Processos Co-Integrados

1. Introdução

- X_t é integrado de ordem d se $\Delta^d X_t$ for estacionário, e escrevemos $X_t \sim \mathrm{I}(d)$.

  Em particular, um processo estacionário é I(0).

  Estudamos, em particular, a classe dos processos ARIMA(p, d, q). Para estes, após tomarmos d diferenças, o processo estacionário resultante é representado por um modelo ARMA(p, q).

- No Capítulo 7 tratamos de modelos VAR estacionários, ou seja, as séries envolvidas são I(0).

A teoria usual de MQO também aplica-se a séries I(0).

Se algumas ou todas as séries de um modelo de regressão são I(1), os resultados estatísticos usuais em geral **não são mais válidos**. Este é o problema da *regressão espúria*, tratado por Granger e Newbold (1974).

Estes autores verificaram, através de simulações, que dadas duas séries completamente não-correlacionadas, mas I(1), a regressão de uma sobre a outra tenderá a produzir uma relação aparentemente significativa.

Veja também Phillips (1986).

Há, portanto, a necessidade de se desenvolver técnicas para analisar relações entre séries não-estacionárias.

- Neste capítulo estaremos interessados em analisar modelos para descrever co-movimentos dinâmicos de duas ou mais séries temporais, como séries de preços de ativos ou taxas de câmbio. É comum que preços de ativos apresentem uma tendência estocástica comum no longo prazo, ou seja, que sejam *co-integrados*.

Preços e taxas (de câmbio, de juros) em geral são I(1) e é usual analisar os logaritmos destas séries, para investigar co-integração. Estabelecida uma relação de equilíbrio de longo prazo entre log-preços, por exemplo, ajusta-se um modelo que corrige desvios de curto prazo da relação de equilíbrio. Este modelo é chamado *modelo de correção de erros* (MCE).

- Se X_t e Y_t forem processos I(d), então a combinação linear $Z_t = Y_t - \alpha X_t$ será, em geral, também I(d).

Mas é possível que Z_t seja integrado de ordem menor, digamos I$(d-b), b>0$. Se $d=b=1$, então X_t e Y_t serão I(1) e Z_t será I(0). Neste caso, dizemos que X_t e Y_t são co-integrados.

Todavia, não é geralmente verdade que exista α tal que $Z_t \sim$ I(0) ou, em geral, $Z_t \sim$ I$(d-b)$.

- No caso de um vetor $\mathbf{X}_t$, de ordem $n \times 1$, dizemos que ele é integrado de ordem d, I(d), se d for a maior ordem de integração das séries individuais.

  Ou seja, se $\mathbf{X}_t = (X_{1t}, \ldots, X_{nt})', X_{it} \sim$I$(d_i)$, então $d =$max$(d_1, \ldots, d_n)$.

- **Definição 1.** As componentes do vetor $\mathbf{X}_t$ são *co-integradas de ordem* (d, b), e escrevemos, $\mathbf{X}_t \sim$ C.I.(d, b), se:

(a) todos as componentes de $\mathbf{X}_t$ são $\mathrm{I}(d)$;

(b) existe um vetor $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)^{'}$, não-nulo, tal que

$$\boldsymbol{\beta}^{'}\mathbf{X}_t = \beta_1 X_{1t} + \ldots + \beta_n X_{nt} \sim \mathrm{I}(d-b),\ d \geq b > 0. \quad (1)$$

O vetor $\boldsymbol{\beta}$, de ordem $n \times 1$, é chamado *vetor co-integrado* (ou vetor de co-integração).

- **Exemplo 1** (Engle e Granger, 1987). Considere $n = 2$ e as séries X_{1t} e X_{2t}, dadas por

$$X_{1t} + \beta X_{2t} = u_t, \quad (2)$$

$$u_t = \phi_1 u_{t-1} + \varepsilon_{1t}, \quad (3)$$

$$X_{1t} + \alpha X_{2t} = v_t, \tag{4}$$

$$v_t = \phi_2 v_{t-1} + \varepsilon_{2t}, \tag{5}$$

onde supomos os ε_{it} independentes e normais, com média zero e com $E(\varepsilon_{it}\varepsilon_{js}) = 0, i, j = 1, 2$. Suponha $\phi_i \neq 0, i = 1, 2$. Então temos os seguintes casos a analisar:

(i) $\phi_i < 1, i = 1, 2$.

Neste caso, X_{1t} e X_{2t} serão I(0), mas os parâmetros α e β não são identificados.

(ii) $\phi_1 = 1$, $\phi_2 < 1$.

As séries são ambas I(1) e $(1, \alpha)'$ é o vetor co-integrado. A equação 4) é identificada.

(iii) $\phi_1 < 1$, $\phi_2 = 1$.

Similar ao anterior, o vetor co-integrado é $(1, \beta)'$ e a equação (2) é identificada.

(iV) $\phi_1 = \phi_2 = 1$.

Ambas as séries são $\mathrm{I}(1)$, mas não são co-integradas, os parâmetros não são identificados.

- **Observações**: (i) O vetor de co-integração $\boldsymbol{\beta}$ não é único, pois se $\lambda \neq 0$, então $\lambda\boldsymbol{\beta}$ é também um vetor de co-integração. Tipicamente, uma das variáveis é usada para normalizar $\boldsymbol{\beta}$, fixando-se seu coeficiente igual a um.

  Usualmente toma-se $\boldsymbol{\beta} = (1, -\beta_2, \ldots, -\beta_n)'$, de modo que

$$\boldsymbol{\beta}'\mathbf{X}_t = X_{1t} - \beta_2 X_{2t} - \ldots - \beta_n X_{nt}.$$

Por exemplo, se $\beta^{'}\mathbf{X}_t \sim I(0)$, temos que

$$X_{1t} = \beta_2 X_{2t} + \ldots + \beta_n X_{nt} + u_t,$$

com $u_t \sim I(0)$. Dizemos que u_t é o *resíduo de co-integração*. Em equilíbrio de longo prazo, $u_t = 0$ e a relação de equilíbrio de longo prazo é

$$X_{1t} = \beta_2 X_{2t} + \ldots + \beta_n X_{nt}.$$

(ii) Todas as variáveis devem ser integradas de *mesma ordem*. Se elas forem integradas de ordens diferentes, não podem ser co-integradas.

(iii) Se $\mathbf{X}_t$ tiver $n > 2$ componentes, podem existir vários vetores de co-integração. Se existirem exatamente r vetores de co-integração linearmente independentes, com

$0 < r \leq n-1$, então eles podem ser reunidos numa matriz $\boldsymbol{A}$, de ordem $n \times r$, com posto r, chamado o *posto de co-integração*. Neste caso,

$$\mathbf{A}'\mathbf{X}_t = \begin{bmatrix} \boldsymbol{\beta}_1'\mathbf{X}_t \\ \vdots \\ \boldsymbol{\beta}_r'\mathbf{X}_t \end{bmatrix} = \begin{bmatrix} u_{1t} \\ \vdots \\ u_{rt} \end{bmatrix}$$

é estacionária, isto é, $\mathrm{I}(0)$.

2. **Tendências Comuns**

- Vimos que log-preços de ativos podem ser modelados por um passeio aleatório, ou seja,

$$\Delta p_t = \mu + \varepsilon_t,$$

onde $\varepsilon_t \sim$ i.i.d.$\mathcal{N}(0, \sigma^2)$. Logo, a melhor previsão de qualquer valor futuro é o valor de hoje mais um "drift". Mas se existe uma relação de co-integração entre dois ou mais log-preços, um modelo multivariado pode dar informação sobre o equilíbrio de longo prazo entre as séries.

- Preços co-integrados têm uma tendência estocástica comum, um fato apontado por Stock e Watson (1988). Ou seja, eles caminharão juntos no longo prazo porque uma combinação linear deles é reversível à média (estacionária).

- **Exemplo 2**. Suponha que

$$X_{1t} = \mu_{1t} + \varepsilon_{1t}, \tag{6}$$

$$X_{2t} = \mu_{2t} + \varepsilon_{2t}, \tag{7}$$

onde μ_{it} é um passeio aleatório representando a tendência estocástica da variável $X_{it}, i = 1, 2$ e $\varepsilon_{it}, i = 1, 2$ é estacionário.

Suponha que X_{1t} e X_{2t} sejam I(1) e que existam constantes β_1 e β_2 tais que $\beta_1 X_{1t} + \beta_2 X_{2t}$ seja I(0), ou seja,

$$\beta_1 X_{1t} + \beta_2 X_{2t} =$$

$$(\beta_1 \mu_{1t} + \beta_2 \mu_{2t}) + (\beta_1 \varepsilon_{1t} + \beta_2 \varepsilon_{2t})$$

seja estacionário. Então devemos ter o primeiro termo do segundo membro nulo, ou seja, $\mu_{1t} = -(\beta_2/\beta_1)\mu_{2t}$, o que mostra que as tendências são as mesmas, a menos de um escalar.

De modo geral, se o vetor $\mathbf{X}_t$ for co-integrado, com r vetores de co-integração, $0 < r < n$, então existirão $n-r$ tendências estocásticas comuns.

- **Exemplo 3.** Considere as séries

$$
\begin{aligned}
X_{1t} &= \beta_2 X_{2t} + u_t, \qquad (8)\\
X_{2t} &= X_{2,t-1} + v_t, \qquad (9)
\end{aligned}
$$

onde u_t e v_t são ambas I(0).

Segue-se que X_{2t} é um passeio casual e representa a tendência estocástica comum, ao passo que (8) representa a relação de equilíbrio de longo prazo.

O vetor de co-integração é $\boldsymbol{\beta} = (1, -\beta_2)'$.

Figura 1: séries simuladas, com $\beta_2 = 1$, $u_t = 0,6u_{t-1} + a_t$, a_t e v_t independentes $\mathcal{N}(0,1)$, independentes entre si.

3. **Modelo de Correção de erro**

- Vimos que se duas ou mais séries são co-integradas, existe uma relação de equilíbrio de longo prazo entre elas.

  Muitas variáveis econômicas apresentam relações de equilíbrio, como preços de um mesmo produto em diferentes mercados.

  Suponha, por exemplo, que P_{1t} e P_{2t} sejam tais preços em dois mercados distintos e que a relação (normalizada) de equilíbrio entre eles seja $P_{1t} - \beta P_{2t} = 0$.

  Suponha, ainda, que variações em P_{1t} dependam de desvios deste equilíbrio no instante $t-1$, ou seja,

$$\Delta P_{1t} = \alpha_1 (P_{1,t-1} - \beta P_{2,t-1}) + a_{1t}, \quad (10)$$

e uma relação similar valha para P_{2t}:

$$\Delta P_{2t} = \alpha_2(P_{1,t-1} - \beta P_{2,t-1}) + a_{2t}. \quad (11)$$

- O mesmo vale para um mecanismo de correção de erro mais geral. Suponha que X_{1t} e X_{2t} sejam duas séries I(1), $u_t = X_{1t} - \beta X_{2t} = 0$ seja a relação de equilíbrio e

$$\begin{aligned}
\Delta X_{1t} &= \alpha_1(X_{1,t-1} - \beta X_{2,t-1}) + a_{11}(1)\Delta X_{1,t-1} + \\
&+ a_{12}(1)\Delta X_{2,t-1} + a_{1t}, \qquad (12) \\
\Delta X_{2t} &= \alpha_2(X_{1,t-1} - \beta X_{2,t-1}) + a_{21}(1)\Delta X_{1,t-1} + \\
&+ a_{22}(1)\Delta X_{2,t-1} + a_{2t}. \qquad (13)
\end{aligned}$$

Este é um modelo VAR(1) nas primeiras diferenças com um termo de correção de erro adicionado.

Os parâmetros α_1 e α_2 são relacionados à velocidade de ajustamento.

Se ambos forem nulos não há relação de longo prazo e não temos um modelo como o acima.

Se $\mathbf{X}_t = (X_{1t}, X_{2t})'$, podemos escrever (12)-(13) como

$$\Delta \mathbf{X}_t = \boldsymbol{\alpha}\boldsymbol{\beta}' \mathbf{X}_{t-1} + \mathbf{A}\Delta \mathbf{X}_{t-1} + \mathbf{a}_t, \qquad (14)$$

com

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} 1 \\ -\beta \end{bmatrix},$$

$$\mathbf{A} = \begin{bmatrix} a_{11}(1) & a_{12}(1) \\ a_{21}(1) & a_{22}(1) \end{bmatrix}.$$

- Vemos que (14) também pode ser escrita

$$\mathbf{X}_t - \mathbf{X}_{t-1} = \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{X}_{t-1} + \mathbf{A}(\mathbf{X}_{t-1} - \mathbf{X}_{t-2}) + \mathbf{a}_t,$$

ou

$$\mathbf{X}_t = (\mathbf{I}_n + \mathbf{A} + \boldsymbol{\alpha}\boldsymbol{\beta}')\mathbf{X}_{t-1} - \mathbf{A}\mathbf{X}_{t-2} + \mathbf{a}_t, \quad (15)$$

logo variáveis que são co-integradas podem ser geradas por um processo VAR.

- Considere, agora, um modelo VAR(1) n-dimensional,

$$\mathbf{X}_t = \mathbf{\Phi}\mathbf{X}_{t-1} + \mathbf{a}_t. \quad (16)$$

O processo $\mathbf{X}_t$ será estacionário se todas as soluções de $|\mathbf{I}_n - \mathbf{\Phi} z| = 0$ estiverem fora do círculo unitário.

Suponha que o processo seja não-estacionário, com uma ou mais raízes sobre o círculo unitário.

Isto é equivalente a dizer que um ou mais auto-valores de $\mathbf{\Phi}$ são iguas a um, os demais estando dentro do círculo unitário.

Como $|\mathbf{I}_n - \mathbf{\Phi}| = 0$, a matriz $\mathbf{\Pi} = \mathbf{I}_n - \mathbf{\Phi}$ é singular. Suponha que o seu posto seja $\rho(\mathbf{\Pi}) = r < n$, de modo que $\mathbf{\Pi}$ pode ser decomposta como $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}^{'}$, onde $\boldsymbol{\alpha}$ e $\boldsymbol{\beta}$ têm ordem $n \times r$ e posto r.

- Suponha que as componentes de $\mathbf{X}_t$ sejam todas I(1). Então, de (16),

$$\mathbf{X}_t - \mathbf{X}_{t-1} = -(\mathbf{I}_n - \mathbf{\Phi})\mathbf{X}_{t-1} + \mathbf{a}_t,$$

ou

$$\Delta \mathbf{X}_t = -\mathbf{\Pi}\mathbf{X}_{t-1} + \mathbf{a}_t. \qquad (17)$$

Portanto, como

$$\boldsymbol{\alpha}\boldsymbol{\beta}^{'}\mathbf{X}_{t-1} = -\Delta \mathbf{X}_t + \mathbf{a}_t,$$

o segundo termo é I(0), logo $\boldsymbol{\alpha}\boldsymbol{\beta}^{'}\mathbf{X}_{t-1}$ é I(0) e continua a ser I(0) se o multiplicarmos por $(\boldsymbol{\alpha}^{'}\boldsymbol{\alpha})^{-1}\boldsymbol{\alpha}^{'}$, resultando

$\boldsymbol{\beta}^{'}\mathbf{X}_{t-1} \sim$I(0) e, finalmente, $\boldsymbol{\beta}^{'}\mathbf{X}_t \sim$I(0).

Segue-se que cada linha de $\boldsymbol{\beta}^{'}\mathbf{X}_t$ representará uma relação de co-integração.

Conclui-se que, a partir de um VAR(1) n-dimensional, obtemos um modelo nas primeiras diferenças com variáveis co-integradas.

- É fácil ver que, para um VAR(2)

$$\mathbf{X}_t = \mathbf{\Phi}_1 \mathbf{X}_{t-1} + \mathbf{\Phi}_2 \mathbf{X}_{t-2} + \mathbf{a}_t, \qquad (18)$$

obtemos

$$\Delta \mathbf{X}_t = \mathbf{D}_1 \Delta \mathbf{X}_{t-1} - \mathbf{\Pi} \mathbf{X}_{t-2} + \mathbf{a}_t, \qquad (19)$$

com $\mathbf{\Pi} = \mathbf{I}_n - \mathbf{\Phi}_1 - \mathbf{\Phi}_2$ e $\mathbf{D}_1 = -(\mathbf{I}_n - \mathbf{\Phi}_1)$.

Outra maneira de escrever (18) é

$$\Delta \mathbf{X}_t = \mathbf{F}_1 \Delta \mathbf{X}_{t-1} - \mathbf{\Pi} \mathbf{X}_{t-1} + \mathbf{a}_t, \qquad (20)$$

onde $\mathbf{F}_1 = -\mathbf{\Phi}_2$, $\mathbf{\Pi}$ como antes.

Esta forma é a chamada *forma de correção de equilíbrio* ou *de correção de erros*.

Por outro lado, tomando-se uma diferença em (16), obtemos

$$\Delta \mathbf{X}_t = \mathbf{\Phi} \Delta \mathbf{X}_{t-1} + \Delta \mathbf{a}_t,$$

que tem uma parte de médias móveis não invertível, logo não obtemos uma representação VAR para as primeiras diferenças. O mesmo ocorre com (18).

- Considere, agora, um modelo VAR(p) genérico,

$$\mathbf{X}_t = \mathbf{\Phi}_0 + \mathbf{\Phi}_1 \mathbf{X}_{t-1} + \ldots + \mathbf{\Phi}_p \mathbf{X}_{t-p} + \mathbf{a}_t. \quad (21)$$

Este diz-se *co-integrado de posto* r se $\mathbf{\Pi} = \mathbf{I}_n - \mathbf{\Phi}_1 - \ldots - \mathbf{\Phi}_p$ tiver posto r e portanto puder ser escrita como $\mathbf{\Pi} = \boldsymbol{\alpha} \boldsymbol{\beta}'$, $\boldsymbol{\alpha}$ e $\boldsymbol{\beta}$ de ordem $n \times r$ e posto r.

Dizemos que $\boldsymbol{\beta}$ é a *matriz de co-integração* ou de *vetores co-integrados* e $\boldsymbol{\alpha}$ é a *matriz de cargas*.

Se $r = 0$, então $\Delta\mathbf{X}_t$ tem uma representação VAR($p-1$) estacionária e se $r = n$, então o vetor $\mathbf{X}_t$ tem uma representação VAR (p) estacionária.

Neste caso, a representação (19) fica

$$\Delta\mathbf{X}_t = \mathbf{\Phi}_0+\mathbf{D}_1\Delta\mathbf{X}_{t-1}+\ldots+\mathbf{D}_{p-1}\Delta\mathbf{X}_{t-p+1}-\mathbf{\Pi}\mathbf{X}_{t-p}+\mathbf{a}_t, \tag{22}$$

com $\mathbf{D}_i = -(\mathbf{I}_n-\mathbf{\Phi}_1-\ldots-\mathbf{\Phi}_i), 1 \leq i \leq p-1$.

O processo (21) terá uma representação ou *modelo de correção de erros* (MCE)

$$\Delta\mathbf{X}_t = \mathbf{\Phi}_0-\boldsymbol{\alpha}\boldsymbol{\beta}^{'}\mathbf{X}_{t-1}+\mathbf{F}_1\Delta\mathbf{X}_{t-1}+\ldots+\mathbf{F}_{p-1}\Delta\mathbf{X}_{t-p+1}+\mathbf{a}_t, \tag{23}$$

onde $\mathbf{F}_i = -(\mathbf{\Phi}_{i+1}+\ldots+\mathbf{\Phi}_p), 1 \leq i \leq p-1$.

- Como $\Delta\mathbf{X}_t \sim \mathrm{I}(0)$ e $\boldsymbol{\beta}^{'}\mathbf{X}_{t-1} \sim \mathrm{I}(0)$, estes termos têm média constante; sejam

$E(\Delta \mathbf{X}_t) = \mathbf{c}$, um vetor $n \times 1$, representando taxas de crescimento, e $E(\boldsymbol{\beta}^{'}\mathbf{X}_{t-1}) = \boldsymbol{\mu}$, um vetor $r \times 1$, representando interceptos nas relações de co-integração. Temos, então, que

$$\boldsymbol{\Phi}_0 = (\mathbf{I}_n - \mathbf{F}_1 - \ldots - \mathbf{F}_{p-1})\mathbf{c} + \boldsymbol{\alpha}\boldsymbol{\mu}.$$

Segue-se que o termo contante é a soma de duas parcelas, uma relacionada com o crescimento dos dados e outra com os interceptos nas relações de co-integração.

Podemos, então, escrever (23) como

$$\Delta \mathbf{X}_t - \mathbf{c} = \sum_{i=1}^{p-1} \mathbf{F}_i(\Delta \mathbf{X}_{t-i} - \mathbf{c}) - \boldsymbol{\alpha}(\boldsymbol{\beta}^{'}\mathbf{X}_{t-1} - \boldsymbol{\mu}) + \mathbf{a}_t, \tag{24}$$

e vemos que há duas formas de correção de equilíbrio em (24):

- uma do crescimento dos dados em relação à sua média,

- outra, dos vetores de co-integração em relação à sua média.

Em análises de séries reais, temos que verificar se $\mathbf{c}$ e $\boldsymbol{\mu}$ são diferentes de zero ou não.

- Podemos obter estimadores de máxima verossimilhança para os parâmetros $\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{F}$ e $\boldsymbol{\Sigma}$ do modelo VAR(p) co-integrado , onde $\boldsymbol{\Sigma}$ é a matriz de covariâncias de $\mathbf{a}_t$. Veja Lütkepohl (1991) para detalhes.

4. **Testes para Co-Integração**

- Para se concluir que duas ou mais séries são co-integradas poderíamos pensar que bastaria analisar os seus gráficos. Todavia isso não é suficiente.

  Co-integração pode ou não existir entre séries que parecem ter uma tendência comum de longo prazo.

  É necessário usar testes formais e nesta seção veremos dois procedimentos para testar a existência de co-integração entre duas ou mais séries.

- Suponha o vetor $\mathbf{X}_t$, de ordem $n \times 1$, com todas as componentes I(1). Podemos destacar duas situações:

  (a) há, no máximo, um vetor de co-integração; este caso foi tratado por Engle e Granger (1987);

(b) há r, $0 \leq r < n$, vetores de co-integração, caso considerado por Johansen (1988).

Além dessas referências, veja Zivot e Huang (2003), que também é uma referência para o uso do program S+FinMetrics.

4.1 Procedimento de Engle e Granger

Seja $u_t = \boldsymbol{\beta}^{'}\mathbf{X}_t$ o resíduo de co-integração. O teste de Engle e Granger consiste em dois passos:

(i) forme os resíduos de co-integração;

(ii) faça um teste de raizes unitárias para determinar se esses resíduos são I(0). Temos as hipóteses:

$$
\begin{aligned}
H_0 &: \quad u_t \sim \mathrm{I}(1) : \text{ não há co-integração}, \\
H_1 &: \quad u_t \sim \mathrm{I}(0) : \text{há co-integração}. \quad (25)
\end{aligned}
$$

Temos, ainda, dois casos a considerar:

[1] **O vetor de co-integração é conhecido e fixado**

Por exemplo, o vetor é especificado pela teoria econômica. Use um teste ADF ou PP para testar H_0 contra H_1.

- **Exemplo 4.** Consideremos as duas séries geradas no exemplo 3, onde $\beta = (1, -1)'$.

  Usando a função **unitroot** do **S+FinMetrics** obtemos o

  Quadro 1: teste ADF com 2 lags e constante na regressão.

  Vemos que o valor observado da estatística ADF é $-12,39$, o que conduz à rejeição da hipótese H_0, que X_{1t} e X_{2t} não sejam co-integradas, com vetor de co-integração $\boldsymbol{\beta} = (1, -1)^{'}$.

- [2] **O vetor de co-integração é estimado**

  Para o caso de duas séries $\mathbf{X}_t = (X_{1t}, X_{2t})^{'}$, considere a regressão

$$X_{2t} = \alpha + \beta X_{1t} + u_t,$$

  e use os resíduos de MQO $\hat{u}_t$ para o teste de raiz unitária.

  No caso geral de $\mathbf{X}_t$, de ordem $n \times 1$, considere $\mathbf{X}_t = (X_{1t}, \mathbf{X}_{2t}^{'})^{'}$, com $\mathbf{X}_{2t} = (X_{2t}, \ldots, X_{nt})^{'}$ e a regressão

$$X_{1t} = \alpha + \boldsymbol{\beta}_2^{'} \mathbf{X}_{2t} + u_t, \qquad (26)$$

  para obter os resíduos de MQO $\hat{u}_t$ e testar H_0 contra H_1.

4.2 **Procedimento de Johansen**

O procedimento de Yohansen é uma generalização multivariada do teste de DF. Considere o modelo (23) re-escrito na forma

$$\Delta \mathbf{X}_t = \mathbf{\Phi_0} \mathbf{D}_t + \boldsymbol{\alpha}\boldsymbol{\beta}' \mathbf{X}_{t-1} + \mathbf{F}_1 \Delta \mathbf{X}_{t-1} + \ldots + \mathbf{F}_{p-1} \Delta \mathbf{X}_{t-p+1} + \mathbf{a}_t, \tag{27}$$

onde agora $\mathbf{\Pi} = \mathbf{\Phi}_1 + \ldots + \mathbf{\Phi}_p - \mathbf{I}_n$ e $\mathbf{D}_t$ contém termos determinísticos (constantes, tendências etc).

- O procedimento de Johansen (1988, 1995) para testar a existência de co-integração é baseado nos seguintes passos:

  (i) verificar a ordem de integração das séries envolvidas; verificar a existência de tendências lineares;

  (ii) especificar e estimar um modelo VAR(p) para $\mathbf{X}_t$, que supomos I(1);

(iii) construir testes da razão de verossimilanças (RV) para se determinar o número de vetores de co-integração, que sabemos ser igual ao posto de $\mathbf{\Pi}$;

(iv) dados os vetores de co-integração (normalizados apropriadamente), estimar o MCE (via EMV).

- Segundo Johansen (1994, 1995), os termos determinísticos em (27) são restritos à forma

$$\mathbf{\Phi}_0 \mathbf{D}_t = \boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t. \tag{28}$$

Para verificarmos o efeitos dos termos determinísticos no modelo VAR, consideremos um caso especial:

$$\Delta \mathbf{X}_t = \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t + \boldsymbol{\alpha}\boldsymbol{\beta}^{'} \mathbf{X}_{t-1} + \mathbf{a}_t, \tag{29}$$

onde $\boldsymbol{\mu}_0$ e $\boldsymbol{\mu}_1$ são ambos vetores $n \times 1$. Vamos decompor estes dois vetores em relação à média das relações de co-integração e em relação à média das taxas de crescimento,

$$\begin{aligned} \boldsymbol{\mu}_0 &= \boldsymbol{\alpha}\boldsymbol{\rho}_0 + \mathbf{c}_0, \\ \boldsymbol{\mu}_1 &= \boldsymbol{\alpha}\boldsymbol{\rho}_1 + \mathbf{c}_1. \end{aligned} \tag{30}$$

Então podemos escrever

$$\Delta \mathbf{X}_t = \boldsymbol{\alpha}\boldsymbol{\rho}_0 + \mathbf{c}_0 + \boldsymbol{\alpha}\boldsymbol{\rho}_1 t + \mathbf{c}_1 t + \boldsymbol{\alpha}\boldsymbol{\beta}^{'}\mathbf{X}_{t-1} + \mathbf{a}_t$$

$$= \boldsymbol{\alpha}(\boldsymbol{\rho}_0, \boldsymbol{\rho}_1, \boldsymbol{\beta}^{'}) \begin{pmatrix} 1 \\ t \\ X_{t-1} \end{pmatrix} + (\mathbf{c}_0 + \mathbf{c}_1 t) + \mathbf{a}_t,$$

ou ainda,

$$\Delta \mathbf{X}_t = \boldsymbol{\alpha} \begin{pmatrix} \boldsymbol{\rho}_0' \\ \boldsymbol{\rho}_1' \\ \boldsymbol{\beta} \end{pmatrix} \mathbf{X}_{t-1}^* + (\mathbf{c}_0 + \mathbf{c}_1 t) + \mathbf{a}_t, \tag{31}$$

com $\mathbf{X}_{t-1}^* = (1, t, \mathbf{X}_{t-1}')'$.

Podemos sempre escolher $\boldsymbol{\rho}_0$ e $\boldsymbol{\rho}_1$ tais que o erro de equilíbrio $(\boldsymbol{\beta}^*)'\mathbf{X}_t^* = \mathbf{v}_t$ tenha média zero, logo

$$E(\Delta \mathbf{X}_t) = \mathbf{c}_0 + \mathbf{c}_1 t. \tag{32}$$

Note que se $\mathbf{c}_0 \neq \mathbf{0}$ temos um crescimento constante nos dados e se $\mathbf{c}_1 \neq \mathbf{0}$ temos uma tendência linear nas diferenças ou tendência quadrática nos níveis das variáveis.

- Há cinco casos a considerar.

Caso 1. constante nula, $\boldsymbol{\mu}_t = \mathbf{0}$; neste caso, $\boldsymbol{\rho}_0 = \boldsymbol{\rho}_1 = \mathbf{0}$ e o modelo não possui qualquer componente determinística, com $\mathbf{X}_t \sim \mathrm{I}(1)$ sem "drift" (não há crescimento dos dados) e as relações de co-integração têm média zero. A menos que $\mathbf{X}_0 = \mathbf{0}$, este caso tem pouco interesse nas aplicações práticas.

Caso 2. constante restrita, $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0 = \boldsymbol{\alpha}\boldsymbol{\rho}_0$; neste caso, $\boldsymbol{\rho}_1 = \mathbf{0}, \mathbf{c}_0 = \mathbf{0}$, mas $\boldsymbol{\rho}_0 \neq \mathbf{0}$ e portanto não há tendência linear nos dados e as relações de co-integração têm média $\boldsymbol{\rho}_0$.

Caso 3. constante irrestrita, $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0$; aqui, $\boldsymbol{\rho}_1 = \mathbf{0}$, as séries de $\mathbf{X}_t$ são $\mathrm{I}(1)$ sem "drift" e as relações de co-integração podem ter médias diferentes de zero.

Caso 4. tendência restrita, $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\alpha}\boldsymbol{\rho}_1 t$; neste caso, $\mathbf{c}_1 = \mathbf{0}$, mas $\mathbf{c}_0, \boldsymbol{\rho}_0, \boldsymbol{\rho}_1$ são irrestritos. As séries são $\mathrm{I}(1)$ com "drift" e as relações de co-integração têm uma tendência linear.

Caso 5. tendência irrestrita, $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t$; não há henhuma restrição sobre $\boldsymbol{\mu}_0$ e $\boldsymbol{\mu}_1$, as séries são $\mathrm{I}(1)$ com tendência linear (logo tendência quadrática nos níveis) e as relações de co-integração têm tendência linear. Previsões podem ser ruins, logo deve-se ter certo cuidado em se adotar esta opção.

- Sabemos que o posto de $\mathbf{\Pi}$ fornece também o número de autovalores não-nulos de $\mathbf{\Pi}$; suponha que os ordenemos $\lambda_1 > \lambda_2 > \cdots > \lambda_n$.

  Se as séries são *não* co-integradas, $\rho(\mathbf{\Pi}) = 0$ e todas os auto-valores serão nulos, ou ainda $\ell n(1-\lambda_i) = 0$, para todo i. Um teste da RV para testar o posto de $\mathbf{\Pi}$ é baseado na *estatística traço*

$$\lambda_{\text{traço}}(r_0) = -T \sum_{i=r_0+1}^{n} \ell n(1 - \hat{\lambda}_i), \quad (33)$$

onde $\hat{\lambda}_i$ são os auto-valores estimados de $\mathbf{\Pi}$ e (33) testa

$$\begin{aligned} H_0 &: r \leq r_0, \\ h_1 &: r > r_0, \end{aligned} \quad (34)$$

sendo r o posto de $\mathbf{\Pi}$.

Se $\rho(\mathbf{\Pi}) = r_0$, então $\hat{\lambda}_{r_0+1}, \ldots, \hat{\lambda}_n$ são aproximadamente nulas e a estatística (33) será pequena; caso contrário, será grande.

Como dissemos acima, a distribuição assintótica de (33) é uma generalização multivariada da distribuição ADF e depende da dimensão $n - r_0$ e da especificação dos termos determinísticos.

Os valores críticos podem ser encontrados em Osterwald-Lenum (1992) para os casos (a)-(e) acima e $n - r_0 = 1, \ldots, 10$.

Johansen também usa a **estatística do máximo auto-valor**

$$\lambda_{\text{max}}(r_0) = -T\ell n(1 - \widehat{\lambda}_{r_0+1}), \qquad (35)$$

para testar

$$\begin{aligned} H_0 &: r = r_0, \\ H_1 &: r = r_0 + 1. \qquad (36) \end{aligned}$$

- A distribuição assintótica de (35) também depende de $n-r_0$ e da especificação de termos determinísticos. Valores críticos podem ser encontrados na referência acima citada.

Supondo-se que o posto de $\mathbf{\Pi}$ é r, Johansen (1988) prova que o estimador de máxima verossimilhança de $\boldsymbol{\beta}$ é dado por $\widehat{\boldsymbol{\beta}}_{\text{MV}} = (\widehat{v}_1, \ldots, \widehat{v}_r)$, onde $\widehat{v}_i$ é o auto-vetor associado ao auto-valor $\widehat{\lambda}_i$ e os estimadores de máxima verossimilhança dos parâmetros restantes são obtidos por meio de uma regressão multivariada com $\boldsymbol{\beta}$ substituído pelo EMV.

Johansen (1995) mostra a normalidade assintótica dos estimadores de $\boldsymbol{\beta}$, com taxa de convergência T^{-1}.