR de ordem elevada e comparar com um VAR de ordem inferior em um desfasamento, e reduzir sempre um desfasamento a cada um dos modelos até se conseguir rejeitar a hipótese nula.

**Exemplo de aplicação:** Retomado os dados das variáveis macroeconómicas portuguesas, Consumo, Investimento e Rendimento disponível (em diferenças de logaritmo). Para testar a hipótese nula da especificação VAR de ordem 2 contra a de ordem 3, os passos a dar são os seguintes:

# MODELOS VAR

## Seleccção da ordem

$H_0 : VAR(2)$  contra  $H_1 : VAR(3)$

### VAR(3) em diferenças de logaritmo (sem restrições)

Determinant resid covariance (dof adj.)	2.55E-13
Determinant resid covariance	1.47E-13
Log likelihood	630.9481
Akaike information criterion	-20.03160
Schwarz criterion	-18.98443

### VAR(2) em diferenças de logaritmo (com restrições)

Determinant resid covariance (dof adj.)	2.33E-13
Determinant resid covariance	1.62E-13
Log likelihood	638.6107
Akaike information criterion	-20.24953
Schwarz criterion	-19.52284

# MODELOS VAR

## Seleccção da ordem

$H_0 : VAR(1)$  contra  $H_1 : VAR(2)$

### VAR(1) em diferenças de logaritmo

Determinant resid covariance (dof adj.)	2.47E-13
Determinant resid covariance	2.02E-13
Log likelihood	642.2448
Akaike information criterion	-20.33048
Schwarz criterion	-19.91877

### Considerações:

- O teste do rácio de verosimilhanças (LR) é um teste baseado na teoria assintótica,
- O teste LR só é aplicável quando um dos modelos é uma versão restrita do outro modelo.
- Como alternativa, foram desenvolvidos outros testes de selecção da ordem de defasamento, como sejam os testes de FPE (*Final Prediction Error*), AIC (*Akaike's Information Criterion*), BIC (*Bayesian Schwarz's Information Criterion*) e HQ (*Hannan-Quinn Information Criterion*).

# MODELOS VAR

## Seleccção da ordem

### Teste do erro de previsão mínimo (FPE)

Baseado na estatística de Akaike (1969, 1971), consiste em escolher a ordem do VAR que minimiza uma função do erro quadrático médio de previsão. Em geral, escolhe-se o VAR( $p$ ) que produz o menor erro final de previsão “final prediction error” obtido através da expressão:

$$FPE = \left( \frac{T + np + 1}{T - np - 1} \right)^n |\Sigma_\varepsilon|,$$

onde  $T$  é o número de observações,  $n$  é o número de variáveis,  $p$  é a ordem do VAR e  $|\Sigma_\varepsilon|$  é o determinante da matriz de variâncias e covariâncias dos resíduos.

# MODELOS VAR

## Seleccção da ordem

### **Teste de Informação de Akaike (AIC),**

Determina a ordem  $p$  do VAR que minimiza o seguinte critério

$$AIC = \ln|\Sigma_\varepsilon| + \frac{2pn^2}{T},$$

### **Teste Bayesiano de Informação de Schwartz (BIC)**

Utiliza o critério

$$BIC = \ln|\Sigma_\varepsilon| + \frac{\ln T}{T} pn^2,$$

### **Teste de Informação de Hannan-Quinn (HC)**

Foi introduzido por Hannan e Quinn (1979) e Quinn (1980) como alternativa aos testes AIC e FPE, por estes sobreestimarem a ordem  $p$  do VAR quando o número de variáveis  $n$  do sistema for reduzido. O critério HC é definido por:

$$HC = \ln|\Sigma_\varepsilon| + \frac{2\ln\ln T}{T} pn^2$$

# MODELOS VAR

## Seleccção da ordem

VAR Lag Order Selection Criteria

Endogenous variables: DLOG(CONS) DLOG(INVEST) DLOG(REND)

Exogenous variables: C

Date: 20/10/13 Time: 16:57

Sample: 1995Q1 2010Q4

Included observations: 58

Lag	LogL	LR	FPE	AIC	SC	HQ
0	564.5891	NA	7.81e-13	-19.36514	-19.25857	-19.32363
1	606.0491	77.20150*	2.55e-13*	-20.48445*	-20.05815*	-20.31840*
2	611.8743	10.24431	2.85e-13	-20.37498	-19.62895	-20.08439
3	615.7963	6.491538	3.42e-13	-20.19987	-19.13413	-19.78474
4	624.3985	13.34825	3.52e-13	-20.18615	-18.80068	-19.64649
5	628.2707	5.608046	4.28e-13	-20.00933	-18.30414	-19.34513

\* indicates lag order selected by the criterion

LR: sequential modified LR test statistic (each test at 5% level)

FPE: Final prediction error

AIC: Akaike information criterion

SC: Schwarz information criterion

HQ: Hannan-Quinn information criterion

# MODELOS VAR

## Avaliação do diagnóstico

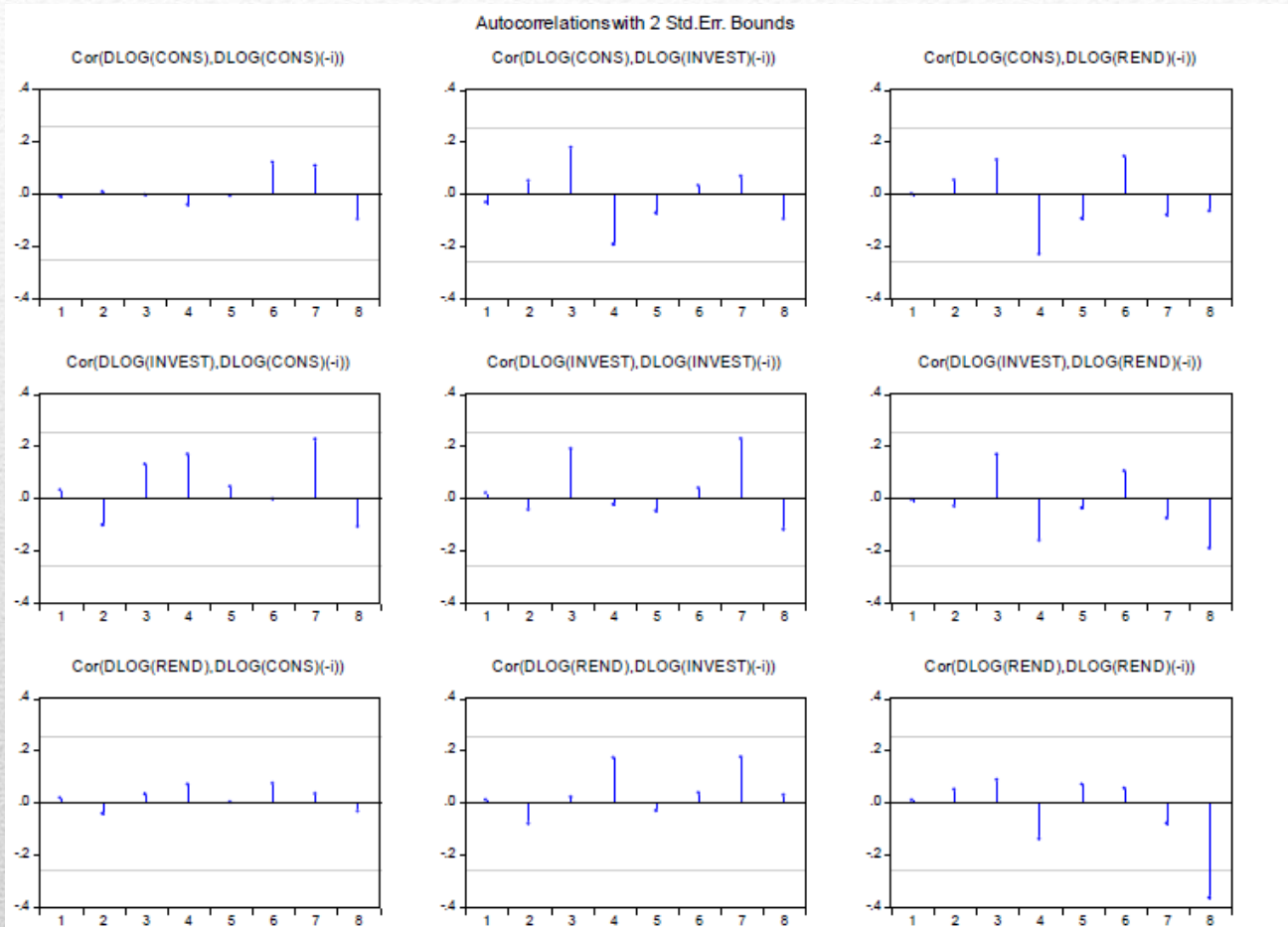
O **correlograma dos resíduos** fornece-nos as correlações cruzadas amostrais dos resíduos estimados do VAR para um número determinado de defasamentos (lags). Seja  $r_{mn}(i)$  o coeficiente de correlação do  $mn$ -ésimo elemento da matriz de autocorrelações dos resíduos,  $R_\varepsilon(i)$ . Um teste à hipótese nula de que o verdadeiro coeficiente de autocorrelação é zero (hipótese de ruído branco),  $H_0: \rho_{mn}(i) = 0$ , contra a alternativa  $H_1: \rho_{mn}(i) \neq 0$ , utiliza a seguinte regra de rejeição:

$$\text{Rejeita-se } H_0 \text{ se } |r_{mn}(i)| > 2/\sqrt{T}$$

**Exemplo de aplicação:** Dados macroeconómicos de Portugal do consumo, investimento e rendimento disponível.

# MODELOS VAR

## Avaliação do diagnóstico



Correlograma dos resíduos do modelo VAR(2) em diferenças de logaritmo: CONS/INVEST/REND



# MODELOS VAR

## Avaliação do diagnóstico

VAR Residual Cross-Correlations  
 Ordered by: variables  
 Date: 09/12/14 Time: 17:45  
 Sample: 1995Q1 2010Q4  
 Included observations: 61

	DLOG(CONS)	DLOG(INVEST)	DLOG(REND)
DLOG(CONS)	1.000000	0.585061	0.383980
DLOG(CONS(-1))	-0.012680	0.028377	0.015092
DLOG(CONS(-2))	0.003981	-0.105584	-0.044829
DLOG(CONS(-3))	-0.007514	0.127011	0.029724
DLOG(CONS(-4))	-0.044878	0.165571	0.066740
DLOG(CONS(-5))	-0.010557	0.041950	-0.001122
DLOG(CONS(-6))	0.117553	-0.005895	0.070571
DLOG(CONS(-7))	0.103953	0.223588	0.030317
DLOG(CONS(-8))	-0.100515	-0.111976	-0.036108
DLOG(INVEST)	0.585061	1.000000	0.457306
DLOG(INVEST(-1))	-0.035850	0.017019	0.007153
DLOG(INVEST(-2))	0.046673	-0.047238	-0.083224
DLOG(INVEST(-3))	0.173499	0.186251	0.018141
DLOG(INVEST(-4))	-0.196655	-0.027893	0.168578
DLOG(INVEST(-5))	-0.077383	-0.051214	-0.034837
DLOG(INVEST(-6))	0.028794	0.035407	0.035144
DLOG(INVEST(-7))	0.064134	0.224161	0.171487
DLOG(INVEST(-8))	-0.099171	-0.122309	0.025934
DLOG(REND)	0.383980	0.457306	1.000000
DLOG(REND(-1))	-0.003824	-0.008381	0.006256
DLOG(REND(-2))	0.049258	-0.033825	0.047563
DLOG(REND(-3))	0.126557	0.164726	0.084398
DLOG(REND(-4))	-0.234260	-0.164663	-0.142036
DLOG(REND(-5))	-0.096915	-0.041283	0.066689
DLOG(REND(-6))	0.140436	0.101301	0.051545
DLOG(REND(-7))	-0.084371	-0.080344	-0.082489
DLOG(REND(-8))	-0.069085	-0.194914	-0.367262

Asymptotic standard error (lag>0): 0.128037

VAR Residual Cross-Correlations  
 Ordered by: lags  
 Date: 09/12/14 Time: 17:47  
 Sample: 1995Q1 2010Q4  
 Included observations: 61

	DLOG(CONS)	DLOG(INVEST)	DLOG(REND)
DLOG(CONS)	1.000000	0.585061	0.383980
DLOG(INVEST)	0.585061	1.000000	0.457306
DLOG(REND)	0.383980	0.457306	1.000000
DLOG(CONS(-1))	-0.012680	0.028377	0.015092
DLOG(INVEST(-1))	-0.035850	0.017019	0.007153
DLOG(REND(-1))	-0.003824	-0.008381	0.006256
DLOG(CONS(-2))	0.003981	-0.105584	-0.044829
DLOG(INVEST(-2))	0.046673	-0.047238	-0.083224
DLOG(REND(-2))	0.049258	-0.033825	0.047563
DLOG(CONS(-3))	-0.007514	0.127011	0.029724
DLOG(INVEST(-3))	0.173499	0.186251	0.018141
DLOG(REND(-3))	0.126557	0.164726	0.084398
DLOG(CONS(-4))	-0.044878	0.165571	0.066740
DLOG(INVEST(-4))	-0.196655	-0.027893	0.168578
DLOG(REND(-4))	-0.234260	-0.164663	-0.142036
DLOG(CONS(-5))	-0.010557	0.041950	-0.001122
DLOG(INVEST(-5))	-0.077383	-0.051214	-0.034837
DLOG(REND(-5))	-0.096915	-0.041283	0.066689
DLOG(CONS(-6))	0.117553	-0.005895	0.070571
DLOG(INVEST(-6))	0.028794	0.035407	0.035144
DLOG(REND(-6))	0.140436	0.101301	0.051545
DLOG(CONS(-7))	0.103953	0.223588	0.030317
DLOG(INVEST(-7))	0.064134	0.224161	0.171487
DLOG(REND(-7))	-0.084371	-0.080344	-0.082489
DLOG(CONS(-8))	-0.100515	-0.111976	-0.036108
DLOG(INVEST(-8))	-0.099171	-0.122309	0.025934
DLOG(REND(-8))	-0.069085	-0.194914	-0.367262

Asymptotic standard error (lag>0): 0.128037

# MODELOS VAR

## Avaliação do diagnóstico

O **teste de autocorrelação de Portmanteau** utiliza a estatística  $Q$  de Ljung-Box/Box-Pierce multivariada para testar a autocorrelação dos resíduos até à ordem  $h$ , com base nas seguintes hipóteses estatísticas

$$H_0 : \mathbf{R}_h = (R_1, R_2, \dots, R_h) = 0 \text{ contra } H_1 : \mathbf{R}_h \neq 0$$

O software EViews fornece duas estatísticas: a estatística  $Q$ ,

$$Q_h = T \sum_{i=1}^h \text{tr}(\hat{R}_i' \hat{R}_\varepsilon^{-1} \hat{R}_i \hat{R}_\varepsilon^{-1}) = T \sum_{i=1}^h \text{tr}(\hat{C}_i' \hat{C}_0^{-1} \hat{C}_i \hat{C}_0^{-1})$$

e a estatística  $Q$  ajustada (isto é, com correcção para pequenas amostras),

$$\bar{Q}_h = T^2 \sum_{i=1}^h (T-i)^{-1} \text{tr}(\hat{C}_i' \hat{C}_0^{-1} \hat{C}_i \hat{C}_0^{-1}), \text{ onde } \hat{C}_i = \frac{1}{T} \sum_{t=i+1}^T \hat{\varepsilon}_t \hat{\varepsilon}_{t-i}'$$

Sobre a hipótese nula de ausência de autocorrelação até à ordem  $h$ , ambas as estatísticas têm distribuição aproximadamente do Qui-quadrado com  $k^2(h-p)$ .

# MODELOS VAR

## Avaliação do diagnóstico

O teste **LM (Lagrange Multiplier) de autocorrelação** assume que o vector dos erros segue um modelo VAR da forma  $\varepsilon_t = D_1\varepsilon_{t-1} + D_2\varepsilon_{t-2} + \dots + D_h\varepsilon_{t-h} + v_t$ , onde  $v_t$  é um ruído branco e  $D_j$  é uma matriz  $k \times k$  diagonal, em que os elementos diagonais são as raízes quadradas dos elementos diagonais de  $C_0$ . Para testar a hipótese nula de ausência de autocorrelação até à ordem  $h$ ,  $H_0 : D_1 = D_2 = \dots = D_h = 0$ , contra a alternativa,  $H_1 : \exists D_j \neq 0, j = 1, 2, \dots, h$ , efectua-se a regressão auxiliar dos resíduos sobre os regressores originais do segundo membro do VAR e sobre os resíduos desfasados. A estatística do teste LM,  $\lambda_{LM}(h)$  (ver Lutkepohl, 2005, p.172-173), tem distribuição assintótica do qui-quadrado com  $hk^2$  graus de liberdade.

Quando a ordem de autocorrelação  $h$  é elevada, o teste LM exige muitos graus de liberdade para correr a regressão auxiliar e a estatística do teste não pode ser calculada nos termos habituais. Daí que seja mais conveniente, proceder ao teste LM de autocorrelação com base na regressão auxiliar,

$$\hat{\varepsilon}_t = \alpha + A_1 Y_{t-1} + \dots + A_p Y_{t-p} + D_h \hat{\varepsilon}_{t-h} + u_t,$$

e testar

$$H_0 : D_h = 0 \text{ contra } H_1 : D_h \neq 0.$$

# MODELOS VAR

## Avaliação do diagnóstico

VAR Residual Portmanteau Tests for Autocorrelations  
Null Hypothesis: no residual autocorrelations up to lag h  
Date: 09/12/14 Time: 17:41  
Sample: 1995Q1 2010Q4  
Included observations: 61

Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	df
1	0.344378	NA*	0.350117	NA*	NA*
2	2.975000	NA*	3.069913	NA*	NA*
3	8.772398	0.4585	9.167177	0.4220	9
4	24.58228	0.1369	26.08652	0.0978	18
5	26.89948	0.4692	28.61062	0.3800	27
6	30.27735	0.7370	32.35699	0.6426	36
7	40.14157	0.6775	43.49990	0.5356	45
8	54.41396	0.4586	59.92660	0.2695	54

\*The test is valid only for lags larger than the VAR lag order.  
df is degrees of freedom for (approximate) chi-square distribution

Teste portmanteau de autocorrelação dos resíduos do modelo VAR(2) em diferenças de logaritmo: CONS/INVEST/REND

# MODELOS VAR

## Avaliação do diagnóstico

Lags	LM-Stat	Prob
1	6.746934	0.6634
2	13.84649	0.1279
3	5.827132	0.7571
4	15.99215	0.0670
5	2.496649	0.9810
6	3.407400	0.9459
7	10.10247	0.3423
8	19.98439	0.0180

Probs from chi-square with 9 df.

Teste LM de autocorrelação dos resíduos do modelo VAR(2) em diferenças de logaritmo: CONS/INVEST/REND

# MODELOS VAR

## Avaliação do diagnóstico

Em seguida, apresentam-se **testes de normalidade multivariada** dos resíduos.

No caso multivariado, é necessário escolher uma factorização dos  $k$  resíduos que sejam ortogonais a si próprios. Seja  $\varepsilon_t$  um processo ruído branco gaussiano  $k$ -dimensional, com  $\varepsilon_t \sim N(0, \Sigma_\varepsilon)$  e seja  $P$  uma matriz  $k \times k$  triangular inferior com elementos diagonais positivos (ver Apêndice A.9.3 de Lutkepohl, 2005), tal que:

$$PP' = \Sigma_\varepsilon$$

e

$$u_t = P^{-1}\varepsilon_t \sim N(0, I_k),$$

onde  $u_t$  é um ruído branco com média zero e matriz de variâncias e covariâncias dada por

$$\Sigma_u = P^{-1}\Sigma_\varepsilon(P^{-1})' = I_k.$$

# MODELOS VAR

## Avaliação do diagnóstico

Para construir o teste, vamos definir os estimadores dos 3º e 4º momentos:

$$m_3 = (m_{13}, m_{23}, \dots, m_{k3})' \text{ com } m_{i3} = \frac{1}{T} \sum_{t=1}^T u_{it}^3, \quad i = 1, 2, \dots, k$$

e

$$m_4 = (m_{14}, m_{24}, \dots, m_{k4})' \text{ com } m_{i4} = \frac{1}{T} \sum_{t=1}^T u_{it}^4, \quad i = 1, 2, \dots, k$$

Então, como  $\varepsilon_t$  é um ruído branco gaussiano com matriz de variâncias e covariâncias não singular, com  $\varepsilon_t \sim N(0, \Sigma_\varepsilon)$ , tem-se o seguinte resultado:

$$\sqrt{T} \begin{bmatrix} m_3 \\ m_4 - \mathbf{3}_k \end{bmatrix} \xrightarrow{d} N \left( 0, \begin{bmatrix} 6I_k & 0 \\ 0 & 24I_k \end{bmatrix} \right)$$

Demonstração: ver Lutkepohl, 2005, p. 174-175.

Como as componentes de  $m_3$  e  $m_4$  são independentes e normalmente distribuídas, podemos construir duas estatísticas do qui-quadrado, somando os quadrados de cada um destes momentos.

# MODELOS VAR

## Avaliação do diagnóstico

Para testar as hipóteses sobre  $m_3$  (skewness),

$$H_0 : E \begin{bmatrix} u_{1t}^3 \\ u_{2t}^3 \\ \vdots \\ u_{kt}^3 \end{bmatrix} = 0 \text{ contra } H_1 : E \begin{bmatrix} u_{1t}^3 \\ u_{2t}^3 \\ \vdots \\ u_{kt}^3 \end{bmatrix} \neq 0,$$

recorre-se ao teste estatístico

$$\lambda_{SKEW} = Tm'_1 m_1 / 6 \xrightarrow{d} \chi^2(k)$$

Para testar as hipóteses sobre  $m_4$  (kurtosis),

$$H_0 : E \begin{bmatrix} u_{1t}^4 \\ u_{2t}^4 \\ \vdots \\ u_{kt}^4 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ \vdots \\ 3 \end{bmatrix} = \mathbf{3}_k \text{ contra } H_1 : E \begin{bmatrix} u_{1t}^4 \\ u_{2t}^4 \\ \vdots \\ u_{kt}^4 \end{bmatrix} \neq \begin{bmatrix} 3 \\ 3 \\ \vdots \\ 3 \end{bmatrix} = \mathbf{3}_k,$$

recorre-se ao teste estatístico

$$\lambda_{KURT} = T(m_2 - \mathbf{3}_k)'(m_2 - \mathbf{3}_k) / 24 \xrightarrow{d} \chi^2(k)$$



# MODELOS VAR

## Avaliação do diagnóstico

Para testar conjuntamente as duas hipóteses nulas,

$$H_{S0} : E \begin{bmatrix} u_{1t}^3 \\ u_{2t}^3 \\ \vdots \\ u_{kt}^3 \end{bmatrix} = 0 \wedge H_{K0} : E \begin{bmatrix} u_{1t}^4 \\ u_{2t}^4 \\ \vdots \\ u_{kt}^4 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ \vdots \\ 3 \end{bmatrix} = \mathbf{3}_k$$

usa-se o teste estatístico

$$\lambda_{SK} = \lambda_{SKEW} + \lambda_{KURT} \xrightarrow{d} \chi^2(2k)$$

O software EViews permite a escolha das seguintes factorizações da matriz  $P$ :

- **Decomposição de Choleski** (Lutkepohl, 2005);
- **Decomposição de Doornik-Hansen** (Doornik e Hansen, 1994);
- **Decomposição de Urzua** (Urzua, 1997).

# MODELOS VAR

## Avaliação do diagnóstico

VAR Residual Normality Tests  
 Orthogonalization: Cholesky (Lutkepohl)  
 Null Hypothesis: residuals are multivariate normal  
 Date: 09/12/14 Time: 17:48  
 Sample: 1995Q1 2010Q4  
 Included observations: 61

Component	Skewness	Chi-sq	df	Prob.
1	-1.258107	16.09214	1	0.0001
2	-0.047046	0.022502	1	0.8808
3	0.155889	0.247065	1	0.6191
Joint		16.36170	3	0.0010

Component	Kurtosis	Chi-sq	df	Prob.
1	5.279751	13.20971	1	0.0003
2	2.770697	0.133641	1	0.7147
3	4.781936	8.070546	1	0.0045
Joint		21.41390	3	0.0001

Component	Jarque-Bera	df	Prob.
1	29.30185	2	0.0000
2	0.156143	2	0.9249
3	8.317611	2	0.0156
Joint	37.77560	6	0.0000

Teste de normalidade dos resíduos (ortogonalização de Cholesky) do modelo VAR(2) em diferenças do logaritmo: CONS/INVEST/REND

# MODELOS VAR

## Avaliação do diagnóstico

VAR Residual Normality Tests  
 Orthogonalization: Residual Correlation (Doornik-Hansen)  
 Null Hypothesis: residuals are multivariate normal  
 Date: 09/12/14 Time: 17:50  
 Sample: 1995Q1 2010Q4  
 Included observations: 61

Component	Skewness	Chi-sq	df	Prob.
1	-0.771014	6.110109	1	0.0134
2	-0.096192	0.113010	1	0.7367
3	-0.036940	0.016713	1	0.8971
Joint		6.239833	3	0.1005

Component	Kurtosis	Chi-sq	df	Prob.
1	3.507858	0.755371	1	0.3848
2	2.675954	0.001290	1	0.9714
3	4.772948	12.03639	1	0.0005
Joint		12.79305	3	0.0051

Component	Jarque-Bera	df	Prob.
1	6.865481	2	0.0323
2	0.114300	2	0.9445
3	12.05310	2	0.0024
Joint	19.03288	6	0.0041

Teste de normalidade dos resíduos (ortogonalização de Doornik-Hansen) do modelo VAR(2) em diferenças do logaritmo:CONS/INVEST/REND

# MODELOS VAR

## Avaliação do diagnóstico

VAR Residual Normality Tests				
Orthogonalization: Residual Covariance (Urzua)				
Null Hypothesis: residuals are multivariate normal				
Date: 09/12/14 Time: 17:51				
Sample: 1995Q1 2010Q4				
Included observations: 61				
Component	Skewness	Chi-sq	df	Prob.
1	-0.629010	4.434895	1	0.0352
2	-0.384884	1.660460	1	0.1975
3	0.111867	0.140272	1	0.7080
Joint		6.235627	3	0.1007
Component	Kurtosis	Chi-sq	df	Prob.
1	3.068201	0.088212	1	0.7665
2	3.114867	0.145173	1	0.7032
3	4.744084	10.98315	1	0.0009
Joint		11.21653	3	0.0106
Component	Jarque-Bera	df	Prob.	
1	4.523107	2	0.1042	
2	1.805634	2	0.4054	
3	11.12342	2	0.0038	
Joint	105.2628	25	0.0000	

Teste de normalidade dos resíduos (ortogonalização de Urzua) do modelo VAR(2) em diferenças do logaritmo: CONS/INVEST/REND

# MODELOS VAR

## Avaliação do diagnóstico

O **teste de heteroscedasticidade de White** é uma extensão do teste de White (White, 1980) para testar a heteroscedasticidade não condicionada nos erros das equações do sistema VAR (Kelejian, 1982 e Doornik, 1995). O teste consiste em correr a regressão de cada um dos produtos cruzados dos resíduos sobre os produtos cruzados dos regressores e testar a significância conjunta da regressão.

A estatística LM do teste tem distribuição do qui-quadrado com  $mn$  graus de liberdade,

$$LM_{white} \sim \chi^2(mn),$$

onde  $m = k(k+1)/2$  é o número de produtos cruzados dos resíduos no sistema e  $n$  é o número de variáveis do segundo membro incluídas na regressão do teste de White.

# MODELOS VAR

## Avaliação do diagnóstico

VAR Residual Heteroskedasticity Tests: No Cross Terms (only levels and squares)  
 Sample: 1995Q1 2010Q4  
 Included observations: 61

Joint test:

Chi-sq	df	Prob.
102.1364	72	0.0113

Individual components:

Dependent	R-squared	F(12,48)	Prob.	Chi-sq(12)	Prob.
res1*res1	0.232002	1.208347	0.3051	14.15212	0.2911
res2*res2	0.380841	2.460372	0.0136	23.23128	0.0258
res3*res3	0.460433	3.413355	0.0012	28.08643	0.0054
res2*res1	0.342482	2.083484	0.0364	20.89140	0.0520
res3*res1	0.410120	2.781038	0.0059	25.01731	0.0147
res3*res2	0.376524	2.415641	0.0153	22.96795	0.0280

# MODELOS VAR

## Avaliação do diagnóstico

VAR Residual Heteroskedasticity Tests: Includes Cross Terms

Date: 09/12/14 Time: 17:54

Sample: 1995Q1 2010Q4

Included observations: 61

Joint test:

Chi-sq	df	Prob.
185.2992	162	0.1014

Individual components:

Dependent	R-squared	F(27,33)	Prob.	Chi-sq(27)	Prob.
res1*res1	0.514292	1.294151	0.2385	31.37182	0.2561
res2*res2	0.574144	1.647817	0.0859	35.02281	0.1383
res3*res3	0.709421	2.983937	0.0016	43.27467	0.0245
res2*res1	0.632934	2.107488	0.0212	38.60899	0.0687
res3*res1	0.630101	2.081981	0.0230	38.43614	0.0712
res3*res2	0.563805	1.579788	0.1053	34.39212	0.1549

# MODELOS VAR ESTRUTURAIS

Embora o modelo VAR sem restrições, seja (por simplicidade) o modelo VAR(1),

$$Y_t = A_0 + A_1 Y_{t-1} + \varepsilon_t,$$

conduza a previsões não enviesadas e de variância mínima, se estiver disponível informação a priori acerca de qualquer um dos coeficientes do modelo, será obviamente possível melhorar a precisão dos estimadores e reduzir a variância dos erros de previsão.

Considere-se um modelo VAR(1) bivariado escrito na **forma reduzida**:

$$\begin{aligned} Y_{1t} &= b_{10} - b_{12} Y_{2t} + c_{11} Y_{1,t-1} + c_{12} Y_{2,t-1} + u_{1t} \\ Y_{2t} &= b_{20} - b_{21} Y_{1t} + c_{21} Y_{1,t-1} + c_{22} Y_{2,t-1} + u_{2t} \end{aligned} \quad (1)$$

ou

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \quad (2)$$



# MODELOS VAR ESTRUTURAIS

ou ainda

$$BY_t = \Gamma_0 + \Gamma_1 Y_{t-1} + u_t. \quad (3)$$

Pré-multiplicando ambos os membros por  $B^{-1}$ , obtem-se o VAR na **forma standard**:

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \quad (4)$$

ou

$$Y_t = A_0 + A_1 Y_{t-1} + \varepsilon_t, \quad (5)$$

onde:  $A_0 = B^{-1}\Gamma_0$ ,  $A_1 = B^{-1}\Gamma_1$  e  $\varepsilon_t = B^{-1}u_t$ .

# MODELOS VAR ESTRUTURAIS

O **objectivo da análise VAR estrutural** é usar a teoria económica para recuperar as inovações estruturais dos resíduos  $\varepsilon_{1t}$  e  $\varepsilon_{2t}$ , pois se pretendermos obter as funções de resposta a impulsos ou as funções de decomposição da variância, é necessário usar os choques estruturais ( $u_{1t}$  e  $u_{2t}$ ) e não os erros de previsão ( $\varepsilon_{1t}$  e  $\varepsilon_{2t}$ ).

A decomposição de Choleski impõe um forte pressuposto acerca dos erros estruturais do modelo. Suponha que na relação (1), se escolhe uma ordenação das variáveis tal que  $b_{21} = 0$ , o que resulta em

$$Y_{1t} = b_{10} - b_{12}Y_{2t} + c_{11}Y_{1,t-1} + c_{12}Y_{2,t-1} + u_{1t}$$

$$Y_{2t} = b_{20} + c_{21}Y_{1,t-1} + c_{22}Y_{2,t-1} + u_{2t}$$

# MODELOS VAR ESTRUTURAIS

Neste caso a relação entre os choques estruturais  $u_{1t}$  e  $u_{2t}$  e os erros de previsão  $\varepsilon_{1t}$  e  $\varepsilon_{2t}$ , passa a ser:

$$\varepsilon_{1t} = u_{1t} - b_{12}u_{2t}$$

$$\varepsilon_{2t} = u_{2t}$$

ou

$$u_{1t} = \varepsilon_{1t} + b_{12}\varepsilon_{2t}$$

$$u_{2t} = \varepsilon_{2t}$$

E portanto, impor a restrição  $b_{21} = 0$  é equivalente a assumir que  $Y_{2t}$  tem um efeito contemporâneo sobre  $Y_{1t}$  mas  $Y_{1t}$  não tem um efeito contemporâneo sobre  $Y_{2t}$  mas apenas ao fim de um período. Este pressuposto deve ser devidamente justificado pela teoria económica, caso contrário, os choques podem não ser correctamente identificados. Mais, nestas condições, as funções de resposta a impulsos e de decomposição da variância deixam de fazer sentido.

# MODELOS VAR ESTRUTURAIS

## Decomposição de Sims-Bernanke

Considere-se um modelo VAR(1) com  $n=3$  variáveis:

$$\begin{bmatrix} 1 & b_{12} & b_{13} \\ b_{21} & 1 & b_{23} \\ b_{31} & b_{32} & 1 \end{bmatrix} \begin{bmatrix} Y_{1t} \\ Y_{2t} \\ Y_{3t} \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \\ b_{30} \end{bmatrix} + \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \\ Y_{3,t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \end{bmatrix}$$

ou

$$BY_t = \Gamma_0 + \Gamma_1 Y_{t-1} + u_t.$$

Pré-multiplicando ambos os membros por  $B^{-1}$ , sai:

$$Y_t = B^{-1}\Gamma_0 + B^{-1}\Gamma_1 Y_{t-1} + B^{-1}u_t$$

ou

$$Y_t = A_0 + A_1 Y_{t-1} + \varepsilon_t,$$

onde  $A_0 = B^{-1}\Gamma_0$ ,  $A_1 = B^{-1}\Gamma_1$  e  $\varepsilon_t = B^{-1}u_t$ .

O problema é obter os valores observados de  $\varepsilon_t$  e restringir o sistema de modo a que se consiga recuperar  $u_t$  através da relação  $u_t = B\varepsilon_t$ .

# MODELOS VAR ESTRUTURAIS

## Decomposição de Sims-Bernanke

### Exemplo de aplicação:

Suponha-se que se tem 5 observações residuais de um processo VAR(1) bivariado:

$t$	1	2	3	4	5
$\varepsilon_{1t}$	1	-0.5	0	-1	0.5
$\varepsilon_{2t}$	0.5	-1	0	-0.5	1

Identifique os choques estruturais através da decomposição de Choleski, impondo a restrição  $b_{12} = 0$ .

E se considerasse a decomposição de Choleski impondo a restrição  $b_{21} = 0$ ?

A decomposição de Choleski é apenas uma das formas possíveis de identificação de restrições. Existem outras alternativas: restrição dos coeficientes, restrição da variância e restrições simétricas.

# MODELOS VAR ESTRUTURAIS

## Decomposição de Blanchard-Quah

Para ilustrar o método de decomposição de Blanchard-Quah, suponha-se que estamos interessados em decompor um processo  $Y_t \sim I(1)$  nas duas componentes, temporária e permanente. Seja  $Z_t \sim I(0)$  uma segunda variável que é afectada pelos mesmos dois choques. Se omitirmos os termos independentes, a representação média do processo bivariado  $Y_t$  e  $Z_t$  vem dada por:

$$\Delta Y_t = \sum_{k=0}^{\infty} c_{11}(k) u_{1,t-k} + \sum_{k=0}^{\infty} c_{12}(k) u_{2,t-k}$$
$$Z_t = \sum_{k=0}^{\infty} c_{21}(k) u_{1,t-k} + \sum_{k=0}^{\infty} c_{22}(k) u_{2,t-k}$$

ou, na forma compactada

$$\begin{bmatrix} \Delta Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} C_{11}(L) & C_{12}(L) \\ C_{21}(L) & C_{22}(L) \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \quad (1)$$

# MODELOS VAR ESTRUTURAIS

## Decomposição de Blanchard-Quah

Para decompor  $Y_t$  nas suas componentes, permanente e temporária, é necessário assumir que **um dos choques tem um efeito temporário** sobre  $Y_t$ . Blanchard e Quah assumem que um choque na procura agregada não tem um efeito de longo prazo sobre o GNP real. Neste caso, o efeito cumulativo de um choque de  $u_{1t}$  sobre  $\Delta Y_t$  deve ser igual a zero, e portanto, os coeficientes de  $c_{ij}(k)$  são tais que:

$$\sum_{k=0}^{\infty} c_{11}(k) u_{1,t-k} = 0,$$

o que para qualquer possível realização de  $u_{1t}$ , implica que:

$$\sum_{k=0}^{\infty} c_{11}(k) = 0.$$

# MODELOS VAR ESTRUTURAIS

## Decomposição de Blanchard-Quah

Como choques do lado da oferta e do lado da procura não são observados, coloca-se o problema de como recuperá-los da estimação do VAR. Como as variáveis são estacionárias, podemos representar o VAR através da forma:

$$\begin{bmatrix} \Delta Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} A_{11}(L) & A_{12}(L) \\ A_{21}(L) & A_{22}(L) \end{bmatrix} \begin{bmatrix} \Delta Y_{t-1} \\ Z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \quad (2)$$

ou

$$X_t = A(L)X_{t-1} + \varepsilon_t$$

A questão crítica é que os resíduos do VAR,  $\varepsilon_{1t}$  e  $\varepsilon_{2t}$ , são composições das inovações puras  $u_{1t}$  e  $u_{2t}$ . Por exemplo,  $\varepsilon_{1t}$  é o erro de previsão a um passo à frente de  $Y_t$ , isto é,  $\varepsilon_{1t} = \Delta Y_t - E_{t-1}(\Delta Y_t)$ . A partir do modelo VMA, o erro de previsão a 1-passo à frente vem dado por  $c_{11}(0)u_{1t} + c_{12}(0)u_{2t}$ . Como as duas representações são semelhantes, temos que

$$\varepsilon_{1t} = c_{11}(0)u_{1t} + c_{12}(0)u_{2t}$$

e, por analogia,

$$\varepsilon_{2t} = c_{21}(0)u_{1t} + c_{22}(0)u_{2t}$$



# MODELOS VAR ESTRUTURAIS

## Decomposição de Blanchard-Quah

Se  $c_{11}(0), c_{12}(0), c_{21}(0)$  e  $c_{22}(0)$  forem conhecidos, é possível recuperar  $u_{1t}$  e  $u_{2t}$  a partir dos resíduos  $\varepsilon_{1t}$  e  $\varepsilon_{2t}$ . Blanchard e Quah mostram que a relação entre o modelo VAR dado em (2) e o modelo VMA dado em (1) mais a relação de longo prazo  $\sum_{k=0}^{\infty} c_{11}(k) = 0$ , formam exactamente 4 restrições que podem ser usadas para identificar estes quatro coeficientes.

Notando que os resíduos do VAR podem ser usados para construir estimativas de  $Var(\varepsilon_{1t})$ ,  $Var(\varepsilon_{2t})$ ,  $cov(\varepsilon_{1t}\varepsilon_{2t}) = cov(\varepsilon_{2t}\varepsilon_{1t})$ , as 4 restrições são as seguintes:

**Restrição 1)**  $var(\varepsilon_{1t}) = c_{11}^2(0) + c_{12}^2(0)$

**Restrição 2)**  $var(\varepsilon_{2t}) = c_{21}^2(0) + c_{22}^2(0)$

**Restrição 3)**  $E(\varepsilon_{1t}\varepsilon_{2t}) = c_{11}(0)c_{21}(0) + c_{12}(0)c_{22}(0)$

**Restrição 4)**  $\left[ 1 - \sum_{k=0}^{\infty} a_{22}(k) \right] c_{11}(0) + \sum_{k=0}^{\infty} a_{12}(k) c_{21}(0) = 0$

# MODELOS VAR ESTRUTURAIS

## Exemplo de Aplicação: Dados macroeconómicos de Portugal

Structural VAR Estimates				
Date: 11/12/12 Time: 14:36				
Sample (adjusted): 1995Q3 2010Q4				
Included observations: 62 after adjustments				
Estimation method: method of scoring (analytic derivatives)				
Convergence achieved after 7 iterations				
Structural VAR is just-identified				
Model: $Ae = Bu$ where $E[uu'] = I$				
Restriction Type: short-run pattern matrix				
A =				
	1	0	0	
C(1)	1	0		
C(2)	C(3)	1		
B =				
C(4)	0	0		
0	C(5)	0		
0	0	C(6)		
	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	-2.549536	0.444391	-5.737148	0.0000
C(2)	-0.084197	0.057246	-1.470788	0.1413
C(3)	-0.027930	0.013222	-2.112286	0.0347
C(4)	0.007303	0.000656	11.13553	0.0000
C(5)	0.025555	0.002295	11.13553	0.0000
C(6)	0.002661	0.000239	11.13553	0.0000
Log likelihood	636.0425			
Estimated A matrix:				
1.000000	0.000000	0.000000		
-2.549536	1.000000	0.000000		
-0.084197	-0.027930	1.000000		
Estimated B matrix:				
0.007303	0.000000	0.000000		
0.000000	0.025555	0.000000		
0.000000	0.000000	0.002661		

**Figura 44:** Decomposição estrutural com restrições de curto prazo VAR(1) em diferenças de logaritmo para CONS/INVES/REND

# MODELOS VAR ESTRUTURAIS

## Exemplo de Aplicação: Dados macroeconômicos de Portugal

Structural VAR Estimates					
Date: 11/12/12 Time: 15:14					
Sample (adjusted): 1995Q3 2010Q4					
Included observations: 62 after adjustments					
Estimation method: method of scoring (analytic derivatives)					
Convergence achieved after 7 iterations					
Structural VAR is just-identified					
Model: $Ae = Bu$ where $E[uu'] = I$					
Restriction Type: long-run pattern matrix					
Long-run response pattern:					
	C(1)	C(2)	C(4)		
	0	C(3)	C(5)		
	0	0	C(6)		
		Coefficient	Std. Error	z-Statistic	
				Prob.	
	C(1)	0.004488	0.000403	11.13553	0.0000
	C(2)	0.003979	0.000673	5.914872	0.0000
	C(3)	0.026996	0.002424	11.13553	0.0000
	C(4)	0.019812	0.001935	10.23680	0.0000
	C(5)	0.056612	0.006132	9.232313	0.0000
	C(6)	0.016719	0.001501	11.13553	0.0000
Log likelihood	636.0425				
Estimated A matrix:					
	1.000000	0.000000	0.000000		
	0.000000	1.000000	0.000000		
	0.000000	0.000000	1.000000		
Estimated B matrix:					
	0.005232	0.002640	0.004358		
	-0.002186	0.025737	0.018237		
	-0.000533	-0.000555	0.002878		

**Figura 45:** Decomposição estrutural com restrições de longo prazo VAR(1) em diferenças de logaritmo para CONS/INVES/REND

# COINTEGRAÇÃO E MODELOS VEC

A **cointegração** refere-se à relação de equilíbrio no longo prazo entre duas ou mais séries temporais não estacionárias. O conceito foi formalmente introduzido por Granger (1981) e Engle e Granger (1987) e traduz-se no seguinte.

Sejam  $\beta = (\beta_1, \beta_2, \dots, \beta_k)'$  um vector de  $k$  parâmetros e  $Y_t = (Y_{1t}, Y_{2t}, \dots, Y_{kt})'$  um vector de  $k$  variáveis económicas com características de equilíbrio no longo prazo, tal que a situação de equilíbrio ocorra quando se verifica a restrição,

$$\beta Y_t = 0.$$

Na maior parte do tempo  $Y_t$  encontra-se fora da situação de equilíbrio, sendo a distância a que o sistema se encontra da situação de equilíbrio, designada por **erro de equilíbrio**, dada por

$$E_t = \beta Y_t.$$

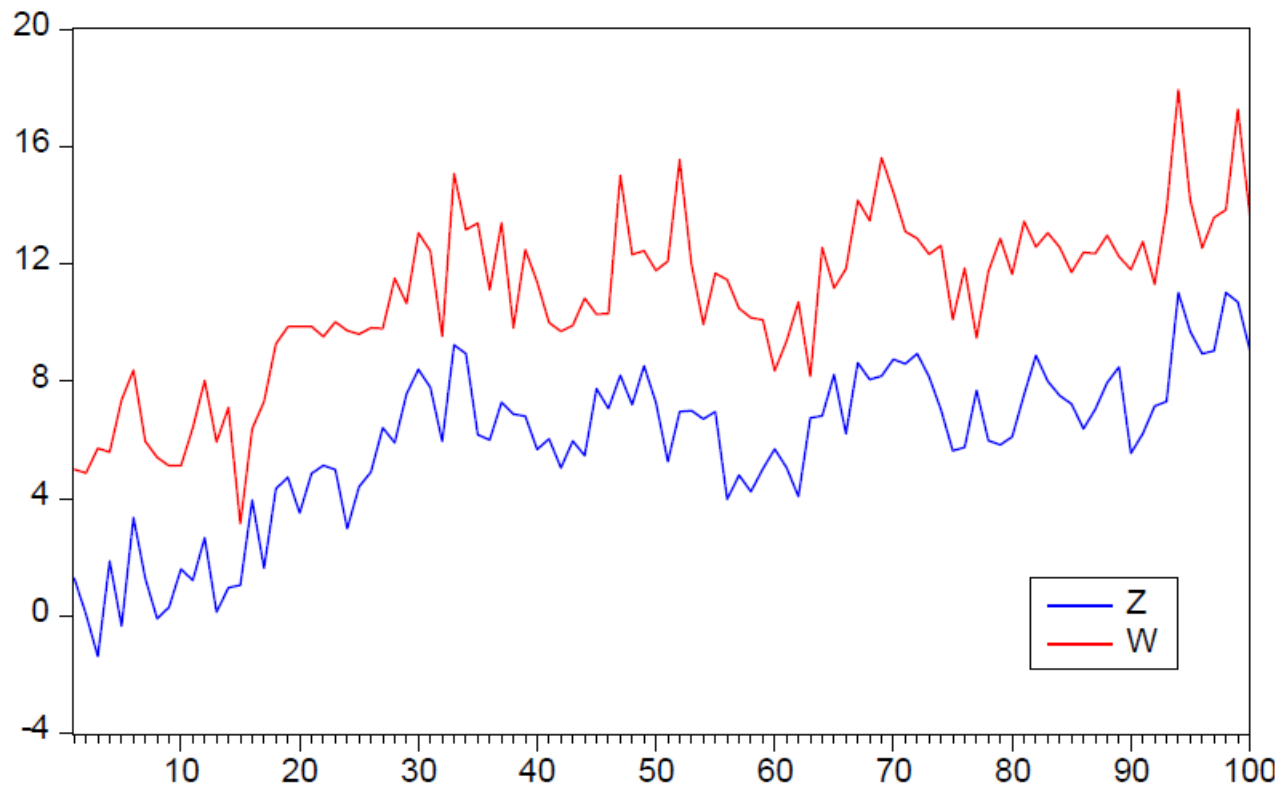
# COINTEGRAÇÃO E MODELOS VEC

Estes desvios devem ser limitados para que o processo do erro de equilíbrio seja estacionário. Assim, diz-se que as componentes do vector  $Y_t = (Y_{1t}, Y_{2t}, \dots, Y_{kt})'$  são **cointegradas de ordem  $(d, b)$** , e designa-se por  $Y_t \sim CI(d, b)$  se:

- Todas as componentes do vector  $Y_t$  são integradas de ordem  $d$  ou  $I(d)$ ;
- Existe um vector  $\beta = (\beta_1, \beta_2, \dots, \beta_k)'$  tal que a combinação linear  $\beta Y_t$  seja integrada de ordem  $d - b$  ou  $I(d - b)$ , com  $b > 0$ .

Nesta combinação linear, o vector  $\beta$  é chamado o **vector cointegrante**. Por exemplo, considerando o caso mais comum em que  $d = b = 1$ , se o vector  $Y_t$  é cointegrado, cada componente ou variável de  $Y_t$  é integrada de ordem 1 ou  $I(1)$ , e a combinação linear dessas variáveis é integrada de ordem 0 ou  $I(0)$ , sendo, portanto, esta última estacionária.

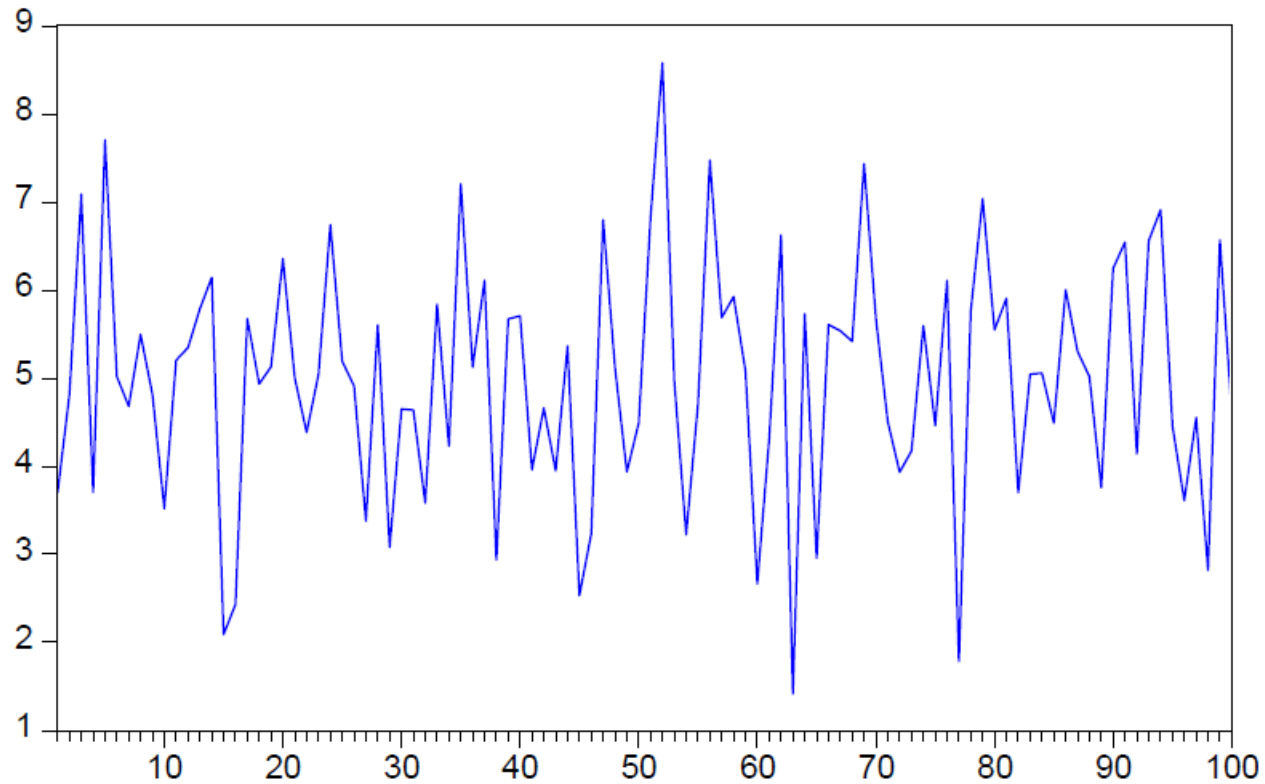
# COINTEGRAÇÃO E MODELOS VEC



Séries simuladas de dois processos integrados:  $Z_t = \mu_t + \varepsilon_{1t}$  e  $W_t = 5 + \mu_t + \varepsilon_{2t}$ ,  
onde  $\mu_t = \mu_{t-1} + \varepsilon_t$  é um passeio aleatório;  $Z_t$  e  $W_t$  são cointegradas ou

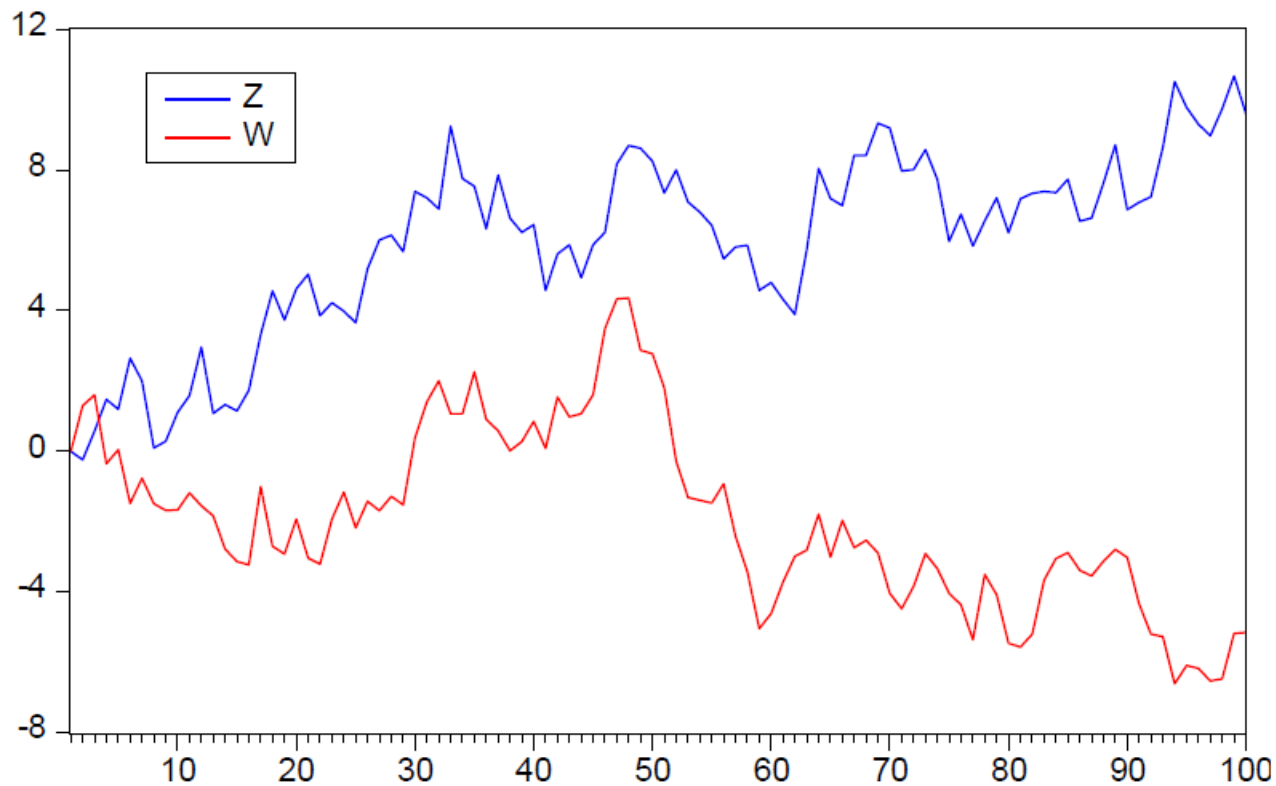
$CI(1,1)$

# COINTEGRAÇÃO E MODELOS VEC



Erro de equilíbrio:  $E_t = W_t - Z_t$

# COINTEGRAÇÃO E MODELOS VEC

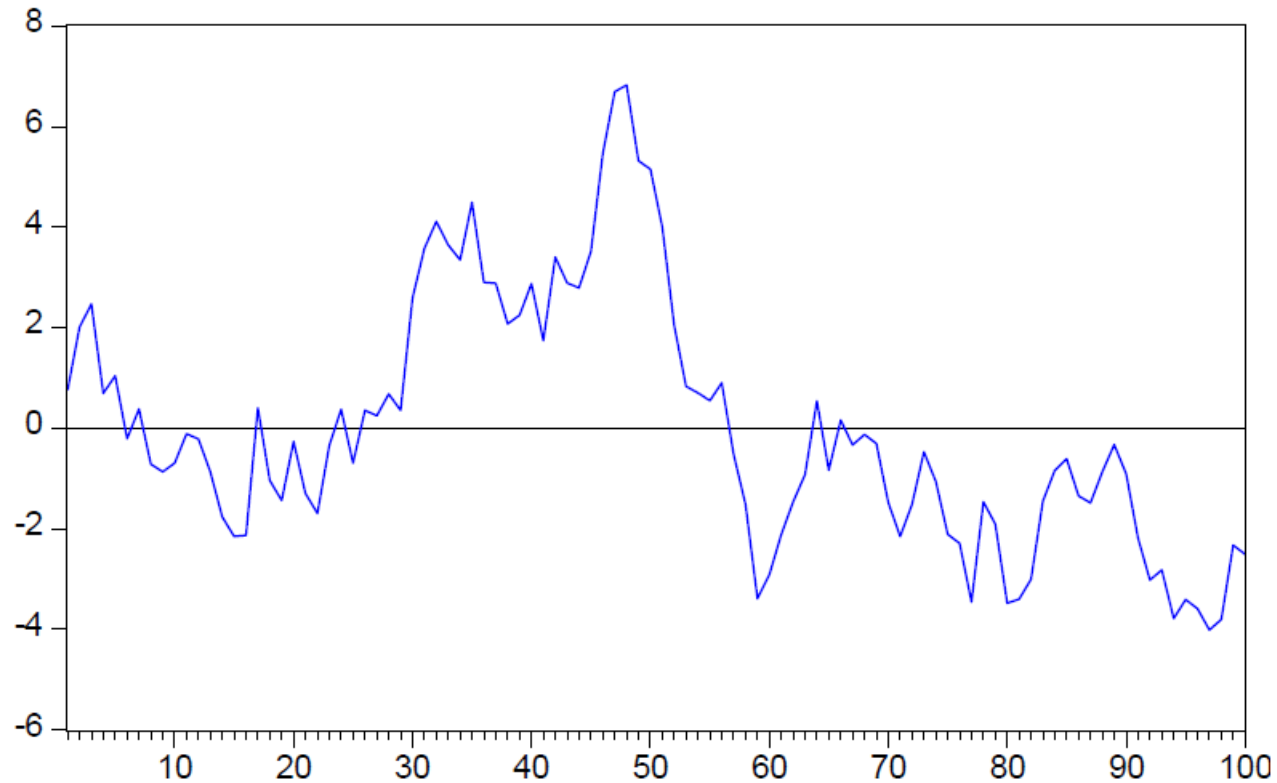


Séries simuladas de dois processos integrados:  $Z_t = Z_{t-1} + \varepsilon_{1t}$  e  $W_t = W_{t-1} + \varepsilon_{2t}$

;  $Z_t$  e  $W_t$  não são cointegradas



# COINTEGRAÇÃO E MODELOS VEC



Resíduos da regressão de cointegração:  $W_t = a + bZ_t + u_t$

# COINTEGRAÇÃO E MODELOS VEC

Se as variáveis  $Y_t$  e  $X_t$  forem não estacionárias ou integradas de ordem um  $I(1)$ , a regressão de cointegração (ou de equilíbrio de longo prazo),

$$Y_t = \beta_0 + \beta_1 X_t + u_t,$$

em geral, conduz a estimativas  $\hat{\beta}_1$  e  $\hat{\beta}_2$  pouco satisfatórias (fenómeno de regressão espúria). Uma forma de ultrapassar este problema consiste em diferenciar as variáveis para assegurar a sua estacionaridade, isto é, de modo a que  $\Delta X_t \sim I(0)$  e  $\Delta Y_t \sim I(0)$ , e depois estimar a regressão,

$$\Delta Y_t = a_0 + a_1 \Delta X_t + \Delta u_t,$$

Este modelo produz estimativas correctas para  $a_0$  e  $a_1$ , resolvendo o problema da regressão espúria. No entanto, o modelo perdeu a informação sobre a relação de equilíbrio de longo prazo, frequentemente estabelecida pela teoria económica, conseguindo apenas captar a relação de curto prazo entre as variáveis. Para resolver este problema, foi introduzido o **modelo de correcção do erro**.

# COINTEGRAÇÃO E MODELOS VEC

Como vimos anteriormente, se  $Y_t$  e  $X_t$  são cointegradas, existe uma combinação linear de  $Y_t$  e  $X_t$  que é estacionária ou  $I(0)$ . Deste modo, os resíduos da equação de cointegração,

$$\hat{u}_t = Y_t - \hat{\beta}_1 - \hat{\beta}_2 X_t,$$

representam uma combinação linear de  $Y_t$  e  $X_t$  no longo prazo.

Nestas condições, como  $\hat{u}_t$  é  $I(0)$ , podemos expressar a relação entre  $Y_t$  e  $X_t$  através de uma especificação de **modelo de correcção do erro** ou ECM (Error-correction model):

$$\Delta Y_t = b_0 + b_1 \Delta X_t + \pi \hat{u}_{t-1} + \varepsilon_t,$$

onde  $\hat{u}_{t-1} = Y_{t-1} - \hat{\beta}_1 - \hat{\beta}_2 X_{t-1}$  representa o erro desfasado um período estimado da equação de cointegração;  $b_1$  representa o impacto da variação de  $X_t$  sobre  $Y_t$  (efeito de curto prazo); e  $\pi$  representa o efeito de ajustamento, isto é, mede a rapidez com que qualquer desequilíbrio no período anterior influencia o equilíbrio de longo prazo.

# COINTEGRAÇÃO E MODELOS VEC

## Formalizando:

Num modelo de correcção do erro (MCE), a dinâmica de curto prazo das variáveis no sistema é influenciada pelo desvio em relação ao equilíbrio de longo prazo. Supondo que as variáveis  $Y_t$  e  $X_t$  são  $I(1)$ , um modelo simples de **correcção do erro** pode ser expresso por:

$$\Delta X_t = \alpha_1(Y_{t-1} - \beta X_{t-1}) + \varepsilon_{1t}, \quad \text{com } \alpha_1 > 0$$

$$\Delta Y_t = -\alpha_2(Y_{t-1} - \beta X_{t-1}) + \varepsilon_{2t}, \quad \text{com } \alpha_2 > 0$$

onde  $\varepsilon_{1t}$  e  $\varepsilon_{2t}$  são ruídos brancos que pode estar correlacionados,  $\alpha_1$  e  $\alpha_2$  são parâmetros que representam a velocidade de ajustamento com que qualquer desvio (ou desequilíbrio) no período anterior afecta o equilíbrio de longo prazo, e  $\beta$  é um coeficiente que representa a resposta de longo prazo.

# COINTEGRAÇÃO E MODELOS VEC

Estes resultados podem ser generalizados para um modelo com  $k$  variáveis. Seja  $Y_t = (Y_{1t}, \dots, Y_{kt})'$  um vector de  $k$  variáveis endógenas que admite uma representação de modelo de correcção do erro da forma:

$$\Delta Y_t = \Pi Y_{t-1} + \Pi_1 \Delta Y_{t-1} + \dots + \Pi_p \Delta Y_{t-p} + \varepsilon_t,$$

onde  $\Pi$  é uma matriz  $k \times k$  com elementos  $\pi_{js}$ , tal que pelo menos um  $\pi_{js} \neq 0$ ;  $\Pi_i$  são matrizes  $k \times k$  com elementos  $\pi_{rs}(i)$ ;  $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{kt})'$  é um vector  $k \times 1$  de perturbações aleatórias. Se todas as variáveis de  $Y_t$  forem  $I(1)$  e existir uma representação de correcção do erro dessas variáveis na forma anterior, então existe uma combinação linear de variáveis  $I(1)$  que é estacionária.

# COINTEGRAÇÃO E MODELOS VEC

Para melhor ilustrar a relação entre a cointegração e o mecanismo de correcção do erro, considere-se um modelo simples VAR(1) bivariado:

$$Y_t = a_{11}Y_{t-1} + a_{12}X_{t-1} + \varepsilon_{1t} \quad (1)$$

$$X_t = a_{21}Y_{t-1} + a_{22}X_{t-1} + \varepsilon_{2t} \quad (2)$$

onde  $\varepsilon_{1t}$  e  $\varepsilon_{2t}$  são termos ruído branco que podem estar correlacionados entre si. Usando o operador atraso, podemos escrever as equações (1) e (2) na forma

$$(1 - a_{11}L)Y_t - a_{12}LX_t = \varepsilon_{1t} \quad (3)$$

$$-a_{21}LY_t + (1 - a_{22}L)X_t = \varepsilon_{2t} \quad (4)$$

Resolvendo em ordem a  $Y_t$  e  $X_t$ , e utilizando notação matricial, vem

$$\begin{bmatrix} 1 - a_{11}L & -a_{12}L \\ -a_{21}L & 1 - a_{22}L \end{bmatrix} \begin{bmatrix} Y_t \\ X_t \end{bmatrix} = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

# COINTEGRAÇÃO E MODELOS VEC

Usando a Regra de Cramer para resolver esta equação matricial, obtém-se as soluções de  $Y_t$  e  $X_t$ :

$$Y_t = \frac{(1 - a_{22}L)\varepsilon_{1t} + a_{12}L\varepsilon_{2t}}{(1 - a_{11}L)(1 - a_{22}L) - a_{12}a_{21}L^2} \quad (5)$$

$$X_t = \frac{a_{21}L\varepsilon_{1t} + (1 - a_{11}L)\varepsilon_{2t}}{(1 - a_{11}L)(1 - a_{22}L) - a_{12}a_{21}L^2} \quad (6)$$

Fazendo  $(1 - a_{11}L)(1 - a_{22}L) - a_{12}a_{21}L^2 = 0$  e resolvendo esta equação característica inversa em ordem a  $L$ , obtém-se duas raízes características inversas. Se fizermos  $\lambda = 1/L$ , podemos escrever a equação característica (como alternativa à equação característica inversa) na forma

$$\lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0 \quad (7)$$

Como as variáveis têm a mesma equação característica, as raízes características da equação (7),  $\lambda_1$  e  $\lambda_2$ , determinam a comportamento temporal de ambas as variáveis.

# COINTEGRAÇÃO E MODELOS VEC

Para assegurar que as variáveis sejam  $I(1,1)$ , uma das raízes características tem que ser igual a 1 e a outra inferior a um em valor absoluto.

Para conseguir que a maior das duas raízes seja igual a 1, basta fazer

$$\frac{(a_{11} + a_{22}) + \sqrt{(a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21})}}{2} = 1$$

ou

$$\frac{1}{2}(a_{11} + a_{22}) + \frac{1}{2}\sqrt{a_{11}^2 + a_{22}^2 - 2a_{11}a_{22} + 4a_{12}a_{21}} = 1$$

Depois de algumas simplificações, pode mostrar-se que

$$a_{11} = \frac{(1 - a_{22}) - a_{12}a_{21}}{1 - a_{22}}$$



# COINTEGRAÇÃO E MODELOS VEC

Em relação à segunda raiz característica, como  $a_{12}$  e/ou  $a_{21}$  têm que ser diferentes de zero se as variáveis forem cointegradas, a condição  $|\lambda_2| < 1$  requer que

$$a_{22} > -1$$

e

$$a_{12}a_{21} + a_{22}^2 < 1$$

Em resumo, as principais **implicações do modelo de correcção do erro** são as seguintes:

# COINTEGRAÇÃO E MODELOS VEC

1. As restrições impostas sobre os parâmetros para assegurarem que as variáveis sejam cointegradas de ordem (1,1) garantem que um modelo de correcção do erro existe. No nosso exemplo,  $Y_t$  e  $X_t$  são processos que contem uma raiz unitária, mas a combinação linear  $Y_{t-1} - \beta X_{t-1}$  é estacionária; o vector cointegrante normalizado é  $[1, -(1-a_{22})/a_{21}]$ ; as variáveis admitem uma representação de modelo de correcção do erro com coeficientes de velocidade de ajustamento  $\alpha_1 = -a_{12}a_{21}/(1-a_{22})$  e  $\alpha_2 = a_{21}$ . Por outro lado, o modelo de correcção do erro para variáveis  $I(1)$  implica necessariamente cointegração (**teorema de representação de Granger**).

# COINTEGRAÇÃO E MODELOS VEC

2. Um sistema cointegrado pode ser visto como um modelo VAR com restrições sobre os seus coeficientes. Seja o modelo VAR

$$\begin{bmatrix} \Delta Y_t \\ \Delta X_t \end{bmatrix} = \begin{bmatrix} a_{11} - 1 & a_{12} \\ a_{21} & a_{22} - 1 \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

ou, em notação abreviada,

$$\Delta Y_t = \Pi X_{t-1} + \varepsilon_t$$

Não é apropriado a estimação de um modelo VAR de variáveis cointegradas usando apenas as primeiras diferenças, isto é,  $\Delta Y_t = \varepsilon_t$ . A não inclusão do termo  $\Pi X_{t-1}$  vai eliminar o mecanismo de correcção do erro do modelo.

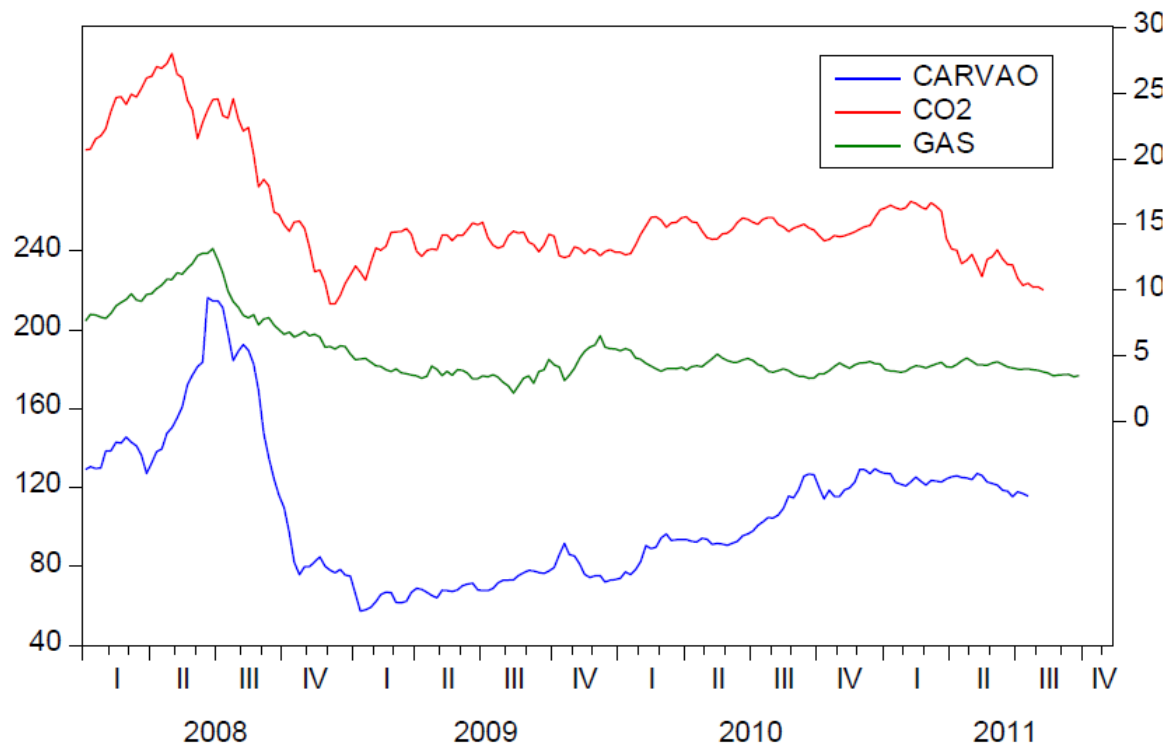
# COINTEGRAÇÃO E MODELOS VEC

3. Interpretação do **conceito de causalidade à Granger num sistema cointegrado** em que uma variável é exogenamente fraca. Num sistema cointegrado,  $Y_t$  não causa à Granger  $X_t$  se os valores desfasados  $\Delta Y_{t-i}$  não entrarem na equação de  $\Delta X_t$  e se  $X_t$  não responder aos desvios em relação ao equilíbrio de longo prazo. Nestas condições, diz-se que  $X_t$  é uma variável **exogenamente fraca**.

# COINTEGRAÇÃO E MODELOS VEC

## Exemplo de aplicação:

Preços do carvão, gás e CO2 entre 2/1/2008 e 11/11/2011 (dados semanais obtidos através das médias das observações diárias).



Séries semanais dos preços do Gás, CO2 e Carvão

# COINTEGRAÇÃO E MODELOS VEC

Vector Error Correction Estimates			
Date: 11/08/12 Time: 16:49			
Sample (adjusted): 23/01/2008 20/07/2011			
Included observations: 183 after adjustments			
Standard errors in ( ) & t-statistics in [ ]			
Cointegrating Eq:	CointEq1		
CARVAO(-1)	1.000000		
CO2(-1)	-20.58283 (4.42290) [-4.65370]		
GAS(-1)	25.42139 (7.68378) [ 3.30845]		
C	79.52030		
Error Correction:	D(CARVAO)	D(CO2)	D(GAS)
CointEq1	-0.011715 (0.00710) [-1.65109]	0.000396 (0.00111) [ 0.35529]	-0.002121 (0.00052) [-4.05026]
D(CARVAO(-1))	0.373780 (0.07419) [ 5.03837]	0.017811 (0.01165) [ 1.52892]	0.009373 (0.00547) [ 1.71223]
D(CARVAO(-2))	0.062693 (0.07204) [ 0.87021]	0.000520 (0.01131) [ 0.04596]	-0.005615 (0.00532) [-1.05619]

Estimação do modelo de correcção do erro VEC(2) tri-dimensional (Carvão, CO2 e Gás)

# COINTEGRAÇÃO E MODELOS VEC

Roots of Characteristic Polynomial	
Endogenous variables: CARVAO CO2 GAS	
Exogenous variables:	
Lag specification: 1 2	
Date: 11/08/12 Time: 16:51	
Root	Modulus
1.000000	1.000000
1.000000	1.000000
0.864514	0.864514
0.592157 - 0.159649i	0.613301
0.592157 + 0.159649i	0.613301
0.095812 - 0.292506i	0.307798
0.095812 + 0.292506i	0.307798
-0.232137 - 0.086937i	0.247882
-0.232137 + 0.086937i	0.247882
VEC specification imposes 2 unit root(s).	

Raízes características do VEC(2)

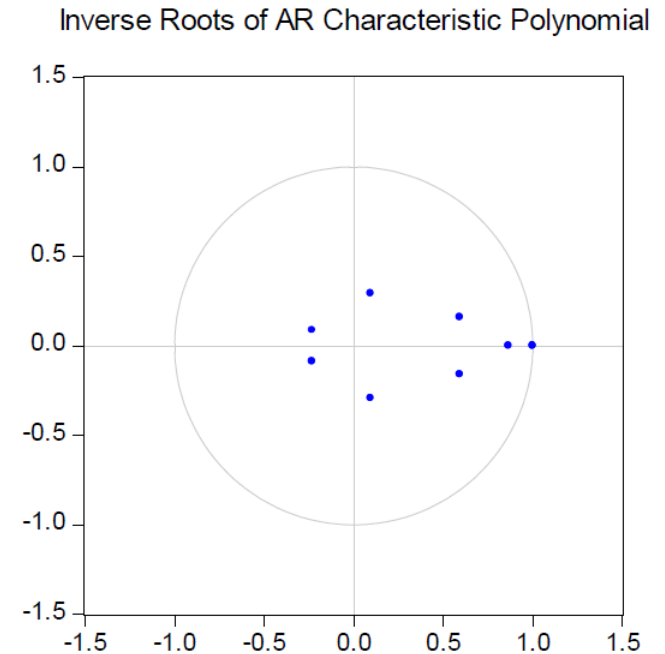


Gráfico das raízes características do VEC(2)

# COINTEGRAÇÃO E MODELOS VEC

## VEC Granger Causality/Block Exogeneity Wald Tests

Date: 11/08/12 Time: 16:52

Sample: 2/01/2008 9/11/2011

Included observations: 183

Dependent variable: D(CARVAO)

Excluded	Chi-sq	df	Prob.
D(CO2)	7.091115	2	0.0289
D(GAS)	7.683081	2	0.0215
All	14.61865	4	0.0056

Dependent variable: D(CO2)

Excluded	Chi-sq	df	Prob.
D(CARVAO)	2.938263	2	0.2301
D(GAS)	0.140415	2	0.9322
All	2.939782	4	0.5680

Dependent variable: D(GAS)

Excluded	Chi-sq	df	Prob.
D(CARVAO)	3.059678	2	0.2166
D(CO2)	12.57908	2	0.0019
All	13.59210	4	0.0087

Teste de causalidade à Granger/exogeneidade do VEC(2)



# COINTEGRAÇÃO E MODELOS VEC

A metodologia de Engle-Granger para testar a cointegração goza de fácil implementação mas tem importantes limitações:

- Os testes de raízes unitárias aos resíduos das regressões de cointegração são equivalentes apenas assintoticamente;
- O método de Engle-Granger não consegue estimar separadamente múltiplos vectores cointegrantes (em sistemas com 3 ou mais variáveis);
- Sendo o método de Engle-Granger um processo de estimação bi-etápico, em que as estimativas da equação da 2ª etapa utilizam os resíduos obtidos na equação da 1ª etapa, qualquer erro introduzido na 1ª etapa passa para a 2ª etapa.

Para ultrapassar estas limitações, Johansen (1988) introduziu uma metodologia de análise de cointegração baseada na característica da matriz dos parâmetros do modelo VAR e suas raízes características, que permite testar todos os vectores cointegrantes para um determinado conjunto de variáveis.

# COINTEGRAÇÃO E MODELOS VEC

Para ilustrar melhor estes conceitos, considere-se um modelo VAR(1) com  $k$  variáveis

$$Y_t = A_1 Y_{t-1} + \varepsilon_t, \quad (1)$$

ou

$$\begin{aligned} \Delta Y_t &= A_1 Y_{t-1} - Y_{t-1} + \varepsilon_t \\ &= (A_1 - I) Y_{t-1} + \varepsilon_t \\ &= \Pi Y_{t-1} + \varepsilon_t, \end{aligned} \quad (2)$$

onde  $\Pi = A_1 - I$  é uma matriz  $k \times k$  e  $\pi_{js}$  representa o elemento da  $j$ -ésima linha e  $s$ -ésima coluna da matriz  $\Pi$ . Como se verá de seguida, a característica de  $\Pi = A_1 - I$  dá-nos o número de vectores cointegrantes.

# COINTEGRAÇÃO E MODELOS VEC

Se a **característica da matriz  $\Pi$  for zero**,  $rank(\Pi) = 0$ , então todos os elementos de  $\Pi$  são nulos. Neste caso, a relação (2) é equivalente a ter um modelo VAR em primeiras diferenças,

$$\Delta Y_t = \varepsilon_t, \quad (3)$$

onde a primeira diferença de cada uma das variáveis no vector  $Y_t$  é estacionária ou integrada de ordem 0,  $\Delta Y_{it} \sim I(0)$ . E assim, cada uma das variáveis  $Y_{it}$  contém uma raiz unitária ( $Y_{it} = Y_{i,t-1} + \varepsilon_{it}$ ) e não existe nenhuma combinação das variáveis que seja estacionária.

Se a **característica da matriz  $\Pi$  for completa (“full rank”)**,  $rank(\Pi) = k$ , a solução de longo prazo da relação (2) é dada pelas  $k$  equações independentes

$$\begin{aligned} \pi_{11}Y_{1t} + \pi_{12}Y_{2t} + \cdots + \pi_{1k}Y_{kt} &= 0 \\ \pi_{21}Y_{1t} + \pi_{22}Y_{2t} + \cdots + \pi_{2k}Y_{kt} &= 0 \\ \vdots & \\ \pi_{k1}Y_{1t} + \pi_{k2}Y_{2t} + \cdots + \pi_{kk}Y_{kt} &= 0 \end{aligned} \quad (4)$$

Cada uma destas  $k$  equações é uma restrição independente da solução de longo-prazo. Como todas as variáveis são estacionárias, estas não podem ser cointegradas.

# COINTEGRAÇÃO E MODELOS VEC

Nos casos intermédios, em que a **matriz de coeficientes  $\Pi$  tem característica reduzida  $r < k$** ,  $rank(\Pi) = r < k$ , existem  $r$  vectores cointegrantes (ou  $r$  equações independentes) e  $k - r$  tendências estocásticas no sistema. Se  $r = 1$ , existe um único vector cointegrante dado por qualquer linha da matriz de coeficientes  $\Pi$ . Neste caso, podemos escrever cada termo  $\Delta Y_{it}$  na forma de correcção do erro. Por exemplo,  $\Delta Y_{1t}$  pode escrever-se como

$$\Delta Y_{1t} = \pi_{11} Y_{1,t-1} + \pi_{12} Y_{2,t-1} + \dots + \pi_{1k} Y_{k,t-1} + \varepsilon_{1t}$$

ou, normalizando em relação a  $Y_{1,t-1}$ , e fazendo  $\alpha_1 = \pi_{11}$  e  $\beta_{1j} = \pi_{1j} / \pi_{11}$ ,

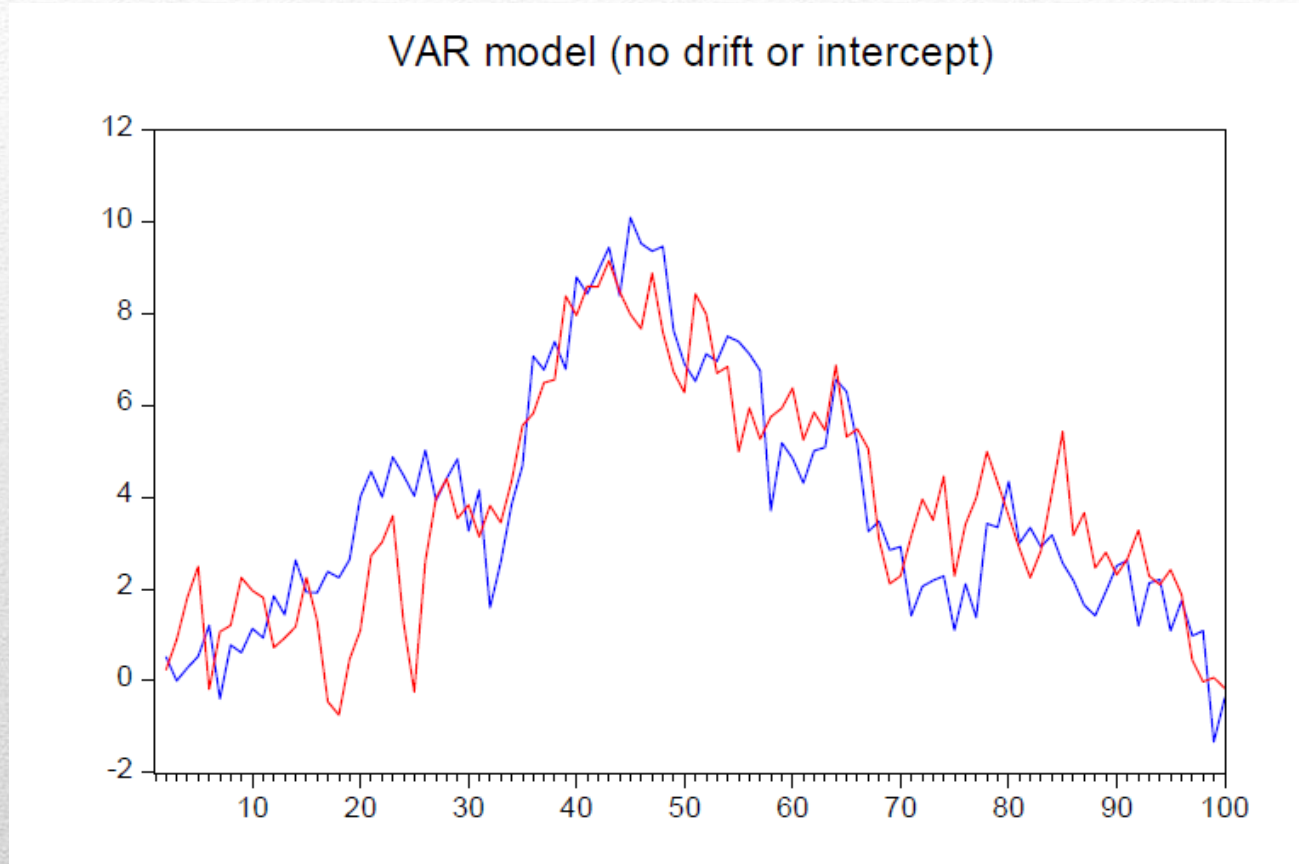
$$\Delta Y_{1t} = \alpha_1 (Y_{1,t-1} + \beta_{12} Y_{2,t-1} + \dots + \beta_{1k} Y_{k,t-1}) + \varepsilon_{1t}$$

No longo prazo, a variável  $Y_{it}$  satisfaz a relação

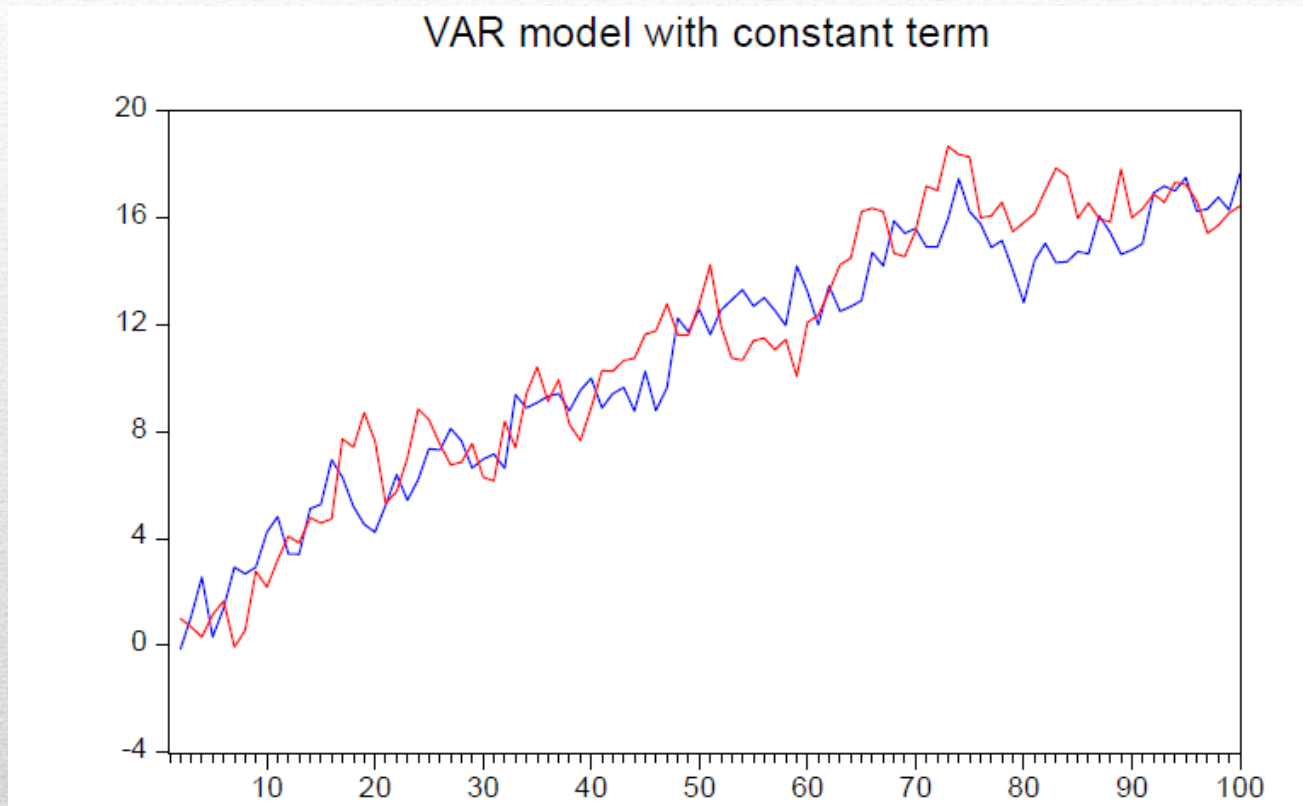
$$Y_{1,t-1} + \beta_{12} Y_{2,t-1} + \dots + \beta_{1k} Y_{k,t-1} = 0$$

Assim,  $[1, \beta_{12}, \beta_{13}, \dots, \beta_{1k}]$  é o vector cointegrante normalizado e  $\alpha_1$  é o parâmetro da velocidade de ajustamento.

# COINTEGRAÇÃO E MODELOS VEC



# COINTEGRAÇÃO E MODELOS VEC



# COINTEGRAÇÃO E MODELOS VEC

Segundo a **metodologia de Johansen** (1991,1995), testar a cointegração entre um grupo de variáveis não estacionárias consiste nos seguintes passos.

Considere-se o modelo VAR de ordem  $p$ :

$$Y_t = A_1 Y_{t-1} + \dots + A_p Y_{t-p} + \varepsilon_t, \quad (7)$$

onde:  $Y_t$  é um vector de  $k$  variáveis não estacionárias ou integradas de ordem 1 ou  $I(1)$ ; e  $\varepsilon_t$  é um vector de perturbações aleatórias.

Este modelo pode escrever-se na seguinte forma, adicionado e subtraindo ao segundo membro o termo  $A_p Y_{t-p+1}$ :

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + (A_{p-1} + A_p) Y_{t-p+1} - A_p \Delta Y_{t-p+1} + \varepsilon_t. \quad (8)$$

Se agora adicionarmos e subtrairmos o termo  $(A_{p-1} + A_p) Y_{t-p+2}$ , resulta

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \dots - (A_{p-1} + A_p) \Delta Y_{t-p+2} - A_p \Delta Y_{t-p+1} + \varepsilon_t. \quad (9)$$

Seguindo o mesmo procedimento com os sucessivos termos, obtém-se

$$\Delta Y_t = \Pi Y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + \varepsilon_t, \quad (10)$$

onde  $\Pi = \sum_{i=1}^p A_i - I$ ;  $\Gamma_i = -\sum_{j=i+1}^p A_j$ .

# COINTEGRAÇÃO E MODELOS VEC

Segundo o **teorema de representação de Granger**, se a matriz de coeficientes  $\Pi$  tiver característica reduzida  $r < k$ , então existem duas matrizes  $\alpha$  e  $\beta$ , cada uma com característica  $r$ , tal que:

$$\Pi = \alpha\beta'$$

e

$$\beta'Y_t \sim I(0)$$

onde:

$\alpha$  é a matriz  $k \times r$  dos **coeficientes da velocidade de ajustamento**;

$\beta$  é a matriz  $k \times r$  dos **coeficientes de longo-prazo**;

$r$  é o **número de relações de cointegração**;

cada coluna de  $\beta$  é um **vector cointegrante**.



# COINTEGRAÇÃO E MODELOS VEC

O **método de Johansen** (Johansen, 1988) consiste em estimar a matriz de coeficientes  $\Pi$  a partir de um modelo VAR sem restrições e testar se as restrições impostas pela característica da matriz  $\Pi$  podem ser rejeitadas.

Suponha que se obtém a matriz  $\Pi$  e se ordena as  $k$  raízes características, tal que  $\lambda_1 > \lambda_2 > \dots > \lambda_k$ :

- Se as variáveis de  $Y_t$  forem não cointegradas, a característica da matriz  $\Pi$  é zero,  $rank(\Pi) = r = 0$ , e todas as suas raízes características serão nulas,  $\lambda_i = 0$  para  $i = 1, 2, \dots, k$ . O mesmo será dizer que  $\ln(1 - \lambda_i) = 0$  para  $i = 1, 2, \dots, k$ .
- Se a característica da matriz  $\Pi$  é um,  $rank(\Pi) = r = 1$ , a maior raiz característica assume um valor no intervalo  $0 < \lambda_1 < 1$  e  $\ln(1 - \lambda_1) < 0$ , enquanto que as outras raízes características são nulas,  $\lambda_i = 0$  para  $i = 2, \dots, k$ , assim como  $\ln(1 - \lambda_i) = 0$  para  $i = 2, \dots, k$ .

# COINTEGRAÇÃO E MODELOS VEC

## Teste do traço

Hipótese nula ( $H_0$ ): O número de vetores cointegrantes distintos é menor ou igual a  $r$  (característica da matriz  $\Pi$ ) contra uma alternativa genérica.

Estatística do teste:  $\lambda_{trace}(r) = -T \sum_{i=r+1}^k \ln(1 - \hat{\lambda}_i),$

onde  $\hat{\lambda}_i$  é o valor estimado da raiz característica (ou valor próprio) da matriz estimada de  $\Pi$  e  $T$  é o número de observações utilizadas.

Valores críticos: São obtidos através de um método de simulação de Monte Carlos (Osterwald-Lenum, 1992).

# COINTEGRAÇÃO E MODELOS VEC

## Teste do máximo

Hipótese nula ( $H_0$ ): O número de vetores cointegrantes é menor ou igual a  $r$  contra a alternativa de  $r+1$  vetores cointegrantes.

Estatística do teste:  $\lambda_{\max}(r, r+1) = -T \ln(1 - \hat{\lambda}_{r+1})$

onde  $\hat{\lambda}_i$  é o valor estimado da raiz característica (ou valor próprio) da matriz estimada de  $\Pi$  e  $T$  é o número de observações utilizadas.

Valores críticos: São obtidos através de um método de simulação de Monte Carlo (Osterwald-Lenum, 1992).

**Nota:** a maioria das tabelas estatísticas (Osterwald-Lenum, 1992), fornecem os valores das distribuições empíricas das estatísticas  $\lambda_{\text{trace}}(r)$  e  $\lambda_{\max}(r, r+1)$  para  $k-r$  e dimensões de teste 1%, 2.5%, 5% e 10%. Os resultados dos testes do traço e do máximo podem não ser coincidentes.

# COINTEGRAÇÃO E MODELOS VEC

O EViews considera as seguintes especificações dos testes de cointegração de Johansen (1995, p. 80-84):

1)  $\Pi Y_{t-1} = \alpha \beta' Y_{t-1}$  (Dados das séries sem tendência determinística e equações de cointegração sem constante ou tendência)

2)  $\Pi Y_{t-1} = \alpha(\beta' Y_{t-1} + \rho_0)$  (Dados das séries sem tendência determinística e equações de cointegração com constante)

3)  $\Pi Y_{t-1} = \alpha(\beta' Y_{t-1} + \rho_0) + \alpha_{\perp} \gamma_0$  (Dados das séries com tendência linear determinística e equações de cointegração com constante)

4)  $\Pi Y_{t-1} = \alpha(\beta' Y_{t-1} + \rho_0 + \rho_1 t) + \alpha_{\perp} \gamma_0$  (Dados das séries com tendência linear determinística e equações de cointegração com constante e tendência)

5)  $\Pi Y_{t-1} = \alpha(\beta' Y_{t-1} + \rho_0 + \rho_1 t) + \alpha_{\perp}(\gamma_0 + \gamma_1 t)$  (Dados das séries com tendência quadrática determinística e equações de cointegração com constante e tendência)

Nota: Os termos associados com  $\alpha_{\perp}$  são termos determinísticos obtidos fora das relações de cointegração.

# COINTEGRAÇÃO E MODELOS VEC

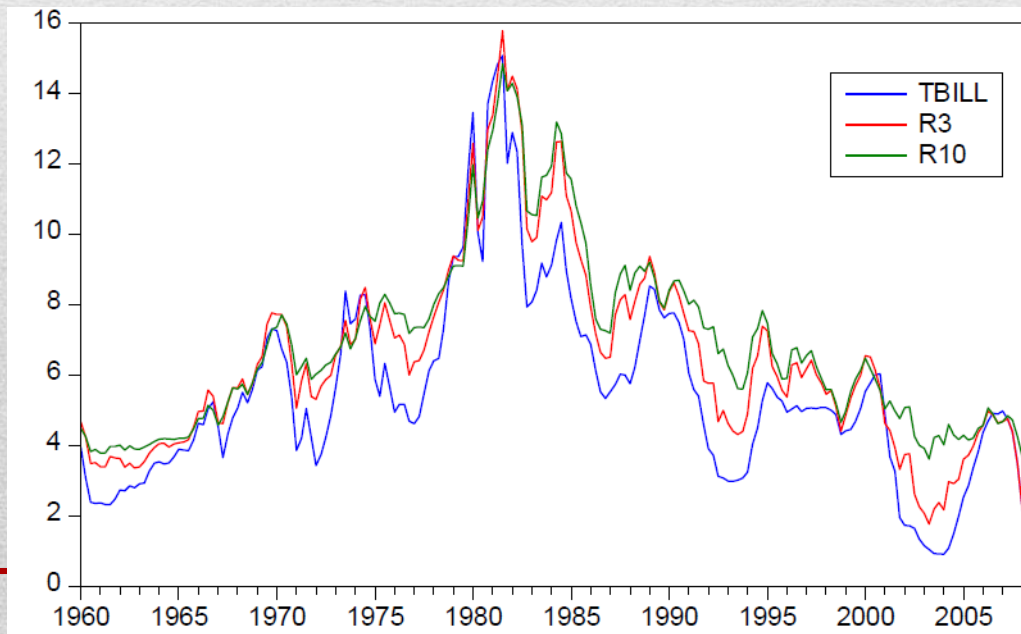
## Exemplo de aplicação:

Considere a base de dados macroeconómica dos Estados Unidos (Quarterly.xls ou Quarterly\_US.wk1) entre 1960Q1 e 2008Q1, na qual se pretende analisar as relações de cointegração entre as variáveis:

TBILL = taxa de juro de bilhetes do tesouro a 3 meses;

R3 = taxa de juro de bilhetes do tesouro a 3 anos;

R10 = taxa de juro de bilhetes do tesouro a 10 anos.



# COINTEGRAÇÃO E MODELOS VEC

Date: 23/08/12 Time: 16:36  
 Sample (adjusted): 1962Q1 2008Q1  
 Included observations: 185 after adjustments  
 Trend assumption: No deterministic trend (restricted constant)  
 Series: TBILL R3 R10  
 Lags interval (in first differences): 1 to 7

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.182744	52.59629	35.19275	0.0003
At most 1	0.069478	15.26272	20.26184	0.2117
At most 2	0.010437	1.940909	9.164546	0.7894

Trace test indicates 1 cointegrating eqn(s) at the 0.05 level  
 \* denotes rejection of the hypothesis at the 0.05 level  
 \*\*MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.182744	37.33357	22.29962	0.0002
At most 1	0.069478	13.32181	15.89210	0.1216
At most 2	0.010437	1.940909	9.164546	0.7894

Max-eigenvalue test indicates 1 cointegrating eqn(s) at the 0.05 level  
 \* denotes rejection of the hypothesis at the 0.05 level  
 \*\*MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegrating Coefficients (normalized by b\*S11\*b=I):

	TBILL	R3	R10	C

# COINTEGRAÇÃO E MODELOS VEC

	2.567876	-2.495452	0.169793	0.855835
	2.289363	-5.442669	3.579285	-2.282104
	0.736026	-1.602675	0.551175	2.538829
Unrestricted Adjustment Coefficients (alpha):				
D(TBILL)	-0.038389	-0.014653	0.061863	
D(R3)	0.049346	0.056195	0.052376	
D(R10)	0.093208	0.034052	0.035700	
1 Cointegrating Equation(s):      Log likelihood      -77.55777				
Normalized cointegrating coefficients (standard error in parentheses)				
TBILL	R3	R10	C	
1.000000	-0.971796	0.066122	0.333285	
	(0.15880)	(0.16667)	(0.23056)	
Adjustment coefficients (standard error in parentheses)				
D(TBILL)	-0.098579			
	(0.12339)			
D(R3)	0.126713			
	(0.11363)			
D(R10)	0.239348			
	(0.08478)			
2 Cointegrating Equation(s):      Log likelihood      -70.89686				
Normalized cointegrating coefficients (standard error in parentheses)				
TBILL	R3	R10	C	
1.000000	0.000000	-0.969104	1.252908	
		(0.05942)	(0.44046)	
0.000000	1.000000	-1.065271	0.946313	
		(0.04275)	(0.31691)	
Adjustment coefficients (standard error in parentheses)				
D(TBILL)	-0.132125	0.175551		
	(0.16527)	(0.28763)		
D(R3)	0.255364	-0.428990		
	(0.15148)	(0.26364)		
D(R10)	0.317305	-0.417930		
	(0.11320)	(0.19703)		