Macroeconometria 2

Mestrado em Economia Monetária e Financeira

Mestrado em Economia

ISCTE-IUL, Dep. de Economia

MODELOS VAR, TESTES DE RAIZES UNITÁRIAS E COINTEGRAÇÃO

Luís Filipe Martins

luis.martins@iscte.pt

http://iscte.pt/~lfsm

Departamento de Métodos Quantitativos,

ISCTE-IUL, Escola de Gestão

Lisboa, Janeiro de 2010

1 Motivação

- Evidência Empirica (*stylized facts*): Algumas series economicas não têm média constante (não estacionaridade); podem apresentar um tendência deterministica (trend); não ter um mecanismo de reversão para a média e a dependencia temporal ser muito persistente (não estacionaridade e memória longa); haver co-movimentos entre séries não estacionárias (cointegração).

Example 1 *AR multivariado: (real log) Stock prices e (log) Dividendos; (log) Dividend-price ratio e (real log) dividend growth rate; (log) Dividend-price ratio, (real log) dividend growth rate e price-moving average earnings ratio; real stock return, log dividend-price ratio e level of the stochastically detrended short-term interest rate; ...*

Example 2 *Cointegração entre Preços e Dividendos (combinação linear entre eles, $P_t - \frac{D_t}{R}$, é estacionária (R =discount rate)); Cointegração entre taxas de juro (e yields) de diferentes maturidades; moeda, preços e rendimento; PPP; ...*

2 Vector Autoregressive Models (VAR)

- Lucas e Sims: (1) Agentes com expectativas forward-looking (parâmetros expectacionais) (2) Parâmetros estruturais obtidos a partir de problemas de optimização intertemporal dos agentes (micro-fundamentos).

- Sims: Praticamente não existem variáveis exógenas num contexto de agentes forward-looking. SOLUÇÃO: Modelos não estruturais (evita equações simultâneas) - VAR - para fornecer evidência empírica sobre as respostas das variáveis a acções de política económico-financeira, sem impôr restrições teóricas e tendo em conta a possível endogeneidade dos instrumentos.

- $AR(p)$ $k-$dimensional: $VAR(p) : t = 1, ..., T$

$$\begin{aligned} y_t &= c + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + ... + \Phi_p y_{t-p} + \varepsilon_t \Leftrightarrow && (1) \\ \Phi(L) y_t &= c + \varepsilon_t, \Phi(L) = I_k - \Phi_1 L - \Phi_2 L^2 - ... - \Phi_p L^p. && (2) \end{aligned}$$

$$\begin{pmatrix} y_{1t} \\ ... \\ y_{kt} \end{pmatrix} = \begin{pmatrix} c_1 \\ ... \\ c_k \end{pmatrix} + \begin{pmatrix} \Phi_{11}^{(1)} & ... & \Phi_{1k}^{(1)} \\ ... & ... & ... \\ \Phi_{k1}^{(1)} & ... & \Phi_{kk}^{(1)} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ ... \\ y_{kt-1} \end{pmatrix} + ... + \begin{pmatrix} \Phi_{11}^{(p)} & ... & \Phi_{1k}^{(p)} \\ ... & ... & ... \\ \Phi_{k1}^{(p)} & ... & \Phi_{kk}^{(p)} \end{pmatrix} \begin{pmatrix} y_{1t-p} \\ ... \\ y_{kt-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ ... \\ \varepsilon_{kt} \end{pmatrix}$$

$$\begin{aligned} y_{it} &= c_i + \Phi_{i1}^{(1)} y_{1t-1} + ... + \Phi_{ik}^{(1)} y_{kt-1} + ... \\ &\quad +\Phi_{i1}^{(p)} y_{1t-1} + ... + \Phi_{ik}^{(p)} y_{kt-1} + \varepsilon_{it}. \end{aligned}$$

- $\varepsilon_t \sim w.n._k$ com $E(\varepsilon_t) = \mathbf{0}_{k\times 1}$,$E(\varepsilon_t\varepsilon_t') = \Omega_{k\times k}$, $E(\varepsilon_t\varepsilon_s') = \mathbf{0}_{k\times k}, t \neq s, \Omega$ SPD.
- Estacionaridade: Raízes de $\left|I_k - \Phi_1 z - \Phi_2 z^2 - ... - \Phi_p z^p\right|$ fora do círculo unitário. Neste caso, pelo Wold Representation Theorem, $VMA\left(\infty\right) : y_t = \mu + C(L)\varepsilon_t, C(L) = \Phi(L)^{-1}$.

2.1 Estimação, Inferência e Previsão

- OLS (=MLE) equação a equação porque cada uma das k equações têm os mesmos regressores e estes são ortogonais em relação aos erros (não estão contemporaneamente correlacionados), onde $\widehat{\Omega} = \frac{1}{T}\sum_{t=1}^{T} \widehat{\varepsilon}_t\widehat{\varepsilon}_t'$ ou $\widehat{\Omega} = \frac{1}{T-kp-1}\sum_{t=1}^{T} \widehat{\varepsilon}_t\widehat{\varepsilon}_t'$. Inferência: Ver OLS/GLS e MLE!
- Escolha do desconhecido $p : SBC$ ou

$$LR = [T - k(p+q)]\left(\log\left|\widehat{\Omega}_{p+q}^{-1}\right| - \log\left|\widehat{\Omega}_{p}^{-1}\right|\right) \tag{3}$$

  com distribuição assimptótica $\chi^2_{(k^2q)}$, para $VAR\left(p\right)$ versus $VAR\left(p+q\right), q = 1, 2, ...$.
- MLE condicional (teste LR) ...
- Previsão: Semelhante a AR (ver ARIMA) ...

2.2 Causalidade à Granger

- Granger causality: Questionar relações de causalidade estatística entre as variáveis.

- Num $VAR(1)$ bi-dimensional,

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}, \quad (4)$$

se y_t não prever x_{t+h}, então y não causa à Granger x, para $\phi_{12} = 0$. Mais geralmente, y não causa x se o erro quadrático médio de previsão a h passos para x_{t+h} baseado apenas em $(x_t, x_{t-1}, ...)$ for o mesmo que o MSE de x_{t+h} com $[(x_t, x_{t-1}, ...), (y_t, y_{t-1}, ...)] \Leftrightarrow \Phi_{12}^{(j)}, j = 1, ..., p$.

2.3 Funções impulso-resposta,FIR

- Como é que, num dado sistema VAR, as variáveis endógenas respondem, em termos dinâmicos, a choques exógenos, e durante quanto tempo? O choque na inovação de $x, \varepsilon_{x,t-h} = 0, \varepsilon_{x,t} = 1, \varepsilon_{x,t+h} = 0, h = 1, 2, ...$ provoca que resposta em y em $t+1, t+2, ..., ?$

- No $AR(1)$ estacionário, $y_t = \phi y_{t-1} + \varepsilon_t = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}$,

$$\begin{array}{llllll} \varepsilon_t : ...0 & 0 & 1 & 0 & 0 & 0... \\ x_t : ...0 & 0 & 1 & \phi & \phi^2 & \phi^3... \end{array}$$

- $VAR(p)$ estacionário,

$$y_t = C(L)\varepsilon_t, C(L) = C_0 + C_1 L + C_2 L^2 + ..., \text{ onde } C_0 = I_k, \quad (5)$$

o elemento genérico

$$c_{ij,h} = \frac{\partial y_{i,t}}{\partial \varepsilon_{j,t-h}} = \frac{\partial y_{i,t+h}}{\partial \varepsilon_{j,t}}$$

é a chamada FIR (em ordem ao lag h) de $\varepsilon_{j,t}$ sobre $y_{i,t}$.

- A FIR obtém-se por simples simulação: A partir de $y_t = \Phi_1 y_{t-1} + ... + \Phi_p y_{t-p} + \varepsilon_t$, fazendo $y_{t-1} = ... = y_{t-p} = 0$ e $\varepsilon_{jt} = 1$, com todos os outros $\varepsilon_{it} = 0$, simula-se o sistema para $t, t+1, ..., t+h$. O valor de y_{t+h} corresponde à coluna j de C_h (ver, por exemplo, com o VAR(1) bivariado).

- Se a FIR de x em relação à inovação de y for sempre 0, então y não causa x.

2.4 Decomposição de Variâncias

- Quais os choques que são a principal causa para a variabilidade das variáveis endógenas do sistema.

- O erro de previsão a h passos é dado por

$$\begin{aligned} e_{t+h|t} &= y_{t+h} - y_{t+h|t} = \sum_{j=0}^{h-1} C_j \varepsilon_{t+h-j}; & (6) \\ y_{t+h|t} &= E(y_{t+h} | \varepsilon_t, \varepsilon_{t-1}, ...). & (7) \end{aligned}$$

Portanto, como $\Omega = diag\left(\sigma_1^2, ..., \sigma_k^2\right)$, a variância do $i-$ésimo elemento de $e_{t+h|t}$ é $\sum_{j=1}^{k} \left(\sigma_j^2 \sum_{l=0}^{h-1} c_{ij,l}^2\right)$.

- A fracção da variância do erro de previsão a h passos de $y_{i,t}$

atribuível a $\varepsilon_{j,t}$ é dada por

$$V_{ij,h} = \frac{\sigma_j^2 \sum_{l=0}^{h-1} c_{ij,l}^2}{\sum_{m=1}^{k} \left(\sigma_m^2 \sum_{l=0}^{h-1} c_{im,l}^2 \right)},$$

calculando-se k valores para $V_{ij,h} : V_{i1,h}, ..., V_{ik,h}$.

3 Testes de Raízes Unitárias

- Como se viu (c.f. apontamentos sobre modelos ARIMA), é importante distinguir entre processos $I(0)$ e $I(1)$. Isso é feito em termos formais através de testes de raízes unitárias ou testes de (não)estacionaridade.
  (a) Regressão entre variáveis $I(1)$ pode ser espúrias (var. indep. mas cuja relação vem significativa)
  (b) Variáveis $I(1)$ podem estar cointegradas
  (c) Distribuição do estimador e testes t não são standard para var. $I(1)$.

3.1 Testes de Dickey-Fuller

- AR(1): $y_t = \rho y_{t-1} + \varepsilon_t, \varepsilon_t \sim wn\left(0, \sigma_\varepsilon^2\right)$. Se $\rho = 1$, então y_t é $I(1)$, ao passo que se $|\rho| < 1$, y_t será $I(0)$. Sabe-se que para $\widehat{\rho} = \frac{\sum y_t y_{t-1}}{\sum y_{t-1}^2}$, $\sqrt{T}\left(\widehat{\rho} - \rho\right) \xrightarrow{d} N\left(0, \left(1 - \rho^2\right)\right)$ que é degenerada quando $\rho = 1$. Neste caso, a normalização é $T\left(\widehat{\rho} - 1\right)$ e o estimador diz-se superconsistente.

- $\triangle y_t = \alpha y_{t-1} + \varepsilon_t$, em que $\alpha = \rho - 1$. DF t-stat para $H_0 : I(1), \alpha = 0$ vs $H_1 : I(0), \alpha < 0$,é

$$\tau_{nc} = T\widehat{\alpha}\frac{\sqrt{(1/T^2)\sum y_{t-1}^2}}{\sqrt{SSR/(T-1)}}, SSR = \sum(\triangle y_t - \widehat{\alpha}y_{t-1})^2. \quad (8)$$

  A distribuição não é a normal (é assimétrica com left fat tail) e a RC está do lado esquerdo.

- A distribuição é outra quando se inclui o termo independente $\triangle y_t = c + \alpha y_{t-1} + \varepsilon_t$. A estatistica DF t-stat para $H_0 : I(1)$ vs $H_1 : I(0)$ é a τ_c.
- No teste anterior, sob H_0, o processo y_t é um passeio aleatório com deriva (DSP porque torna-se estacionário às primeiras diferenças): $\triangle y_t = c + \varepsilon_t$. Isto é equivalente a $y_t = y_0 + ct + \sum_{j=1}^{t} \varepsilon_j$. Portanto, o teste τ_c não tem potência contra uma alternativa TSP (estacionário em torno da tendência) $y_t = \mu + ct + u_t$. Os modelos DSP e TSP são non-nested. Solução:
- Modelo genérico

$$
\begin{aligned}
y_t &= \gamma_0 + \gamma_1 t + z_t, z_t = \rho z_{t-1} + \varepsilon_t \Leftrightarrow \\
y_t &= \underbrace{[(1-\rho)\gamma_0 + \rho\gamma_1]}_{\mu} + \underbrace{\gamma_1(1-\rho)}_{\delta} t + \rho y_{t-1} + \varepsilon_t \Leftrightarrow \\
\triangle y_t &= \mu + \delta t + \alpha y_{t-1} + \varepsilon_t.
\end{aligned}
$$

Teste DF t-stat τ_{ct} para $H_0 : DSP$ (passeio aleatório com deriva) vs $H_1 : TSP$ (trend stationary). Se $|\rho| < 1$, teremos um TSP, se $\rho = 1$ resulta um DSP.

3.2 Testes ADF

- Erros ε_t não são wn pois estão autocorrelacionados. Então, para um $AR(p)$, a equação geral de teste é

$$\triangle y_t = \mu + \delta t + \alpha y_{t-1} + \delta_1 \triangle y_{t-1} + ... + \delta_p \triangle y_{t-p} + \varepsilon_t. \quad (9)$$

A distribuição dos testes $\tau_{nc}, \tau_c, \tau_{ct}$ assim aumentados

(Augmented Dickey-Fuller) permanece a mesma.

- $p =?$: AIC ou BIC
- Testes com pouca potência para T pequeno.

3.3 Testes de Phillips-Perron

- Alternativa ao ADF sem corrigir parametricamente (p) a autocorrelação dos erros: Correção de forma não-paramétrica, sugerida por Phillips e Perron (1988), em que a variância de longo prazo de ε_t é estimada a partir dos resíduos da regressão de teste.
- Z_1 corrige τ_c e Z_2 corrige τ_{ct}.
- Problemas em amostras finitas justifica o uso dos testes eficientes ADF-GLS de Elliott, Rothenberg e Stock (ERS) (1996) e os testes M de Ng e Perron (1996).

3.4 Testes de Estacionaridade (KPSS)

- É também possível construir testes para a hipótese nula de que a série é estacionária:

$$\begin{aligned} y_t &= \mu + \delta t + \xi_t + u_t, \qquad u_t \text{ estacionário} \\ \xi_t &= \xi_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim wn(0, \sigma_\varepsilon^2). \end{aligned} \tag{10}$$

- A hipótese nula de estacionaridade é $H_0 : \sigma_\varepsilon^2 = 0$ contra a alternativa de não estacionaridade $H_1 : \sigma_\varepsilon^2 > 0$. O teste LBI ("locally best invariant") é um teste na forma LM $L = \frac{\frac{1}{T^2}\sum_{t=1}^{T} S_t^2}{\hat{\sigma}_u^2}$, em que $S_t = \sum_{j=1}^{t} e_j$, sendo e_t os resíduos da regressão de y_t numa constante e em t (se $\delta \neq 0$) e $\hat{\sigma}_u^2$ uma estimativa consistente da variância de longo prazo dos erros u_t.

- $KPSS_1$ corresponde a τ_c e $KPSS_2$ a τ_{ct}, invertendo as hipóteses de teste.

4 Cointegração

- Certas teorias supôem, explícita ou implicitamente, que existem relações de equilíbrio de longo prazo entre variáveis. Por exemplo, apesar de o consumo per capita e o rendimento per capita serem não estacionários, o seu rácio tem-se mantido estável. Seria implausível que uma das séries apresente um comportamento divergente relativamente à outra. Outros exemplos: relação entre taxas de juro de curto e longo prazo (estrutura temporal), preços e salários, preços de acções e dividendos, moeda, preços e rendimento, PPP, etc.

- Objectivo da análise de cointegração: detectar e analisar relações de longo prazo (em níveis) entre variáveis de interesse. Em geral, $x_t \sim I(1), y_t \sim I(1) \Rightarrow ay_t + bx_t \sim I(1)$ e, por isso, análise de curto prazo (diferenças) $a\triangle y_t + b\triangle x_t \sim I(0)$. Excepção: Cointegração, $x_t \sim I(1), y_t \sim I(1) \Rightarrow ay_t + bx_t \sim I(0)$ com vector de cointegração $(a, b)'$ e erro de equilibrio $u_t = ay_t + bx_t$.

- DEF: y_t é um vector de k variáveis $y_t = (y_{1t}, ..., y_{kt})'$. y_t é cointegrado de ordem (d, b), $y_t \tilde{} CI(d, b)$ se (i) cada $y_{it} \sim I(d), i = 1, ..., k$; (ii) existe pelo menos um vector $\beta \neq \mathbf{0}$:$\beta' y_t = \beta_1 y_{1t} + ... + \beta_k y_{kt} = u_t \sim I(d - b)$, com $b > 0, d \geq b$.

- Notas:
(a) A norma é $d = b = 1$ mas d, b podem ser não inteiros.

(b) Normalização $\beta_1 = 1$ porque para qualquer escalar a, $a\beta' y_t \sim I(0)$.

(c) Se $k > 2$, podem existir $r \leq k-1$ vectores de cointegração linearmente independentes tal que $B' y_t \sim I(0)$, em que $B_{k\times r} = (\beta_1, ..., \beta_r)$ com $rank(B) = r$.

(d) β conhecido versus desconhecido.

4.1 Estimação

- Regressão uniequacional $y_t = \beta^{'} x_t + \gamma^{'} z_t + u_t$, em que z_t contém termos determinísticos (constante, tendência, dummies, ...) e $x_t \sim I(1), y_t \sim I(1)$.

- O OLS é superconsistente: $T\left(\widehat{\beta} - \beta\right)$ tem uma distribuição limite não degenerada; OLS enviesado em amostras pequenas; OLS não é eficiente, em especial se $r > 1$; Tal como nas regressões espúrias, em geral, a teoria assimptótica standard não é válida.

- DOLS (dynamic OLS) de Saikkonen (1991) e Stock e Watson (1993): OLS a

$$y_t = \beta^{'} x_t + \gamma^{'} z_t + \sum_{j=-p}^{p} \delta_j \triangle y_{t-j} + \varepsilon_t, \qquad (11)$$

  para eliminar os efeitos perturbadores da dinâmica de curto prazo. A inferência habitual é assimptoticamente válida. Alternativa: FM-OLS (fully-modified OLS) de Phillips and Hansen (1990) ...

4.2 Testes de Cointegração

- Testes de cointegração a partir de testes de raízes unitárias sobre os erros $u_t = y_t - \beta^{'} x_t - \gamma^{'} z_t$.
- $H_0 : u_t \sim I(1)$ (não existe cointegração) vs $H_1 : u_t \sim I(0)$ (existe cointegração) sobre os residuos $\widehat{u}_t$ ("residual-based tests"). Não se podem utilizar as distribuições DF porque a distribuição é outra.
- Testes Engle-Granger (DF) (ou AEG (ADF)) sobre a regressão $\triangle \widehat{u}_t = \alpha \widehat{u}_{t-1} + \varepsilon_t$. As distribuições dependem de k e dos termos determinísticos na regressão de cointegração.
- Versões PP: Phillips e Ouliaris, 1990.

4.3 Representações VAR e VECM e Previsão

- Teorema da representação de Granger: Em relações de equilíbrio de longo prazo entre variáveis (cointegração), existe uma representação em Modelo de Correcção de Erros (e vice-versa)

$$\triangle y_t = \delta^{'} z_t + \lambda(y_{t-1} - \beta' x_{t-1}) + \gamma \triangle x_t + u_t. \tag{12}$$

  Este modelo distingue claramente entre dinâmica de curto (diferenças) e longo prazo (termo em níveis de correcção de erro de equilíbrio no período anterior).

- Estimação:
  (a) A 2 passos: $\widehat{\beta}$ da regressão estática $y_t = \beta^{'} x_t + u_t$ e depois utilizando $\widehat{u}_{t-1}$ no ECM.

(b) Directa sob $\triangle y_t = \delta' z_t + \lambda y_{t-1} - \zeta' x_{t-1} + \gamma \triangle x_t + u_t$ onde $\widehat{\beta} = -\widehat{\zeta}/\widehat{\lambda}$. A inferência habitual não é válida.

- y_t cointegrado (por isso $I(1)$) admitindo uma representação $VAR(p)$ não estacionário

$$\Phi(L) y_t = \varepsilon_t, \Phi(L) = I_k - \Phi_1 L - ... - \Phi_p L^p,$$

chega-se a uma representação VECM

$$\triangle y_t = \Pi y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \triangle y_{t-j} + \varepsilon_t, \tag{13}$$

$$\Gamma_j = \Phi_1 - I_k - \Pi + \sum_{i=2}^{j} \Phi_i; \Pi = -\Phi(1)$$

- Portanto, se
  (a) $rank(\Pi) = 0$, $\Pi = 0$, não existindo cointegração e, dado que $y_t \sim I(1)$, o VAR tem que ser escrito em diferenças;
  (b) $rank(\Pi) = k$, $|\Pi| \neq 0$, pelo que $\Phi(L)$ não contém uma raíz unitária, ou seja, $y_t \sim I(0)$ e então o processo tem uma representação MA e VAR em niveis.
  (c) $rank(\Pi) = r$, $0 < r < k$, então $\Pi = \alpha\beta'$, em que α, β são matrizes $k \times r$ de ajustamento e de cointegração, respectivamente. Neste caso,

$$\triangle y_t = \alpha Z_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \triangle y_{t-j} + \varepsilon_t; Z_t = \beta' y_t. \tag{14}$$

- Previsão:
  (a) Não Cointegração: Usar VAR em niveis ou em diferenças.
  (b) Cointegração: Usar VECM ...

4.4 Método de Johansen

- Johansen (1988): MLE baseado no método de regressão ”reduced rank” e testes LR de cointegração.
- Estimação concentrada na dinâmica de curto prazo Γ_i :
  (a) Regride-se $\triangle y_t$ em $\triangle y_{t-i}$ $(i = 1, ..., p-1)$, obtendo-se os resíduos R_{0t}, procedendo-se de forma semelhante em relação a y_{t-1} e $\triangle y_{t-i}$, ficando R_{1t}.
  (b) A regressão de interesse é agora $R_{0t} = \alpha\beta' R_{1t} + \varepsilon_t$. As matrizes dos momentos entre R_{0t} e R_{1t} (correlações canónicas) resultam da típica matriz de regressão multivariada de elementos $k \times k$, $\begin{pmatrix} S_{00} & S_{01} \\ S_{10} & S_{11} \end{pmatrix}$, em que $S_{ij} = \frac{1}{T} \sum R_i R_j'$.
  (c) $\widehat{\beta} = (\widehat{q}_1, ..., \widehat{q}_r)$, as primeiras r colunas da matriz de vectores próprios associados ao valores próprios $\widehat{\lambda}_1 \geq ... \geq \widehat{\lambda}_r \geq \widehat{\lambda}_{r+1} \geq ... \geq \widehat{\lambda}_k$ que são solução da equação caracteristica

$$\left|\lambda S_{11,T} - S_{10,T} S_{00,T}^{-1} S_{01,T}\right| = 0. \tag{15}$$

- Testes de cointegração trace e lambda-max de Johansen:
  (a) lambda-max test: $H_0 : r$ *versus* $H_1 : r+1, r = 0, ..., k-1$.

(b) trace test: At most r cointegrating vectors, $H_0 : r$ *versus* $H_1 : r = k$.

- As distribuições do MLE e dos testes não são standard e dependem de k, r e da existência de componentes deterministicas. Casos:

(a) (raro) $\beta' y_t$ e VECM sem componentes deterministicas;

$$\triangle y_t = \alpha\beta' y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \triangle y_{t-j} + \varepsilon_t. \tag{16}$$

(b) $\beta' y_t$ com constante e VECM sem componentes deterministicas (y_t sem tendência);

$$\triangle y_t = \alpha \left(\rho_0 + \beta' y_{t-1}\right) + \sum_{j=1}^{p-1} \Gamma_j \triangle y_{t-j} + \varepsilon_t \tag{17}$$

$$= \alpha\rho_0 + \alpha\beta' y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \triangle y_{t-j} + \varepsilon_t. \tag{18}$$

(c) $\beta' y_t$ e VECM com constante (y_t com tendência estocástica

/ tendência linear);

$$\triangle y_t = \alpha_\perp \gamma_0 + \alpha\left(\rho_0 + \beta' y_{t-1}\right) + \sum_{j=1}^{p-1} \Gamma_j \triangle y_{t-j} + \varepsilon_t \tag{19}$$

$$= \mu_0 + \left(\rho_0 + \beta' y_{t-1}\right) + \sum_{j=1}^{p-1} \Gamma_j \triangle y_{t-j} + \varepsilon_t; \tag{20}$$

$$\mu_0 = \alpha_\perp \gamma_0. \tag{21}$$

(d) $\beta' y_t$ com constante e trend e VECM com constante (y_t com tendência deterministica / tendência linear);

$$\triangle y_t = \alpha_\perp \gamma_0 + \alpha\left(\rho_0 + \rho_1 t + \beta' y_{t-1}\right) + \sum_{j=1}^{p-1} \Gamma_j \triangle y_{t-j} + \varepsilon_t \tag{22}$$

$$= \mu_0 + \left(\rho_0 + \rho_1 t + \beta' y_{t-1}\right) + \sum_{j=1}^{p-1} \Gamma_j \triangle y_{t-j} + \varepsilon_t; \tag{23}$$

$$\mu_0 = \alpha_\perp \gamma_0. \tag{24}$$

(e) (raro) $\beta' y_t$ e VECM com constante e trend (y_t com tendência quadrática).

$$\triangle y_t = \alpha_\perp\left(\gamma_0 + \gamma_1 t\right) + \alpha\left(\rho_0 + \rho_1 t + \beta' y_{t-1}\right) + \sum_{j=1}^{p-1} \Gamma_j \triangle y_{t-j} + \varepsilon_t \tag{25}$$

$$= \mu_1 + \left(\rho_0 + \rho_1 t + \beta' y_{t-1}\right) + \sum_{j=1}^{p-1} \Gamma_j \triangle y_{t-j} + \varepsilon_t; \tag{26}$$

$$\mu_1 = \alpha_\perp\left(\gamma_0 + \gamma_1 t\right). \tag{27}$$

- Testes a restrições lineares ao espaço de cointegração H_0 : $\beta = H_{k\times s}Q_{s\times r}, s \leq r$. Estatistica $LR \xrightarrow{d} \chi^2_{r(k-s)}$.
- Identificação de α exige a imposição de r^2 restrições.