CO-INTEGRAÇÃO E SUAS REPRESENTAÇÕES: UMA RESENHA*

Pedro L. Valls Pereira**

Resumo

Neste artigo apresenta-se uma resenha das possíveis representações de processos co-integrados num contexto multivariado. O conceito de co-integração num contexto multivariado possibilita a existência de relações lineares entre variáveis integradas que apresentam ordem de integração menor que a das variáveis presentes nesta relação linear. O caso mais estudado é aquele em que a relação entre as variáveis resulta numa variável estacionária, isto é, $I(0)$, já que neste caso pode-se aplicar a análise de regressão convencional. Discute-se as diversas representações de processos co-integrados, dentre os quais, o mais útil, do ponto de vista da teoria econométrica, é o do modelo de Mecanismo de correção dos erros (M.C.E.), onde os agentes econômicos utilizam "servo-mecanismo" que corrige os erros cometidos previamente, de acordo com uma relação de equilíbrio pré-estabelecida.

Abstract

This survey article presents some possible representations of multivariate co-integrated processes. The concept of multivariate co-integration allows the existence of linear relationship between integrated variables which has order of integration smaller than the variables that belong to this relationship. Most of the literature has studied the case in which the relationship between the variables is a stationary variable, *i.e.* $I(0)$, because, in this case, the convencional regression approach can be used. It is presented several representations of co-integrated processes, and from the econometric point of view, the most important is the Error Correction Model approach, which allows adjustment to some long-run equilibrium described by economic theory. The short-run adjustment is a "servo-mechanism" which econometricians are free to model pragmatically.

1. Introdução.

O fato dominante em macroeconomia, pelo menos para economias desenvolvidas, nos últimos séculos, é o crescimento do produto.

---

*Este artigo é uma versão modificada do capítulo III da tese de Livre Docência apresentada ao IME-USP. O autor agradece os comentários dos participantes do XII Encontro Brasileiro de Econometria realizado em Brasilia, e em particular, a um leitor anônimo desta revista, que em muito contribuíram para a versão final deste artigo. Os erros e omissões remanescentes são, contudo, de minha inteira responsabilidade.

**Professor Associado do Departamento de Estatística da USP e do Instituto de Economia da UNICAMP.

R. de Econometria Rio de Janeiro v. 11, nº 2, p.185-216 novembro 1991

Por exemplo, para os E.U.A., o Produto Nacional Bruto Real cresceu a uma média de 3,7% durante o período de 1820 a 1986, mas a extrapolação desta taxa indica que a mesma não pode ter sido sustentada por décadas.

Quais são, então, as causas da flutuação do crescimento?

Blanchard e Fisher (1989) apresentam fatos estilizados para tentar caracterizar estes movimentos.

O primeiro fato diz respeito ao comportamento estocástico do Produto e Desemprego, já que os movimentos destas variáveis estão associados a flutuações no bem-estar. Isto é, períodos de *boom*, expansão e declínio do desemprego estão associados a períodos de euforia quando governos são reeleitos; períodos de recessão e depressão estão associados a crises e descontentamentos.

Necessita-se, portanto, obter regularidades empíricas que caracterizem estas variáveis. A primeira tentativa de estudar os ciclos econômicos de séries temporais foi feita por Burns e Mitchell (1946). O objetivo era tentar caracterizar o ciclo econômico através do comprimento das expansões e contrações, assim como da amplitude das flutuações. Foram construídos índices da atividade econômica, conhecidos por indicadores antecedentes, que tiveram um papel importante para sumariar e fazer previsões para o nível e o ciclo da atividade macroeconômica.

Atualmente, macroeconometristas procuram regularidades nas flutuações do produto, através da separação da tendência e ciclo que compõem esta variável. Nesta abordagem, tenta-se caracterizar dois tipos de choques que podem afetar o nível da atividade econômica.

O primeiro, chamado de choque permanente, é aquele cujos efeitos, no produto, são duradouros. Este tipo de choque está associado ao componente de tendência da série e é, por construção, não-estacionário. Melhorias na produtividade ou aumento na força de trabalho são exemplos de choques permanentes no produto, já que seus efeitos no produto são duradouros. O segundo tipo, chamado de choque transitório, tem efeito que desaparece ao longo do tempo. Este segundo tipo de choque está associado ao componente de ciclo, e é, por construção, estacionário. As quebras de safra agrícola e as guerras no oriente médio, que afetaram as exportações brasileiras, são exemplos de choques transitórios no produto.

Como o conceito de choque permanente está associado ao componente de tendência da série econômica e, se esta for estocástica, tem-se uma relação entre os movimentos permanentes destas séries e o conceito de integração. Este conceito está relacionado a um resultado conhecido (ver Granger 1966) na literatura de que a maioria das séries temporais econômicas são não-estacionárias e requerem diferenças de pelo menos primeira ordem para induzir estacionariedade.

Por outro lado, as flutuações no produto podem estar associadas às flutuações de outras variáveis macroeconômicas, por exemplo, o desemprego. Portanto, deve-se tentar obter uma decomposição dos choques que leve em consideração os co-movimentos destas variáveis. Contudo, a Teoria Econômica raramente sugere equilíbrios que não sejam funções estacionárias das variáveis envolvidas. Isto implica que existem forças econômicas fundamentais que, ao longo do tempo, fazem com que estas variáveis evoluam estocasticamente juntas. Em outras palavras, visto que todas as variáveis econômicas individuais envolvidas na teoria podem ser não-estacionárias, os desvios de um dado equilíbrio são limitados. Este conceito está relacionado ao de co-integração, já que, se as variáveis co-integram, os choques permanentes caracterizam a relação de longo prazo entre as variáveis e os desvios deste longo prazo, que caracterizam a dinâmica de curto prazo, estão relacionados aos choques transitórios.

Um outro problema associado às relações estáticas é o de regressão espúria (ver Granger e Newbold 1974 e Phillips 1986). As soluções adotadas na literatura, talvez influenciadas pela metodologia de Box e Jenkins (1976), são as de transformar todas as variáveis em estacionárias antes de fazer as regressões. Neste tipo de procedimento, os modelos resultantes descartam as relações de longo prazo sugeridas pela teoria econômica que possam existir entre as variáveis.

Uma solução para os problemas apresentados é o uso do Modelo de Mecanismo de Correção dos Erros. Este modelo foi sugerido por Sargan (1964)[1] e tem a vantagem de reter informações sobre o nível das séries, de modo que as relações de longo prazo entre as variáveis do modelo permaneçam presentes. Granger (1983) mostra que existe

[1]Sargan introduziu o modelo M.C.E. na estimação da Curva de Phillips para a Inglaterra, mas foi com os trabalhos de Hendry (ver Hendry 1987), para uma descrição metodológica destes modelos que este tipo de modelo foi difundido. Este modelo pode ser interpretado como aquele em que existe uma regra de controle ótimo, e o erro do período anterior é parcialmente corrigido no próximo período.

uma equivalência entre M.C.E. e séries co-integradas. Isto é, se um conjunto de variáveis co-integradas for encontrado, existe um M.C.E. que representa estas variáveis e vice-versa.

Neste artigo, estendem-se os conceitos de processos integrados a um contexto multivariado, analisando-se o conceito de co-integração, isto é, a possibilidade da existência de relações lineares entre variáveis integradas que apresentam ordem de integração menor que a das variáveis presentes nesta relação linear. O caso mais estudado é aquele em que a relação entre as variáveis resulta numa variável estacionária, isto é, $I(0)$, já que neste caso pode-se aplicar a análise de regressão convencional. Discutem-se, também, diversas representações de processos co-integrados, dentre os quais o mais útil, do ponto de vista da teoria econométrica, é o do modelo de M.C.E., onde os agentes econômicos utilizam "servo-mecanismos"[2] que corrigem os erros cometidos previamente, de acordo com uma relação de equilíbrio pré-estabelecida.

A estrutura do artigo é a seguinte. A seção 2 apresenta formas de se decompor uma série temporal em componentes integrados. Na seção 3 são apresentados os teoremas de representação para processos co-integrados. Na seção 4 são feitas extensões dos teoremas de representação para os casos em que o vetor tem mais de uma raiz unitária ou possui raízes unitárias nas freqüências sazonais. Na última seção são apresentadas algumas conclusões e sugestões não-técnicas para a possível utilização dos resultados apresentados neste artigo.

2. Decomposição de uma série temporal em componentes integrados.

Granger (1981) formalizou o conceito de uma série ser integrada de ordem d (denotada por $I(d)$, se a n-ésima diferença for estacionária, em geral $d = 1$ é suficiente. Esta definição de processo integrado será estendida a uma série temporal vetorial. Seja $\{Y\}_0^\infty$

[2] Alternativamente, Nickell (1985) justifica M.C.E. como resultante de respostas ótimas de agentes econômicos e certos ambientes dinâmicos. Hendry e Ericsson (1991a) discutem como M.C.E. generalizam modelos convencionais de ajustamento parcial e podem ser consistentes com ajustamento do tipo Ss; cf. (Baumol 1952 e Miller e Orr 1966). Hendry e Ericsson (1991b) reinterpretam M.C.E. como formação de expectativas *forward-looking*, embora com *"data-based"* em vez de *"model-based"*.

um processo estocástico k-dimensional e, se todos os componentes de Y_t são $I(0)$ e não tem componente determinístico, pode-se, através da Decomposição de Wold, representar este processo por um processo Média-Móvel Vetorial de ordem infinita. Seja $C(L)$ uma matriz $k \times k$ polinominal no operador defasagem L e ε_t um vetor, $k \times 1$, de variáveis aleatórias, serialmente e contemporaneamente não-correlacionadas, então:

$$Y_t = C(L)\varepsilon_t \tag{1}$$

com $\mathbf{E}(\varepsilon_t \varepsilon_t') = I, \mathbf{E}(\varepsilon_t \varepsilon_{t-i}') = 0$ para $i \neq 0$ e $C(0)$ triangular inferior. O processo Y_t admite uma representação espectral multivariada dada por:

$$F_{YY}(\omega) = C(e^{i\omega})C'(e^{-i\omega}) \tag{2}$$

onde ω é uma freqüência e C' denota a transposta da matriz C. Os elementos diagonais de $F_{YY}(\omega)$ representam os espectros univariados de cada componente do vetor aleatório. Pela hipótese de que todos os componentes de Y_t são $I(0)$, tem-se que estes espectros univariados não são nem zero, nem infinito. Em particular, isto deve ser verdade para $\omega = 0$, a freqüência zero ou o componente de longo prazo da série que é reduzido por diferenciação. O espectro na freqüência zero é dado por:

$$F_{YY}(0) = C(1)\ C'(1) \tag{3}$$

que terá elemento zero na diagonal se a linha de $C(1)$ é identicamente igual a zero e terá elemento diagonal igual a infinito se qualquer elemento de $C(1)$ for infinito.

Similarmente, se todos os elementos de Y_t são $I(1)$, então

$$\Delta Y_t = C(L)\varepsilon_t \tag{4}$$

sob as mesmas condições sobre $C(L)$ e ε_t. Pode-se, através de (4), obter uma representação para o processo que só dependa do nível do mesmo, isto é, pode-se escrever $Y_t = \Delta^{-1}C(L)\varepsilon_t$, ou equivalentemente:

$$Y_t = Y_0 + (1 + L + L^2 + ... + L^{t-1})C(L)\varepsilon_t \tag{5}$$

que é um processo de média móvel com um componente Y_0, que pode ser interpretado como determinístico condicional nas informações de pré-amostra. Observe agora que a soma dos coeficientes da parte média-móvel é dada por $tC(1)$, que cresce quando t cresce, exceto para $C(1) = 0$. Isto implica que o espectro de Y_t, para t grande, tenderá para infinito na freqüência zero.

Suponha que se diferencie mais uma vez o processo (4), obtém-se, então:

$$\Delta^2 Y_t = (1 - L)C(L)\varepsilon_t \tag{6}$$

cuja representação média móvel terá coeficientes que somam zero e, portanto, o espectro será zero na freqüência zero.

Estes casos ilustram o seguinte fato:

(i) subdiferenciação implica que a freqüência zero do espectro tem um pólo e

(ii) superdiferenciação implica que a freqüência zero do espectro é zero.

Portanto, é necessário caracterizar a ordem de integração do processo para que os dois casos ilustrados acima não aconteçam.

Suponha, por outro lado, que este processo é tal que, após a extração do componente determinístico, tenha a seguinte representação ARIMA multivariada:

$$\Phi(L)Y_t = \Theta(L)\varepsilon_t \qquad \varepsilon_t \sim N_k(0, \Omega) \tag{7}$$

onde $\Phi(L)$ e $\Theta(L)$ são matrizes, $k \times k$, de polinômios no operador de defasagem L com ordens p e q, respectivamente, e $\mathbf{E}(\varepsilon_t \varepsilon_t') = 0 \; t \neq s$, isto é, os erros só podem ser contemporaneamente correlacionados, mas não temporalmente. O componente auto-regressivo do processo (1) é tal que $\text{Det}(\Phi(L))$ tenha todas as raízes fora do círculo unitário, exceto por um número d de raízes iguais à unidade. Assume-se, também, que $\Phi(0) = \Theta(0) = I$, e a matriz polinominal $K(L) = \Phi^{-1}(L)\Theta(L)$ tem uma expansão em série de potência convergente. Esta normalização é usada para que se possa identificar, no contexto de parametrização, o processo (ver, por exemplo, Deistler 1985).

Dados o processo (7) e suas condições sobre as matrizes polinominais, $(Y_t)_0^\infty$ é um processo integrado de ordem d, $I(d)$, e tem a seguinte representação (representação de Wold):

$$(1 - L)^d Y_t = C(L)\varepsilon_t \tag{8}$$

onde $C(L) = \dfrac{\Phi^*(L)\Theta(L)}{\det\left[\dfrac{\Phi(L)}{(1-L)^d}\right]}$ sendo $\Phi^*(L)$ a matriz adjunta de $\Phi(L)$.

Observe que $C(L)$ é uma fração racional, que admite um desenvolvimento em série, com coeficientes decrescentes em forma exponencial.

Pela decomposição de Wold, dada por (8), o processo $\{Y\}_0^\infty$ pode ser decomposto em componentes integrados em ordem decrescente. Esta representação é obtida através do desenvolvimento de Taylor da matriz $C(L)$ até a ordem $(d-1)$ e, posteriormente, divide-se os termos deste desenvolvimento por $(1-L)^d$, obtendo-se:

$$\begin{aligned} Y_t &= C(1)(1-L)^{-d}\varepsilon_t - C^{(1)}(1)(1-L)^{-d-1}\varepsilon_t + \\ &\cdots + \frac{C^{(d-1)}(1)(-1)^{d-1}}{(d-1)!}(1-L)^{-1}\varepsilon_t + C^{**}(L)\varepsilon_t \end{aligned} \tag{9}$$

Desta forma é óbvio que a matriz $C^{**}(L)$ é tal que as raízes da equação $\det(C^{**}(L))$ estão fora do círculo unitário. Logo, a representação (9) decompõe Y_t em vários componentes integrados de ordem decrescente, a partir da ordem d. Pode-se, portanto, reescrever Y_t da seguinte forma:

$$Y_t = Y_t^{(d)} + Y_t^{(d-1)} + \cdots + Y_t^{(1)} + Y_t^{(0)} \tag{10}$$

onde $Y_t^{(k)} = \dfrac{(-1)^{d-k}C^{(d-k)}(1)(1-L)^{-k}}{(d-k)!}\varepsilon_t.$

A representação (10) do processo $\{Y\}_0^\infty$ não é única, já que se pode reordenar os diferentes componentes, por exemplo $(X_t^{(d-k)} + X_t^{(d-k-1)})/2$, de tal sorte que se tenha a mesma ordem de integração que $X_t^{(d-k)}$. Uma forma de obter esta representação é através da Forma Canônica de Jordam da Matriz C, onde esta matriz é triangular inferior. Sims *et allii* (1987) mostram como construir esta representação, que será chamada de representação canônica.

Dado que o processo Y_t pode ser caracterizado pela representação (10), podem existir combinações lineares dos componentes de Y_t, de forma que alguns (ou todos) os componentes integrados desapareçam. Este conceito é conhecido por co-integração e é definido abaixo.

DEFINIÇÃO 1: O vetor $Y_t, k \times 1$, de variáveis aleatórias é dito ser co-integrado de ordem $b (b > 0)$ se:

(i) cada componente de Y_t é $I(d), (d > b)$ e

(ii) existe um vetor α tal que $\alpha' Y_t \sim I(d-b)$, onde $\alpha \neq 0$. Denota-se $Y_t \sim CI(d,b)$ e α é chamado de vetor co-integrado.

A definição acima é devida a Engle e Granger (1987) que adaptam a definição de co-integração (de Granger 1981 e Granger e Weiss 1983).

Uma forma alternativa de definir co-integração é através dos espaços vetoriais definidos por $C^{(k)}(1)$ da expressão (10). Tem-se, então, a seguinte proposição:

Proposição 1. *Dizemos que Y_t é co-integrada de ordem $b, CI(d,b)$ se e somente se existe um vetor α tal que $\alpha \in C_{d-b+1}$* mas $\alpha \notin$ *C_{d-b}, onde $C_{d-b} = \bigcap_{k=0}^{s} Ker[C^{(k)}(1)]$ e Ker representa o núcleo da transformação.*

Demonstração: Pela definição acima e por (10) tem-se:

$$\alpha' Y_t = \alpha' Y_t^{(d)} + \cdots + \alpha' Y_t^{(1)} + \alpha' Y_t^{(0)} \sim I(d-b) \tag{11}$$

agora pela definição de $Y_t^{(d-i)}$ dada em (10) e como a soma de dois processos integrados é um processo integrado com a ordem igual ao máximo das ordens dos dois processos, (11) implica que:

$$\alpha' C(1) = \alpha' C^{(1)}(1) = \alpha' C^{(b-1)}(1) = 0 \tag{12}$$

implicando que o núcleo de $C^{(i)}(1)$ é diferente de zero, para $i = 0, \cdots, b-1$ e segue-se o resultado. □

Um caso particular de interesse é aquele em que Y_t é $CI(1,1)$. Neste caso, tem-se que $\alpha' C(1) = 0$, mas este vetor α não é necessariamente único. No caso em que Y_t tem dimensão 2 e Y_t é $CI(1,1)$ este vetor é único, a menos do sinal.[3] Estes dois resultados também valem quando Y_t é $CI(d,b)$.

Uma forma de caracterizar o equilíbrio entre um conjunto de variáveis é através da restrição $\alpha' Y_t = 0$. Deste modo, a relação entre co-integração e equilíbrio é óbvia.

[3] Para uma discussão sobre não-unicidade do vetor co-integrado e suas implicações veja Swamy *et allii* (1989).

Por exemplo, se uma proporção λ de qualquer aumento no nível dos preços é finalmente repassada ao salário nominal, em equilíbrio, $w = k + \lambda p$, onde w e p denotam o salário nominal e o nível dos preços, respectivamente, e k é uma constante. Portanto, se

$$w - k - \lambda p = 0$$

em qualquer período de tempo, então o mercado de trabalho estará em equilíbrio se $\lambda = 1$ e o salário estiver medido em unidades de eficiência, já que, neste caso, o salário real será constante. Evidentemente, o salário nominal leva algum tempo para se ajustar a possíveis mudanças no nível dos preços, mas o escalar $z_t = w_t - k - \lambda p_t$ reflete o desvio do equilíbrio, ou desequilíbrio, no instante de tempo t. Se w e p são co-integradas, então estes desvios do equilíbrio são limitados. Portanto, uma forma de testar a teoria é determinar a ordem de integração da variável z_t. Se a hipótese nula de raiz unitária não for rejeitada para z_t, então os salários não tenderão para o equilíbrio e neste caso, a estimação da relação de equilíbrio não faz sentido.

O exemplo a seguir, devido a Engle e Granger (1987), mostra que a existência de co-integração possibilita a identificação de modelos que num contexto estacionário seriam não-identificados.

EXEMPLO 1: Suponha que o processo gerador de $Y_t = (Y_{1t}, Y_{2t})$ é dado por:

$$\begin{aligned} Y_{1t} + \beta Y_{2t} &= v_t & v_t &= \rho_1 v_{t-1} + \varepsilon_{1t} \\ Y_{1t} + \alpha Y_{2t} &= u_t & u_t &= \rho_2 u_{t-1} + \varepsilon_{2t} \end{aligned}$$

$\varepsilon_{it} \sim \mathrm{NID}(0, \sigma_i^2)$, para $i = 1, 2$ e $\mathrm{E}(\varepsilon_{it}\varepsilon_{js}) = 0$, para $i, j = 1, 2$.

Observe que se Y_{1t} e Y_{2t} forem estacionárias, então α e β não seriam identificados. Este exemplo, quando $\rho_i = 0$ para $i = 1, 2$, corresponderia ao modelo de determinação simultânea de quantidades e preços num mercado em equilíbrio, que é o exemplo clássico para introduzir o problema de identificação de modelos de equações simultâneas.

Suponha agora que $\rho_i \neq 0$, para $i = 1,2$ e pode-se considerar quatro casos:

(i) $\rho_i < 1$ para $i = 1, 2$;
(ii) $\rho_1 = 1$ e $\rho_2 < 1$;
(iii) $\rho_1 < 1$ e $\rho_2 = 1$; e
(iv) $\rho_1 = \rho_2 = 1$.

Para o primeiro caso, temos que Y_{1t} e Y_{2t} são $I(0)$ e este caso é equivalente ao caso em que $\rho_i = 0$, isto é, os parâmetros de interesse, α e β, não são identificados.

Para o segundo caso, tem-se que Y_{1t} e Y_{2t} são $I(1)$ e $(1, \alpha)$ representa o vetor co-integrado. Portanto a segunda equação do sistema é identificada.

Para o terceiro caso, temos o caso simétrico do caso anterior e, portanto, pode-se identificar a primeira equação do sistema e $(1, \beta)$ representa o vetor co-integrado para as variáveis Y_{1t} e Y_{2t}, que são $I(1)$.

Para o quarto caso, Y_{1t} e Y_{2t} são $I(1)$ mas não co-integram; portanto, os parâmetros do modelo não são identificados.

O exemplo a seguir ilustra um modelo onde as variáveis podem estar em nível ou em diferença. Este exemplo poderia ser associado a uma equação de demanda de moeda, onde se relaciona o estoque nominal de moeda ao nível dos preços e a taxa de inflação.

EXEMPLO 2: Suponha que o processo gerador de $Y_t = (Y_{1t}, Y_{2t})$ é dado por:

$$Y_{1t} = \beta Y_{2t} + \gamma \Delta Y_{2t} + \varepsilon_{1t}$$

$$\Delta^2 Y_{2t} = \varepsilon_{2t}$$

$\varepsilon_{it} \sim \mathrm{NID}(0, \sigma_i^2)$, para $i = 1, 2$ e $\mathbf{E}(\varepsilon_{it}\varepsilon_{js}) = 0$, para $i, j = 1, 2$.

É fácil mostrar que Y_{1t} e Y_{2t} são $I(2)$ e da primeira equação deriva-se que $(1, -\beta, -\gamma)$ é o vetor co-integrado que torna a combinação linear de Y_{1t}, Y_{2t} e ΔY_{2t} estacionária, isto é, $I(0)$. Observe que Y_{1t} e Y_{2t} são $CI(2,1)$, já que pela primeira equação $Y_{1t} - \beta Y_{2t}$ e $I(1)$ e necessita-se de uma outra variável $I(1)$, no caso ΔY_{2t}, para que se possa ter co-integração entre as variáveis. Neste exemplo, Y_{1t} representaria o estoque nominal de moeda e Y_{2t} o nível de preços, o que implicaria que ΔY_{2t} seria a taxa de inflação. Caso $\beta = 1$, o sistema acima estaria obtendo a relação entre base real e inflação.

3. Teoremas de representações.

Ao se definir co-integração adotou-se a representação MA do processo. Pode-se utilizar outras representações, isto é, AR ou M.C.E., e examinar que restrições devem ser impostas nestas representações para que se tenha co-integração das variáveis. Apresentam-se, inicialmente, os Teoremas de Representações para o caso em que Y_t

é $CI(1,1)$ e, na seção seguinte, apresentam-se as extensões para os outros casos.

Seja dado o processo estocástico k-variado, $(Y_t)_0^\infty$, tal que existam r vetores co-integrados, $r < k-1$, a serem agrupados na matriz Γ de dimensão $k \times r$. Tem-se, então, as seguintes representações:

1º Caso – Representação MA:

Fazendo-se $d = 1$ em (8), tem-se que o processo $(Y_t)_0^\infty$ segue um processo MA, isto é:

$$(1-L)Y_t = C(L)\varepsilon_t \tag{13}$$

e por (12) tem-se:

$$\alpha' C(1) = 0 = \Gamma' C(1) = 0 \tag{14}$$

e, portanto, as colunas de Γ pertencem ao núcleo de $C(1)$, e o posto de $C(1)$ tem dimensão $k-r$.

Podemos ter quatro casos, conforme Swamy et allii (1989), que são:

(i) $C(1)$ é simétrica com todos os autovalores distintos;

(ii) $C(1)$ é simétrica e seu autovalor zero tem multiplicidade $r > 1$;

(iii) $C(1)$ não é simétrica e todos autovalores são distintos e

(iv) $C(1)$ não é simétrica e seu autovalor zero tem multiplicidade maior que zero.

Caso (i):

Como os autovalores de $C(1)$ são distintos e, por (14), o posto de $C(1)$ é menor do que k, existe um e somente um autovalor zero e, portanto, a dimensão de Γ é um, isto é, $r = 1$. Isto mostra que neste caso o posto de $C(1)$ é $k-1$. Mas o vetor α não é único, já que se α é autovetor de $C(1)$, associado ao autovalor zero, $-\alpha$ também será autovetor associado ao autovalor zero. Desta forma o vetor co-integrado não é único.

Caso (ii):

Como a multiplicidade do autovalor zero é $r > 1$, tem-se que a matriz $\Gamma, k \times r$, e a matriz de autovetores associados ao autovalor zero, e esta matriz tem posto igual a multiplicidade do autovalor zero, isto é, r. Logo o posto de $C(1)$ é $k - r$. Como no caso anterior, os autovetores não são únicos, o que implica que a matriz Γ não é única.

Caso (iii):

No caso de matrizes simétricas, tem-se que os autovetores à direita e à esquerda são iguais, mas isto não é verdade para matrizes não-simétricas. Seja, então, P_1 e P_2 duas matrizes ortogonais, tais que $P_1^{-1}C(1)P_1 = P_2^{-1}C'(1)P_2 = \Lambda$, matriz de autovalores de $C(1)$. P_1 e P_2 correspondem às matrizes de autovetores à direita e à esquerda, respectivamente. Como os autovalores de $C(1)$ são todos distintos e o posto de $C(1)$ tem que ser $k - 1$, e P_1 e P_2 são vetores $k \times 1$ distintos, o vetor co-integrado não é único, já que α pode ser dado pelo autovetor à direita que corresponde ao autovalor zero de $C(1)$, ou ao autovetor à esquerda que corresponde ao autovalor zero de $C'(1)$, e estes dois vetores são distintos.

Caso (iv):

Como a multiplicidade da raiz zero é $r > 1$, tem-se uma indeterminação para a ordem da matriz Γ, já que o número de autovalores não-zero não é igual ao posto de $C(1)$. Isto se deve ao fato de que os autovetores de $C(1)$ associados ao autovalor λ, juntamente com o valor nulo, geram um subespaço. Este subespaço é chamado de auto-espaço de $C(1)$, associado ao autovalor λ, e a dimensão deste subespaço é chamada de multiplicidade geométrica de λ. Por outro lado, a multiplicidade da λ, como raiz da equação $|C(1) - \lambda I| = 0$ é chamada de multiplicidade algébrica de λ. O número de colunas de Γ é igual à multiplicidade geométrica de λ, que não é necessariamente igual à multiplicidade algébrica do autovalor zero de $C(1)$. Pode-se mostrar que a multiplicidade algébrica é maior ou igual à multiplicidade geométrica de cada autovalor de $C(1)$, sendo igual quando a matriz $C(1)$ é semi-simples (ver Rao 1973, p.77), e como necessita-se da multiplicidade geométrica do autovalor zero de $C(1)$ para obter a ordem de Γ, tem-se uma indeterminação. Neste caso, além de não se ter a unicidade de Γ, não se tem também a existência.

Uma forma de garantir a existência de Γ no Caso (iv), é usar a matriz $C(1)$ $C'(1)$, já que esta matriz é simétrica e, desta forma, recai-se no caso (ii).

$2^{\underline{o}}$ Caso – Representação AR:

Fazendo-se $d = 1$ em (8), tem-se que o processo $(Y_t)_0^\infty$ pode ser representado por um processo AR, isto é:

$$A(L)Y_t = \varepsilon_t \tag{15}$$

e $A(L)$ é definido por:

$$A(L)C(L) = 1 - L \Rightarrow A(1)C(1) = 0 \tag{16}$$

e, portanto, as colunas de $A(1)$ pertencem ao núcleo de $C(1)$, e o posto de $A(1)$ é r. Observe que como $A(1)$ e Núcleo $(C(1))$, existe B, matriz $k \times r$, tal que

$$A(1) = B\Gamma' \tag{17}$$

Primeiro, dado (8) temos que mostrar que (15) existe. Tem-se o seguinte Lema:

Lema 1. *Seja $C(L)$ uma matriz polinominal racional que é finita para todo L no e dentro do círculo unitário, então:*

$$M(L) = U(L)C(L)V(L) \tag{18}$$

onde:

(i) Todas as raízes de $\det[U(L)] = 0$ e $\det[V(L)] = 0$ estão fora do círculo unitário e

(ii) $M(L)$ é diagonal com todas as raízes de $\det[M(L)] = 0$ no ou dentro do círculo unitário.

DEMONSTRAÇÃO: Ver Yoo (1986): Por (18) segue-se que a matriz $C(L)$ da representação (8) pode ser escrita da seguinte forma:

$$C(L) = U^{-1}(L)M(L)V^{-1}(L) \tag{19}$$

se $C(L)$ não tiver pólos, e porque $U(L)$ e $V(L)$ são inversíveis. Usando-se (19), (8) pode ser escrito da seguinte forma:

$$(1 - L)Y_t = U^{-1}(L)M(L)V^{-1}(L)\varepsilon_t \tag{20}$$

Como o posto de $C(1)$ é $k - r$, tem-se que o posto de $M(1)$ também é $k - r$. Pela definição de $M(L)$ no Lema 1 acima, tem-se que esta matriz pode ser expressa da seguinte forma:

$$M(L) = \begin{bmatrix} I_{k-r} & 0 \\ 0 & \Delta I_r \end{bmatrix} \tag{21}$$

Pré-multiplicando-se (20) por $U(L)$ obtém-se:

$$(1 - L)U(L)Y_t = M(L)V^{-1}(L)\varepsilon_t \tag{22}$$

onde as últimas r linhas de (22) tem $(1 - L)$ em ambos os lados; portanto, (22) pode ser reescrito da seguinte forma:

$$\overline{M}(L)U(L)Y_t = V^{-1}(L)\varepsilon_t \tag{23}$$

onde $\overline{M}(L)$ é definido por:

$$\overline{M}(L) = \begin{bmatrix} \Delta I_{k-r} & 0 \\ 0 & I_r \end{bmatrix} \tag{24}$$

e finalmente pré-multiplicando (23) por $V(L)$, obtém-se:

$$A(L)Y_t = V(L)\overline{M}(L)U(L)Y_t = \varepsilon_t \tag{25}$$

que é a representação AR dada por (15) o que prova a existência desta representação.

Pela definição de $A(L)$ em (25) e de $C(L)$ em (19) é fácil mostrar (16). Também pela definição das matrizes $U(L)$, $M(L)$ e $V(L)$ tem-se que $A(L)$ tem posto completo para todo L dentro do círculo unitário. Agora, para $L = 1$, o posto de $A(1)$ é dado pelo posto de $\overline{M}(1)$ que é r. Como $A(1) \in$ Núcleo $(C(1))$ e Γ é a matriz de geradores deste subespaço, existe B, matriz $k \times r$, tal que $A(1) = B\Gamma'$, o que prova (17).

Como $A(1)$ é singular, a representação Auto-regressiva Vetorial, que será denotada daqui por diante por VAR, será tal que os procedimentos numéricos utilizados para calcular a representação MA serão altamente instáveis, indicando multicolinearidade entre as variáveis, o que dificulta a análise dos multiplicadores nas funções impulso-resposta.

3º Caso – Representação MCE:

Considere a representação AR do caso anterior dada pela expressão (15). Particionando-se V e U da mesma forma que M, tem-se:

$$V(L) = [V_1(L)|\gamma(L)] \tag{26}$$

$$U(L) = [U_1(L)|\alpha(L)] \tag{27}$$

e

$$A(1) = V(1)\overline{M}(1)U(1) = \gamma(1)\alpha'(1) \equiv \gamma\alpha' \tag{28}$$

onde $\gamma \equiv \gamma(1)$ e $\alpha \equiv \alpha(1)$. Como U e V têm posto completo para todo L dentro ou no círculo unitário, $U(1)$ e $V(1)$ têm posto completo e, portanto, γ e α também têm posto completo, que é igual a r. Agora, fazendo-se o desenvolvimento de Taylor até primeira ordem para $A(L)$, na vizinhança de $L = 1$, obtém-se:

$$A(L) = A(1) + A^{**}(L)(1 - L) = A(1) + A^{**}(L)\Delta \tag{29}$$

onde as raízes de $\det[A^{**}(L)]$ estão fora do círculo unitário. Por (28) e (29) obtém-se a seguinte representação para o processo Y_t:

$$A(L)Y_t = [A(1) + A^{**}(L)\Delta]Y_t = \varepsilon_t$$

$$A^{***}(L)\Delta Y_t = -\gamma\alpha' Y_{t-1} + \varepsilon_t \tag{30}$$

onde $A^{***}(L) = A(1)+A^{**}(L)$ e observe que $A^{***}(0) = A(1)+A^{**}(0)$ e por (29), como $A(0) = I$, tem-se que $A^{***}(0) = I$.

Observe que as variáveis incluídas nesta representação são todas $I(0)$. Observe que a variável dependente em (30) é um VAR nas variações de Y_t, que são variáveis $I(0)$, e a variável do lado direito inclui combinações lineares do nível de Y_t, que são estacionárias. Isto é, $z_t = \alpha' Y_t$ corresponde àquela combinação linear de Y_t onde os pesos são o vetor co-integrado.

O modelo (30) é conhecido na literatura como M.C.E. e esta representação é devida a Granger (1983). Observe que todas as variáveis, condicionais em α, são estacionárias e, portanto, a estimação e inferência dos parâmetros A^* e γ, podem ser feitas utilizando-se os métodos usuais, ou seja, a teoria assintótica convencional vale neste caso, já que a distribuição dos parâmetros é normal. Stock

(1987) mostrou que além disto, os parâmetros estimados convergem rapidamente para as suas distribuições limites e as estimativas são consistentes com viés de ordem T^{-1}.

Se não houvesse co-integração, então o termo de correção dos erros não estaria presente em (30). Por outro lado, se é assumido que os dados seguem um VAR em diferenças, fica implícito que as variáveis não co-integram e, portanto, o modelo estará mal-especificado caso haja co-integração. Neste caso, os modelos VAR em níveis, e que ignoram as restrições dadas por (17) apresentam problemas quando se deseja obter a representação MA, tal como foi comentado naquele caso.

Sims, Stock e Watson (1987) mostram que num modelo VAR onde existem relações de co-integração entre as variáveis e, sempre que qualquer das variáveis puder ser escrita como uma variável estacionária com média zero, isto é, a ausência de tendência determinística no modelo, os coeficientes estimados terão uma distribuição assintótica normal.

A representação (30) ilustra também a conexão existente entre o conceito de co-integração e o de causalidade no sentido de Granger. Por exemplo, num sistema bivariado, se o segundo elemento da matriz B, que neste caso é uma matriz 2×1, for nulo, deverá existir causalidade na direção $Y_{1t} \rightarrow Y_{2t}$, ou se o primeiro elemento de B for nulo, a direção de causalidade será de $Y_{2t} \rightarrow Y_{1t}$. No caso em que ambos os elementos de B são não-nulos, existe causalidade em ambas as direções. Nenhuma das variáveis pode ser considerada como exógena fraca[4] para os parâmetros da outra equação. Neste caso, como existem restrições não-lineares entre as equações, para se estimar os parâmetros eficientemente é necessário utilizar procedimentos de informação completa.

Engle (1987) apresenta uma derivação alternativa para a representação via M.C.E., que será descrita a seguir.

Pela definição de $\overline{M}(L)$, dada por (24), esta matriz pode ser reescrita da seguinte forma:

$$\overline{M}(L) = \Delta I_k + L \begin{bmatrix} 0 & 0 \\ 0 & I_k \end{bmatrix} \tag{31}$$

[4]Para uma definição de exogeneidade fraca veja Engle, Hendry e Richard (1983).

e portanto, (25) pode ser escrito da seguinte forma:

$$V(L)\ U(L)\Delta Y_t = -\gamma(L)\ \alpha'(L)Y_{t-1} + \varepsilon_t \tag{32}$$

que consiste num polinômio vetorial co-integrado, primeiramente sugerido por Yoo (1986). Observe que este modelo pode ser interpretado como uma extensão da classe de modelos Auto-regressivos-Defasagens Distribuídas sugeridos por Hendry, Pagan e Sargan (1984). Observe que na representação (32) é assumido que $V(L)\ U(L)$ é estável, para que a parte auto-regressiva da representação seja $I(1)$.

4º Caso – Representação tendência comum:

A idéia desta representação, devido a Stock e Watson (1988), é o fato estilizado de que séries temporais macroeconômicas podem ser descritas por modelos ARIMA univariados e, portanto, é necessário diferenciá-las para se obter estacionariedade. Mas, quando é considerado um modelo multivariado, o número de raízes unitárias na série temporal multivariada pode ser menor do que o número de raízes unitárias das séries univariadas que constituem o modelo. Uma outra explicação para este fato é que, embora as séries univariadas contenham tendências estocásticas, estas tendências podem ser comuns a várias das variáveis que compõem o processo vetorial e estas tendências comuns desapareceriam quando usadas certas combinações das variáveis.

Seja a representação MA, dada por (13),[5] mas onde é permitido que o processo tenha média, μ, não-nula. Tem-se então:

$$\Delta Y_t = \mu + C(L)\varepsilon_t \tag{33}$$

onde $C(z) = \sum_{i=0}^{\infty} C_i\ z^i$ com $C(0) = I_k$, ε_t é uma seqüência de vetores aleatórios *i.i.d.* com média zero e matriz de covariância G. Assume-se que $C(1)$ tem posto $q < k$ e, logo, Y_t é co-integrado, isto é, existe uma matriz $k \times r, \alpha$ (onde $r = k - q$) tal que $\alpha' C(1) = 0$ e $\alpha'\mu = 0$.

[5]Adotou-se, em (13), as condições sobre a seqüência $\{\varepsilon_t\}$ e $C(L)$ que serão modificadas. Pode-se mostrar, (ver Hannan 1970) que as duas representações são equivalentes.

Seja $\nu_t = G^{-1/2}\varepsilon_t$ e $\xi_t = \sum_{s=1}^{t} \nu_s$. Adotando-se a hipótese de que $\varepsilon_s = 0$ para $s \leq 0$, mas permitindo-se que Y_t tenha um valor inicial não-aleatório Y_0, pode-se, após sucessivas substituições, reescrever (33) da seguinte forma:

$$Y_t = Y_0 + \mu t + C(1)G^{1/2}\xi_t + C^{**}(L)G^{1/2}\nu_t \tag{34}$$

onde $C^{**}(L) = (1-L)^{-1}(C(L) - C(1))$ e $C_j^{**} = \sum_{i=j+1}^{\infty} C_i$. Como $\alpha' C(1) = 0$ e $\alpha'\mu = 0$, tem-se que:

$$z_t = \alpha' Y_t = \alpha' Y_0 + \alpha' C^{**}(L)G^{1/2}\nu_t \tag{35}$$

e assumindo-se que $C(L)$ é 1-somável, segue-se que $C^{**}(L)$ é absolutamente somável e, desta forma, z_t terá variância limitada.

Uma representação alternativa a (34), que explicita o processo co-integrado Y_t em termos de um número menor de passeios aleatórios comuns e de um componente adicional estacionário, será apresentada a seguir.

Stock e Watson (1988) observam que como $C(1)$ tem posto $q < k$ existe uma matriz H_1, com posto $r = k-q$, tal que $C(1)H_1 = 0$. Além disto, se H_2 é uma matriz, $k \times q$, com posto q e com colunas ortogonais às de H_1, então a matriz $A = C(1)H_2$ tem posto q. Considere a matriz, $k \times k, H$ definida por $H = (H_1 \mid H_2)$ que é não-singular e observe que $C(1)H = (0 \mid A)$. Defina, então, a matriz, $q \times k$, de seleção S_q por $(0_{k-q} \mid I_q)$ então $C(1)H = AS_q$. Mas como Y_t co-integra, existe α tal que, $\alpha' C(1) = 0$ e $\alpha'\mu = 0$. Portanto, μ pertence ao subespaço das colunas de $C(1)$ e pode ser escrito da seguinte forma: $\mu = C(1)\tilde{\mu}$. Agora (34) é dado por:

$$Y_t = Y_0 + C(1)[\tilde{\mu}t + G^{1/2}\xi_t] + C^{**}(L)G^{1/2}\nu_t$$

$$Y_t = Y_0 + C(1)H[H^{-1}\tilde{\mu}t + H^{-1}G^{1/2}\xi_t] + a_t \tag{36}$$

com $a_t = C^{**}(L)G^{1/2}\nu_t$.

Definindo-se $t_t = S_qH^{-1}\tilde{\mu}t + S_qH^{-1}G^{1/2}\xi_t, \pi = S_qH^{-1}\tilde{\mu}$ e $n_t = S_qH^{-1}G^{1/2}\nu_t$, pode-se reescrever (36) da seguinte forma:

$$Y_t = Y_0 + A\tau_t + a_t \tag{37}$$

$$\tau_t = \pi + \tau_{t-1} + n_t \tag{38}$$

que expressa Y_t como uma combinação linear de k passeios aleatórios (τ_t), com *drift* π e mais um componente transitório, a_t, não-integrado. Pelas equações (37–38), fica evidente a denominação Tendência Comum desta representação.

Apresenta-se, a seguir, um exemplo que procura mostrar as diferenças entre as quatro representações de processo co-integrados.

EXEMPLO 3: Suponha que $Y_t = (Y_{1t}, Y_{2t})$ tenha a seguinte representação MA

$$(1 - L)\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = (1 - 0,68L)^{-1}\begin{bmatrix} 1 - 0,84L & 0,4L \\ 0,064L & 1 - 0,84L \end{bmatrix}\varepsilon_t$$

Observe que ambos os componentes de Y_t são $I(1)$ e $CI(1,1)$, já que a matriz $C(1)$ é dada por:

$$C(1) = \begin{bmatrix} 0,5 & 1,25 \\ 0,2 & 0,5 \end{bmatrix} \Longrightarrow \text{Posto } [C(1)] = 1$$

e, portanto, o vetor $\alpha' = (1, -0,4)$, o autovetor à direita associado ao autovalor zero é um dos vetores co-integrados. É fácil mostrar que $\alpha = (1, -2,5)$ é o autovetor à esquerda associado ao autovalor zero. Portanto, não se tem unicidade do vetor co-integrado, o que corresponde ao segundo caso estudado na representação MA, já que $C(1)$ não é simétrica e tem autovalores distintos, que são $\lambda = 0$ e $\lambda = 1$.

Invertendo-se a matriz MA, obtém-se a representação AR, que é dada por:

$$\begin{bmatrix} 1 - 0,84L & -0,4L \\ -0,064L & 1 - 0,84L \end{bmatrix}\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \varepsilon_t$$

e observe que $A(L)\ C(L) = 1 - L$ e que $A(1)$ é dada por:

$$A(1) = \begin{bmatrix} 0,16 & -0,4 \\ -0,064 & 0,16 \end{bmatrix} \Longrightarrow \text{Posto } [A(1)] = 1$$

e observe que esta matriz é dada pelo produto do vetor $B = \begin{bmatrix} 0,16 \\ -0,064 \end{bmatrix}$ pelo autovetor à esquerda de $C(1)$, isto é, $\alpha' = [1, -2,5]$.

A representação M.C.E. é, portanto, dada por:

$$(1-L) = \begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} 0,16 \\ -0,064 \end{bmatrix} [1, -2,5] \begin{bmatrix} Y_{1t-1} \\ Y_{2t-1} \end{bmatrix} + \varepsilon_t$$

e neste caso existe causalidade em ambas direções, já que os dois componentes de B são diferentes de zero.

A representação Tendência Comum é dada por:

$$Y_t = Y_0 + \begin{bmatrix} 1,54 \\ 0,58 \end{bmatrix} \tau_t + a_t$$

$$\tau_t = \tau_{t-1} + \eta_t$$

com $\eta_t = [0,1] \begin{bmatrix} 1,0 & 0,4 \\ 0,4 & 1,0 \end{bmatrix}^{-1} \varepsilon_t$ e $a_t = C^*(L)\varepsilon_t$ com $C_j^* = \sum_{i=j+1}^{\infty} C_i$ e $C_i = (0,68)^i \begin{bmatrix} -0,5 & 1,25 \\ 0,2 & -0,5 \end{bmatrix}$.

4. Extensão dos teoremas de representação.

Esta seção apresenta os teoremas de representação, com especial atenção ao do M.C.E., para os casos em que o vetor Y_t tem mais de uma raiz unitária ou possui raízes unitárias nas freqüências sazonais.

Considere o caso de d raízes unitárias, isto é, a representação MA é dada por:

$$(1-L)^d Y_t = C(L)\varepsilon_t \tag{39}$$

onde $C(L)$ tem o seguinte desenvolvimento:

$$C(L) = \gamma_1 + \gamma_2(1-L) + \cdots + \gamma_d(1-L)^{d-1} + C^*(L)\ (1-L)^d \tag{40}$$

com $\gamma_i = \dfrac{(-1)^{i-1}C^{(i-1)}(1)}{(i-1)!}, i = 1, \cdots, d$ e $C^*(L)$ invertível.

Suponha que existe um vetor co-integrado α, tal que:

$$\alpha'\gamma_i = 0 \qquad \text{para} \qquad i = 0, \cdots, d \tag{41}$$

Pela representação VAR, obtém-se que:

$$A(L)\ C(L) = (1-L)^d \Longrightarrow A(1)\ C(1) = 0 \tag{41}$$

o que implica que existe uma matriz D_0 tal que:

$$A(1) = D_0\alpha' \tag{42}$$

Derivando-se (41) com respeito a L e, quando $L = 1$, tem-se a expressão:

$$A^{(1)}(1)\; C(1) + A(1)\; C^{(1)}(1) = 0 \tag{43}$$

e usando-se (42) em (43) obtém-se:

$$A^{(1)}(1)\; C(1) + B\alpha' C^{(1)}(1) = 0 \tag{44}$$

mas, como γ_2 é igual a $C^{(1)}(1)$, a menos de constante, e por (41), tem-se que a segunda expressão do lado esquerdo da igualdade em (44) é zero, o que implica que (44) reduz-se a:

$$A^{(1)}(1)\; C(1) = 0 \tag{45}$$

Portanto, $A^{(1)}(1)$ também pertence ao Núcleo $[C(1)]$, logo existe uma matriz D_1, tal que:

$$A^{(1)}(1) = D_1\alpha' \tag{46}$$

Derivando-se (45) com respeito a L e quando $L = 1$, tem-se a expressão:

$$A^{(2)}(1)\; C(1) + A^{(1)}(1)\; C^{(1)}(1) = 0 \tag{47}$$

que é semelhante a (43) e, por argumentos análogos aos utilizados acima, obtém-se:

$$A^{(2)}(1)\; C(1) = 0 \tag{48}$$

e, portanto, $A^{(2)}(1)$ também pertence ao Núcleo $[C(1)]$ e logo, existe uma matriz D_2 tal que:

$$A^{(2)}(1) = D_2\alpha' \tag{49}$$

e, por indução, existem matrizes D_j tal que:

$$A^{(j)}(1) = D_j\alpha' \quad \text{para } j = 1, \cdots, d. \tag{50}$$

que podem ser também escritas da seguinte forma:

$$\frac{(-1)^j A^{(j)}(1)}{j!} = D_j \alpha'$$

Agora a representação VAR é dada por:

$$A(L)Y_t = \left[\sum_{j=0}^{d-1} D_j \alpha'(1-L)^j + A^{**}(L)(1-L)^d\right] Y_t = \varepsilon_t \tag{51}$$

onde $A^{**}(L)$ tem raízes fora do círculo unitário. Definindo-se

$$A^{***}(L) = A^{**}(L) + \sum_{j=0}^{d-1} D_j \alpha' \tag{52}$$

e, como $A^{***}(0) = I$, após alguma álgebra, (51) é equivalente a:

$$\begin{aligned} A^{**}(L)\ (1-L)^d Y_t = &- \left[\sum_{j=0}^{d-1} D_j \alpha'\right] (1-L)^{d-1} Y_{t-1} \\ &- \left[\sum_{j=0}^{d-2} D_j \alpha'\right] (1-L)^{d-2} Y_{t-1} - \cdots - D_0 \alpha' Y_{t-1} \end{aligned} \tag{53}$$

que é a representação M.C.E.

Considera-se a seguir o caso de raízes unitárias em freqüências sazonais, mais particularmente freqüências semestrais, isto é, $(1-L^2)$.

Para o caso de freqüências trimestrais, (ver Hylleberg *et allii* 1988) e para o caso de freqüências regulares e sazonais, (ver Osborn *et allii* 1988).

Considere a representação de Wold que é dada por:

$$(1-L^2)Y_t = C(L)\varepsilon_t \tag{54}$$

Observe que a matriz $C(L)$ pode ser dada por:

$$C(L) = \gamma_1(1-L) + \gamma_2(1+L) + C^{**}(L)(1-L^2) \tag{55}$$

onde $\gamma_1 = C(-1)/2, \gamma_2 = C(1)/2$ e $C^{**}(L)$ é invertível.

Pela representação VAR tem-se a seguinte condição:

$$A(L)\,C(L) = (1-L)(1+L) \Longrightarrow A(1)\,C(1) = A(-1)\,C(-1) = 0 \quad (56)$$

Haverá co-integração na raiz $(1-L)$ se existir um vetor co-integrado α_1 tal que $\alpha_1'\gamma_1 = \alpha_1'C(-1) = 0$. Similarmente, haverá co-integração na raiz $(1-L)$ se existir um vetor α_2 tal que $\alpha_2'\gamma_2 = \alpha_2'C(1) = 0$. Pode-se, portanto, expressar $A(-1)/2 = D_1\alpha_1'$ e $A(1)/2 = D_2\alpha_2'$.

Pode-se utilizar um desenvolvimento para $A(L)$ semelhante a (55) e obtém-se a representação M.C.E., que será dada por:

$$A^{**}(L)\Delta_2 Y_t = D_2\alpha_2'\Delta Y_{t-1} - D_1\alpha_1'(Y_{t-1} + Y_{t-2}) + \varepsilon_t \quad (57)$$

com $A^{**}(0) = I$.

Observe que se não existir co-integração em $(1-L)$, então $\alpha_1 = 0$ e o segundo termo do lado direito da expressão (57) desaparece. Similarmente, se não houver co-integração em $(1+L)$, $\alpha_2 = 0$ e o primeiro termo do lado direito da expressão (57) desaparece.

Engle (1987) propõe uma representação que permite generalizar os dois casos apresentados acima. Esta representação segue a decomposição de $C(L)$ dada no Lema 1 e obtém-se a representação M.C.E. para modelos com várias raízes unitárias ou raízes em freqüências sazonais ou, ainda, componentes de tendência determinísticas do tipo polinômio de ordem maior ou igual a um.

Sejam $\delta_1(L)$ e $\delta_2(L)$ dois polinômios no operador defasagem L, com $\delta_1(L) = 0$ para $L = \theta_1$ e $\delta_2(L) = 0$ para $L = \theta_2$, onde θ_1 e θ_2 podem ser raízes complexas que podem estar no ou dentro do círculo unitário. Admite-se que o processo $(Y_t)_0^\infty$ tem não-estacionariedade devida a estas raízes e, portanto, diferenciando o processo pelo produto dos polinômios $\delta_1(L)\delta_2(L)$, obtém-se uma representação Média-Móvel finita dada por:

$$\delta_1(L)\delta_2(L)Y_t = C(L)\varepsilon_t \quad (58)$$

Observe que os componentes do processo vetorial Y_t podem ter ordem de integração diferentes e os casos anteriores são formas particulares de (58). Por exemplo, (39) com $d = 2$ corresponde ao caso em que $\delta_1(L) = \delta_2(L) = \Delta = 1 - L$ e (54) corresponde ao caso em que $\delta_1(L) = \Delta$ e $\delta_2(L) = (1 + L)$ sendo que este último pode ser estendido para periodicidade sazonal igual a s com $\delta_1(L) = \Delta$ e $\delta_2(L) = (1 + L + L^2 + \cdots + L^{s-1})$.

Pela decomposição de $C(L)$, dada pelo Lema 1, tem-se que (58) pode ser escrito da seguinte forma:

$$\delta_1(L)\delta_2(L)Y_t = U^{-1}(L)M(L)V^{-1}(L)\varepsilon_t \tag{59}$$

com

$$M(L) = \begin{bmatrix} I_{k-r_1-r_2-r_3} & & & \\ & \delta_1(L)I_{r_1} & & \\ & & \delta_2(L)I_{r_2} & \\ & & & \delta_1(L)\delta_2(L)I_{r_3} \end{bmatrix} \tag{60}$$

e observe que o Posto $[C(\theta_1)] = k - r_1 - r_3$ e Posto $[C(\theta_2)] = k - r_2 - r_3$. Multiplicando-se ambos os lados de (59) por $U(L)$, obtém-se a expressão:

$$\overline{M}(L)\ U(L)Y_t = V^{-1}(L)\varepsilon_t \tag{61}$$

com

$$\overline{M}(L) = \begin{bmatrix} \delta_1(L)\delta_2(L)I_{k-r_1-r_2-r_3} & & & \\ & \delta_3(L)I_{r_1} & & \\ & & \delta_1(L)I_{r_2} & \\ & & & I_{r_3} \end{bmatrix} \tag{62}$$

e particionando $U(L) = [U_1(L), \alpha_1(L), \alpha_2(L), \alpha_3(L)]$, (61) pode ser escrito da seguinte forma:

$$\begin{bmatrix} U_1' \quad (L)\delta_1(L)\delta_2(L)Y_t \\ \alpha_1'(L)\delta_2(L)Y_t \\ \alpha_2'(L)\delta_1(L)Y_t \\ \alpha_3'(L)Y_t \end{bmatrix} = V^{-1}(L)\varepsilon_t \tag{63}$$

e cada coluna de (63) representa um processo estacionário. Por exemplo, $\alpha_1(L)$ elimina a raiz θ_1 e quando multiplicado por $\delta_2(L)$, nenhum componente não-estacionário permanece neste componente do processo vetorial. O último componente do vetor definido em (63) implica que $\alpha_3(L)$ elimina ambas as raízes θ_1 e θ_2. Observe também que pela definição de $U(L)$, todas as submatrizes de $U(L)$ devem ter posto completo.

Observe que o vetor co-integrado definido em (63) é do tipo polinômio co-integrado, mas é sempre possível obter um vetor co-integrado, substituindo-se o valor da raiz relevante. Por exemplo, para $\alpha_1(L)$, tem-se:

$$\alpha_1(L) = \alpha_1(\theta_1) + \delta_1(L)\alpha_1^*(L) \tag{64}$$

que define $\alpha_1^*(L)$. Agora, substituindo-se (64) no segundo componente de (63), obtém-se:

$$\alpha_1'(L)\delta_2(L)Y_t = \alpha_1'(\theta_1)\delta_2(L)Y_t + \alpha_1^{*\prime}(L)\delta_1'(L)\delta_2(L)Y_t \tag{65}$$

e como o lado esquerdo da igualdade e o segundo termo do lado direito da desigualdade em (65) são $I(0)$, o primeiro termo do lado direito da igualdade também tem que ser $I(0)$ e, portanto, um vetor co-integrado com coeficientes constantes é obtido e tem as mesmas propriedades que o polinômio co-integrado.

O argumento acima vale também para a outra raiz, isto é, para o outro polinômio co-integrado $\alpha_2(L)$, mas não vale para o polinômio $\alpha_3(L)$, já que este depende das duas raízes. Pode-se, no entanto, obter uma expansão alternativa para $\alpha_3(L)$ que permite obter um vetor co-integrado através deste polinômio co-integrado. Esta expansão é dada por:

$$\alpha_3'(L) = \frac{\alpha_3'(\theta_1)\delta_2(L)}{\delta_2(\theta_1)} + \frac{\alpha_3'(\theta_2)\delta_1(L)}{\delta_1(\theta_2)} + \alpha_3^{*\prime}(L)\delta_1(L)\delta_2(L) \tag{66}$$

e observe que quando $L = \theta_1$ ou $L = \theta_2$, (66) é satisfeita. Como $\alpha_3'(L)Y_t$ é $I(0)$ e $\alpha_3^{*\prime}(L)\delta_1(L)\delta_2(L)Y_t$ também é $I(0)$, a soma dos dois primeiros termos do lado direito de (66) também são $I(0)$, mas não se pode garantir a existência de um vetor co-integrado constante. Esta é uma característica dos polinômios co-integrados quando se elimina mais de uma raiz.

Pelo sistema (63), pode-se obter a representação M.C.E. Observe que ao se pré-multiplicar (61) por $V(L)$ e separando-se $\overline{M}(L)$ em dois componentes, obtém-se:

$$V(L)U(L)\delta_1(L)\delta_2(L)Y_t = V(L)[\delta_1(L)\delta_2(L)I_k - \overline{M}(L)]Y_t + \varepsilon_t \tag{67}$$

e pela partição de $V(L) = [V_1(L), \gamma_1(L), \gamma_2(L), \gamma_3(L)]$ a representação M.C.E. é dada por:

$$\begin{aligned} V(L)U(L)\delta_1(L)\delta_2(L)Y_t =& \gamma_1(L)\alpha_1'(L)\delta_2(L)(1-\delta_1(L))Y_t + \\ &+ \gamma_2(L)\alpha_2'(L)\delta_1(L)(1-\delta_2(L))Y_t + \\ &+ \gamma_3(L)\alpha_3'(L)(1-\delta_1(L)\delta_2(L))Y_t + \\ &+ \varepsilon_t \end{aligned} \tag{68}$$

onde os três termos de correção dos erros estão presentes. Observe que com $\delta_1(0) = 1$, para $i = 1, 2$, todos os três termos só dependem de Y_t defasados.

5. Conclusões.

Séries macroeconômicas são, em geral, não-estacionárias, já que apresentam algum tipo de tendência na média e/ou na variância. A inflação no Brasil medida, por exemplo, pelo IPA-DI, tem um comportamento bastante ilustrativo. No período de 1966 a 1973, observa-se um patamar inflacionário de 20%, passando a um patamar de 35% no período de 1974–1979, mudando para um de 70% no período de 1980–82 e atingindo um de 120% no período de 1983–85. Pode-se também observar que a média no período de 1973–85 cresce, continuamente, de 20% até 50% e a variância, neste mesmo período, cresce de 9% até 1200%, implicando que ambas as medidas não são invariantes no tempo, caracterizando-se, portanto, algum tipo de não-estacionariedade nesta série.

Granger (1983) formalizou o conceito de variável ser integrada de ordem d (denotada por $I(d)$) se a d-ésima diferença for estacionária. Na prática, em geral, $d = 1$ é suficiente para variáveis reais e $d = 2$ para variáveis nominais. No caso da inflação pode-se mostrar, veja Pereira (1988), que esta série é $I(1)$, isto é, o nível dos preços é $I(2)$. Este procedimento, diferenciar a série para induzir estacionariedade, é o mais usado na literatura de séries temporais, mas como foi mostrado na seção 2, esta transformação pode induzir um pólo, isto é, um valor ilimitado na freqüência zero do espectro se a série for subdiferenciada ou um zero na freqüência zero do espectro se a série for superdiferenciada. Estes resultados, como também foi visto na seção 2, também são válidos ao se trabalhar com séries multivariadas. Foi também observado que uma série temporal multivariada pode ser decomposta em vários componentes integrados de ordem decrescente, a partir da ordem d, conforme (10). Por exemplo, ao se considerar o vetor $Y_t = (w_t, p_t)$ onde w_t representa os salários e p_t o nível dos preços e supondo que $Y_t \sim I(2)$, podemos representar este vetor bivariado pela soma de três componentes a saber, um cuja ordem de integração é 2, outro cuja ordem é 1 e um terceiro que é estacionário. Através desta representação, pode-se mostrar que, embora a série seja não-estacionária, pode existir uma combinação linear dos componentes de Y_t de tal sorte que alguns ou todos os componentes integrados desapareçam. Desta forma, pode existir uma combinação linear dos componentes que seja estacionária. Por exemplo, embora w_t e p_t sejam $I(2)$, pode existir um $\delta = (1, -\lambda)$ tal que $w_t - \lambda p_t = u_t$ seja estacionário. Granger denotou esta propriedade de séries temporais por co-integração. A intuição econômica deste conceito é dada pela relação de equilíbrio entre as variáveis, por exemplo, no equilíbrio $w = \lambda p$ e $u_t = w_t - \lambda p_t$ são os desvios deste equilíbrio que, como w_t e p_t co-integram, devem ser limitados.

Uma outra implicação do conceito de co-integração é a possibilidade de identificação em modelos de equações simultâneas, conforme exemplo 1 apresentado no texto. Aquele exemplo pode ser racionalizado pelo modelo de determinação de quantidades e preços num mercado em equilíbrio que, na econometria usual, é um exemplo clássico de falta de identificação de modelos de equações simultâneas. Mas, como foi observado, se ambas as variáveis forem $I(1)$ e, sob certas condições sobre os erros das duas equações, pode-se ter a segunda equação, que pode ser interpretada como sendo de demanda, identificável.

Dado que os componentes do processo estocástico k-variado co-integram, pode-se reescrever este processo de quatro formas: representação MA, representação AR, representação M.C.E. e de tendência comum. A primeira destas representações é utilizada para se obter respostas a choque na parte não-sistemática do modelo. A segunda, representação AR, é utilizada em modelos VAR e, como foi observado, caso não se leve em consideração a existência de co-integração, pode-se ter problemas na parte de estimação dos parâmetros, já que os procedimentos numéricos são altamente instáveis, comprometendo a análise de multiplicadores nas funções impulso-resposta. A terceira representação M.C.E., que é dada pela expressão (30) do artigo, tem uma racionalidade econômica que torna a representação mais interessante. Como todas as variáveis que aparecem naquela representação são estacionárias, tem-se que $\alpha' Y_{t-1}$ pode ser interpretado como o vetor de desvios do equilíbrio e como estes desvios são estacionários, a dinâmica de curto prazo será dada pelas defasagens de ΔY_t e a matriz γ pode ser interpretada como uma medida da rapidez pela qual o sistema corrige os desvios do equilíbrio no período anterior. Esta representação também mostra que em geral um VAR pode ou não ser apropriado para descrever o processo gerador dos dados, tem-se, então, três casos: (a) se o posto de $A(1)$ é pleno, todas as variáveis são estacionárias e a utilização de um VAR irrestrito em níveis é adequada; (b) quando o posto de $A(1)$ é zero, não há co-integração e a aplicação de um VAR irrestrito em primeiras diferenças é apropriado; (c) mas quando o posto da matriz $A(1)$ é um valor entre 0 e k, existe pelo menos um vetor co-integrado e neste caso a utilização de VAR irrestrito em nível ou em diferenças é inapropriado, isto é, tem uma especificação inadequada do modelo (*misspecified model*). Na última representação de tendência comum, o sistema multivariado tem r vetores co-integrados que, sob certas condições, podem ser escritos como k tendências, Y_0 em (37), mais $k-r$ passeios aleatórios com drifts (as tendências comuns) e n componentes transitórios não-integrados.

Na última seção são apresentadas as extensões para os casos em que se tem mais de uma raiz unitária e/ou raízes unitárias em freqüências sazonais. A racionalidade dada acima também pode ser extendida para estes casos mais complexos.

Embora nesta resenha não se tenha comentado como se deve proceder quanto à estimação e testes de co-integração, apresentam-se as possíveis representações de sistemas multivariados e espera-se que estes procedimentos sejam mais usados, dado que o conceito de co-integração é de fundamental importância e tem sido a tendência recente na análise econômica empírica.

Referências

Baumol, W. J. 1952. "The transactions demand for cash: an inventory theoretic approach." *Quarterly Journal of Economics* **66**: 545–556.

Blanchard, O. & Fisher, S. 1989. *Lectures in Macroeconomic.* Cambridge: MIT Press.

Burns, A. F. & Mitchell, W.C. 1946. *Measuring Business Cycles.* New York: NBER.

Box, G. E. P. & Jenkins, G. M. 1976. *Time Series Analysis.* Forecasting and control. San Francisco: Holden Day.

Deistler M. 1985. "Parametrization of ARMA and state-space systems." In Hannan, E. J., Krishnaiah, P. R. & Rao, M. M., eds., *Handobook of statistics V.* Amsterdam: North Holland.

Engle, R. F. 1987. "On the theory of cointegrated time series." Invited Paper in the European Meeting of the Econometric Society, Copenhagem.

Engle, R. F. & Granger C. W. J. 1987. "Co-integration and error correction: representation, estimation and testing." *Econometrica* **55**: 251–276.

Engel, R. F. , Hendry, D. F. & Richard, J. F 1983. "Exogeneity." *Econometrica* **51**: 277–304.

Granger, C. W. J. 1966. "The typical spectral shape of an economic variable." *Econometrica* **34**: 150–161.

__________ 1981. "Some properties of time series data and their use in econometric model specification." *Journal of Econometrics* **16**: 121-130.

__________ 1983. "Co-integrated variables and error-correction models." Discussion paper, nº 83–13. University of California, San Diego.

Granger, C.W.J. & Newbold, P. 1974. "Spurious regressions in econometrics." *Journal of Econometrics* **2**: 111–120.

Granger, C.W.J. & Weiss, A. A. 1983. "Time series analysis of error correction models." In Karlin, S., Amemiya, T. & Goodman, L. A., eds., *Studies in Econometric Time Series and Multivariate Statistics.* New York: Academic Press.

Hannan, E. J. 1970. *Multiple Time Series.* New York: John Wiley & Sons.

Hendry, D. F. 1987. "Econometric methodology: a personal perspective." In Bewley, T. F., ed., *Advances in Econometrics II.* Cambridge: Cambridge University Press.

Henry, D. F & Ericsson, N. R 1991a. "An econometric analysis of 'U. K. money demand in monetary trends in the United States and the United Kingdom' by Milton Friedmand and Anna J. Schwartz." *American Economic Review* **81**: 8–38.

——————— 1991b. "Modelling the demand for narrow money in the United Kingdom and the United States." In *European Economic Review.* A sair.

Hendry, D. F., Pagan, A. R. & Sargan, J. D. 1984. "Dynamic specification." In Griliches, Z. & Intriligador, M. D., eds., *Handbook of Econometrics II.* Amsterdam: North Holland.

Hylleberg, S., Engle, R. F., Granger, C. W. J. & Yoo, B. S. 1988. "Seasonal integration and cointegration." Discussion paper nº 32. University of California, San Diego.

Miller, M. H. & Orr, D. 1966. "A model of the demand for money by firms." *Quarterly Journal of Economics* 80: 413–435.

Nickell, S. 1985. "Error correction, partial adjustment and all that: an expository note." *Oxford Bulletin of Economics and Statistics* **47**: 119–129.

Osborn, D. R., Chui, A. P. L., Smith, J. P. & Birchenhall, C. R. 1988. "Seasonality and the order of integration for consumption." *Oxford Bulletin of Economics and Statistics* **50**: 361–377.

Phillips, P. C. B. 1986. "Understanding spurious regressions in econometrics." *Journal of Econometrics* **33**: 311–340.

Rao, C. R. 1973. *Linear Statistical Inference and its Applications* 2ª ed. New York: John Wiley & Sons.

Sargan, J. D. 1964. "Wages and Prices in the United Kingdom: a Study in Econometric Methodology." In Hart, P. E. et. al., eds., *Econometric Analysis for National Economic Planning.* London: Butterworth.

Sims, C., Stock, J. H. & Watson, M. W. 1987. "Inference in linear time series models with some unit roots." Hoover Institution, Working Paper in Economics E-87-1.

Stock, J. H. 1987. "Asymptotic properties of least squares estimators of cointegrating vectors." *Econometrica* **55**: 1035–1056.

Stock, J. H. & Watson, M. W. 1988. "Testing for common trends." *Journal of the American Statistical Association* **83**: 1097–1107.

Swamy, P. A. V. B., von zur Muehlen, P. & Mehta, J. S. 1989. "Co-integration: is it a property of the real world?" Finance and Economics Discussion Series, nº 96, Federal Reserve Board Washington.

Yoo, S. 1986. "Multi-cointegrated time series and generalized error-correction models." Economics Department Working Paper. University of California, San Diego.