

**Modelos de análise factorial dinâmica: uma  
abordagem no domínio da frequência**

**Maria do Rosário Fraga de Oliveira Martins  
Sofia Santos Rosa**

**Working Paper n° 79**

**ISSN: 0872-895X**

**Depósito Legal n°: 90631/95**

Instituto Superior de Estatística e Gestão de Informação  
Universidade Nova de Lisboa

Julho de 2003

MODELOS DE ANÁLISE FACTORIAL DINÂMICA:

*Uma abordagem no domínio da frequência*

Autores: Maria do Rosário Fraga de Oliveira Martins  
Sofia Santos Rosa

# Índice

Capítulo 0:	Introdução .....	0
Capítulo 1:	Análise de Fourier .....	1-11
1.1.	Introdução	
1.2.	Funções Ortogonais	
1.3.	Representação de Fourier de Funções Temporais Discretas Finitas	
1.4.	Representação de Fourier de Funções Temporais Discretas Periódicas	
1.5.	Representação de Fourier de Funções Temporais Discretas Não Periódicas – A Transformada de Fourier em Tempo Discreto	
1.6.	Representação de Fourier de Funções em Tempo Contínuo	
1.6.1.	Representação de Fourier de Funções Periódicas	
1.6.2.	Representação de Fourier de Funções Não Periódicas – A Transformada de Fourier em Tempo Contínuo	
Capítulo 2:	Análise Espectral de Processos Estocásticos Univariados .....	12-34
2.1.	O Espectro	
2.1.1.	O Espectro e as Suas Propriedades	
2.1.2.	A Representação Espectral da Função de Autocovariância – A Função de Distribuição Espectral	
2.1.3.	A Representação Espectral dos Processos Estacionários	
2.2.	O Espectro de alguns Processos Comuns	
2.2.1.	Representações Autorregressiva e de Médias Móveis de Séries Temporais	
2.2.2.	O Espectro e a Função Geradora da Autocovariância	
2.2.3.	O Espectro dos Modelos ARMA	
2.2.4.	O Espectro da Soma de Dois Processos Independentes	
2.2.5.	O Espectro de Filtros Lineares	
2.3.	Estimação do Espectro	
2.3.1.	Análise do Periodograma	
2.3.1.1.	O Periodograma e as Suas Propriedades Amostrais	
2.3.1.2.	Testes para a Detecção de Componentes Periódicas	
2.3.2.	O Espectro Amostral	
2.3.3.	O Espectro Alisado	
2.3.3.1.	Alisamento no Domínio Frequência – A Janela Espectral	
2.3.3.2.	Alisamento no Domínio Tempo – A Janela Retardada	
2.3.3.3.	Principais Janelas Utilizadas	
2.3.3.4.	Escolha da Janela	

Capítulo 3:	Análise Espectral de Processos Estocásticos Multivariados.....	35-47
3.1.	Conceitos Gerais	
3.2.	O Caso Bivariado	
3.2.1.	O Espectro Cruzado	
3.2.2.	Interpretação das Funções Espectrais Cruzadas	
3.2.3.	Filtros Lineares Invariantes no Tempo	
3.3.	O Caso Multivariado	
3.4.	Estimação do Espectro Cruzado	
Capítulo 4:	Análise em Componentes Principais Dinâmicas .....	48-55
4.1.	Introdução	
4.2.	Componentes Principais Estáticas	
4.3.	Componentes Principais Dinâmicas	
4.4.	Relação entre a Análise em Componentes Principais Dinâmicas e em Componentes Principais Estáticas	
Capítulo 5:	Análise Factorial Dinâmica Generalizada .....	56-76
5.1.	Introdução	
5.2.	Modelos Factoriais para a Análise de Painéis de Dados de Séries Temporais	
5.2.1.	Notação e Definições	
5.2.2.	Breve Descrição de Alguns Modelos Factoriais	
5.2.2.1.	O Modelo Factorial Estático	
5.2.2.2.	O Modelo Factorial Dinâmico	
5.2.3.	O Modelo Factorial Dinâmico Generalizado	
5.2.3.1.	Definição do Modelo e Fundamentos Teóricos	
5.2.3.2.	Identificação e Estimação Bilateral (2000)	
5.2.3.3.	Identificação e Estimação Unilateral (2001)	
5.2.3.3.1.	Hipóteses Adicionais	
5.2.3.3.2.	Estimação e Previsão	
5.2.3.3.2.1.	Problema de Valores Próprios Generalizados	
5.2.3.3.2.2.	1ª Etapa: Estimação Consistente do Espaço de Factores Comuns	
5.2.3.3.2.3.	2ª Etapa: Construção do Estimador Ótimo	
5.2.3.3.3.	Variáveis Não Estacionárias	
Referências .....		75-76

## Capítulo 0: Introdução

A Análise em Factores Comuns baseia-se num método de estatística multivariada que procura simplificar um conjunto de dados através da redução do número de variáveis necessárias para os descrever. O seu objectivo principal é descrever, se possível, a estrutura da matriz de variâncias-covariâncias de um conjunto elevado de variáveis originais em função de outras variáveis, não observáveis, e em menor número, os chamados factores comuns. Estas novas variáveis expressam o que existe de comum nas variáveis originais.

O estudo de modelos dinâmicos e a sua aplicação a problemas económicos é uma das áreas, dentro da Análise em Factores, que recentemente mais interesse tem suscitado junto dos econométricos. A hipótese subjacente a este tipo de modelos é a de considerar que o movimento conjunto de um número elevado de painéis de séries económicas pode ser explicado por um pequeno número de factores não observáveis.

O objectivo deste trabalho é o de apresentar as ferramentas teóricas necessárias para compreender o funcionamento dos modelos em Factores Dinâmicos. Dar-se-á ênfase às metodologias que utilizam conceitos ligados à Análise Espectral. A ideia fundamental subjacente à análise no domínio da frequência é de que qualquer processo estacionário pode ser visto como a soma de um número infinito de componentes periódicas não correlacionadas. A análise no domínio da frequência permite-nos estudar a relação entre as componentes periódicas das variáveis associadas a uma determinada frequência  $\omega$ .

Numa primeira parte (3 primeiros capítulos) serão delineados os principais conceitos ligados à Análise Espectral de Processos Estocásticos Univariados e Multivariados. De seguida será apresentado um capítulo dedicado à Análise em Componentes Principais Dinâmicas. Finalmente, será exposto o Modelo em Factores ou Modelo Factorial Dinâmico desenvolvido recentemente por Mário Forni, Marco Lippi e Lucrezia Reichlin. A metodologia apresentada neste último capítulo é actualmente utilizada pelo CEPR no cálculo de um indicador coincidente para o ciclo económico da área euro, o  $\text{€COIN}$ .

Este trabalho foi financiado pelo Projecto da FCT: POCTI/41502/ECO/2001, participado pelo Feder.

# Capítulo 1: Análise de Fourier

Usualmente, o estudo de séries temporais é realizado no domínio tempo, recorrendo à análise das funções autocovariância e autocorrelação. Estas funções resultam do confronto, em termos médios ou esperados, entre valores da sucessão com diferentes atrasos. A análise no domínio frequência, que pode ser relacionada com a análise no tempo, procura investigar, em que medida, diferentes frequências ou componentes periódicas, com comportamento sinusoidal, contribuem para a variabilidade das sucessões cronológicas.

A análise espectral resulta de uma modificação da teoria de integrais e séries de Fourier de forma a que se possa transitar do caso das funções determinísticas para o caso dos processos estocásticos. Começa-se, por isso, por referir alguns conceitos da Análise de Fourier.

## 1.1. Introdução

O objectivo desta secção é o de mostrar como é que determinadas funções podem ser escritas recorrendo a funções com comportamento sinusoidal. De facto, por vezes é vantajoso representar uma função como sendo a combinação linear de outras funções elementares que constituem uma base. Ir-se-á utilizar como tais funções elementares as funções seno e cosseno e as funções exponenciais complexas.

Inicialmente, serão consideradas funções ou séries temporais discretas.

## 1.2. Funções Ortogonais

Sejam  $\phi_k(t)$  e  $\phi_j(t)$  duas funções complexas com domínio  $D \subset \mathfrak{R}$ .

Estas funções  $\phi_k(t)$  e  $\phi_j(t)$ , quando discretas e definidas num conjunto discreto, dizem-se ortogonais se:

$$(1) \quad \sum_{t \in D} \phi_k(t) \phi_j^*(t) \begin{cases} = 0, & k \neq j \\ \neq 0, & k = j \end{cases}$$

onde  $\phi_j^*(t)$  representa o conjugado de  $\phi_j(t)$ .

As funções  $\phi_k(t)$  e  $\phi_j(t)$ , se contínuas e definidas num intervalo real, dizem-se ortogonais se:

$$(2) \quad \int_D \phi_k(t) \phi_j^*(t) dt \begin{cases} = 0, & k \neq j \\ \neq 0, & k = j \end{cases}$$

Neste parágrafo, são considerados dois sistemas equivalentes de funções ortogonais, úteis no estudo das séries temporais.

Admita-se que as funções trigonométricas  $\text{sen}(2\pi kt/n)$  e  $\text{cos}(2\pi kt/n)$  estão definidas em  $n$  pontos,  $n$  finito. Para  $k = 0, 1, 2, \dots, [n/2]$ , onde  $[x]$  representa o maior inteiro menor ou igual a  $x$ , o sistema

$$(3) \quad \{\text{sen}(2\pi kt/n), \text{cos}(2\pi kt/n) : k = 0, 1, \dots, [n/2]\}$$

contém  $n$  funções não nulas. É possível demonstrar que o referido sistema é um conjunto de funções ortogonais, pois verificam-se os seguintes resultados:

$$(4) \quad \sum_{t=1}^n \text{cos}(2\pi kt/n) \text{cos}(2\pi jt/n) = \begin{cases} n, & k = j = 0 \text{ ou } n/2 \text{ (se } n \text{ par)} \\ n/2, & k = j \neq 0 \text{ e } n/2 \text{ (se } n \text{ par)} \\ 0, & k \neq j \end{cases}$$

$$(5) \quad \sum_{t=1}^n \text{sen}(2\pi kt/n) \text{sen}(2\pi jt/n) = \begin{cases} 0, & k = j = 0 \text{ ou } n/2 \text{ (se } n \text{ par)} \\ n/2, & k = j \neq 0 \text{ e } n/2 \text{ (se } n \text{ par)} \\ 0, & k \neq j \end{cases}$$

$$(6) \quad \sum_{t=1}^n \text{sen}(2\pi kt/n) \text{cos}(2\pi jt/n) = 0, \quad \forall k, j.$$

Recordando a relação de Euler,  $e^{i\omega} = \text{cos } \omega + i \text{sen } \omega$ , e as identidades  $\text{sen } \omega = (e^{i\omega} - e^{-i\omega})/2i$  e  $\text{cos } \omega = (e^{i\omega} + e^{-i\omega})/2$ , é evidente que o sistema de funções trigonométricas pode ser representado na forma de exponenciais complexas. Nesta representação, o sistema toma a seguinte forma:

$$(7) \quad \left\{ e^{i2\pi kt/n} : -\frac{n}{2} + 1 \leq k \leq \frac{n}{2} \text{ se } n \text{ for par e } -\frac{(n-1)}{2} \leq k \leq \frac{(n-1)}{2} \text{ se } n \text{ for ímpar} \right\}$$

Neste caso, também existem  $n$  funções, que constituem um sistema de funções ortogonais, verificando o resultado

$$\sum_{t=1}^n e^{i2\pi kt/n} e^{-i2\pi jt/n} = \begin{cases} n, & k = j \\ 0, & k \neq j \end{cases}$$

### 1.3. Representação de Fourier de Funções Temporais Discretas Finitas

Seja  $Z_1, Z_2, \dots, Z_n$  uma sequência de  $n$  números. Esta sequência pode ser considerada um conjunto de coordenadas de um ponto num espaço de dimensão  $n$ . Qualquer vector

desse espaço pode ser escrito como uma combinação linear de  $n$  vectores que constituam uma base. Sabe-se que  $n$  vectores ortogonais constituem uma base de um espaço de dimensão  $n$ . Assim sendo, a sequência de pontos  $\{Z_t\}$  pode ser escrita como combinação linear das  $n$  funções trigonométricas ortogonais referidas na secção anterior:

$$(9) \quad Z_t = \sum_{k=0}^{\lfloor n/2 \rfloor} [a_k \cos(2\pi kt/n) + b_k \sin(2\pi kt/n)], \quad t = 1, 2, \dots, n.$$

A equação anterior é chamada série de Fourier da sequência  $Z_t$ . Os coeficientes  $a_k$  e  $b_k$  denominam-se coeficientes de Fourier e são dados pelas seguintes expressões:

$$(10) \quad a_k = \begin{cases} \frac{1}{n} \sum_{t=1}^n Z_t \cos(2\pi kt/n), & k = 0, \text{ e } k = n/2 \text{ se } n \text{ é par} \\ \frac{2}{n} \sum_{t=1}^n Z_t \cos(2\pi kt/n), & k = 1, 2, \dots, \left[ \frac{n-1}{2} \right] \end{cases},$$

$$(11) \quad b_k = \frac{2}{n} \sum_{t=1}^n Z_t \sin(2\pi kt/n), \quad k = 1, 2, \dots, \left[ \frac{n-1}{2} \right].$$

As expressões para  $a_k$  e  $b_k$  são obtidas multiplicando a série de Fourier, de ambos os lados, por  $\cos(2\pi kt/n)$  e  $\sin(2\pi kt/n)$ , respectivamente, e tendo em conta os resultados, apresentados na secção anterior, que estabelecem a ortogonalidade das funções trigonométricas da base.

As frequências  $\omega_k = 2\pi k/n$ ,  $k = 1, 2, \dots, \lfloor n/2 \rfloor$ , têm o nome de frequências de Fourier.

Usando agora, como base do espaço  $n$ -dimensional, o sistema de exponenciais complexas apresentado na secção anterior, pode-se escrever a série de Fourier de  $\{Z_t\}$  da seguinte forma:

$$(12) \quad Z_t = \begin{cases} \sum_{k=-(n-1)/2}^{(n-1)/2} c_k e^{i\omega_k t}, & \text{se } n \text{ é ímpar} \\ \sum_{k=-n/2+1}^{n/2} c_k e^{i\omega_k t}, & \text{se } n \text{ é par} \end{cases},$$

sendo os coeficientes de Fourier  $c_k$  dados por:

$$(13) \quad c_k = \frac{1}{n} \sum_{t=1}^n Z_t e^{-i\omega_k t}.$$

Pelo exposto, conclui-se que qualquer sequência finita pode ser escrita como a combinação linear das funções que constituem os dois sistemas de funções ortogonais apresentados na secção anterior.

#### 1.4.Representação de Fourier de Funções Temporais Discretas Periódicas

Seja  $f: \mathfrak{R} \rightarrow \mathfrak{R}$  uma função tal que

$$(14) \quad f(x) = f(x + kP), \quad \forall x \in \mathfrak{R}, \quad k \in \mathbb{Z}.$$

Nestas condições,  $f$  diz-se periódica. Seja  $P$  o menor valor a verificar a relação (14). Então, diz-se que a função periódica  $f$  tem período fundamental ou, simplesmente, período  $P$ .

Admita-se que a série temporal discreta  $Z_t$  é periódica com período  $n$ , ou seja,  $Z_{t+n} = Z_t$  para todo o  $t$  inteiro. Então,  $Z_t$  fica totalmente definida se se conhecer o seu comportamento num período de tempo com duração igual ao seu período, ou seja, se considerarmos  $t=1,2,\dots,n$ .

Usando a representação de Fourier de uma sequência finita de pontos apresentada na secção anterior, pode-se escrever  $Z_t$ ,  $t=1,2,\dots,n$ , como a combinação linear do sistema de funções trigonométricas ortogonais:

$$(15) \quad Z_t = \sum_{k=0}^{\lfloor n/2 \rfloor} [a_k \cos \omega_k t + b_k \text{sen} \omega_k t], \quad t = 1, 2, \dots, n,$$

$$(16) \quad a_k = \begin{cases} \frac{1}{n} \sum_{t=1}^n Z_t \cos \omega_k t, & k = 0, \text{ e } k = n/2 \text{ se } n \text{ é par} \\ \frac{2}{n} \sum_{t=1}^n Z_t \cos \omega_k t, & k = 1, 2, \dots, \left[ \frac{n-1}{2} \right] \end{cases},$$

$$(17) \quad b_k = \frac{2}{n} \sum_{t=1}^n Z_t \text{sen} \omega_k t, \quad k = 1, 2, \dots, \left[ \frac{n-1}{2} \right],$$

ou como combinação linear das funções exponenciais complexas ortogonais:

$$(18) \quad Z_t = \begin{cases} \sum_{k=-(n-1)/2}^{(n-1)/2} c_k e^{i\omega_k t}, & \text{se } n \text{ é ímpar} \\ \sum_{k=-n/2+1}^{n/2} c_k e^{i\omega_k t}, & \text{se } n \text{ é par} \end{cases}$$

$$(19) \quad c_k = \frac{1}{n} \sum_{t=1}^n Z_t e^{-i\omega_k t}.$$

As equações anteriores são válidas para qualquer  $t$ .

O menor valor  $n$ , para o qual é válida a representação de Fourier de  $Z_t$  (14) ou (16), é chamado período fundamental. A correspondente frequência  $2\pi/n$  é a frequência fundamental. Os termos das representações (14) e (16) correspondentes a  $k = 1$  e  $k = -1$  têm ambos período fundamental  $n$  e são os primeiros componentes harmônicos. Os termos correspondentes a  $k = j$  e  $k = -j$  têm frequência  $j(2\pi/n)$  e são chamados os  $j$ -ésimos componentes harmônicos. Assim sendo, todos os componentes harmônicos têm frequências múltiplas da frequência fundamental.

Seja  $\sum_{t=1}^n Z_t^2$  a energia associada a uma série periódica  $Z_t$  de período  $n$ .

Usando a representação de Fourier, a partir das funções trigonométricas, é possível mostrar que:

$$(20) \quad \sum_{t=1}^n Z_t^2 = \begin{cases} na_0^2 + \frac{n}{2} \sum_{k=1}^{[(n-1)/2]} (a_k^2 + b_k^2), & \text{se } n \text{ é par} \\ na_0^2 + \frac{n}{2} \sum_{k=1}^{[(n-1)/2]} (a_k^2 + b_k^2) + na_{n/2}^2, & \text{se } n \text{ é ímpar} \end{cases}$$

ou, usando a representação de Fourier a partir das funções exponenciais complexas:

$$(21) \quad \sum_{t=1}^n Z_t^2 = \begin{cases} n \sum_{k=-(n-1)/2}^{(n-1)/2} |c_k|^2, & \text{se } n \text{ é par} \\ n \sum_{k=-n/2+1}^{n/2} |c_k|^2, & \text{se } n \text{ é ímpar} \end{cases}$$

onde  $|c_k|^2 = c_k c_k^*$ .

Se se considerar a energia por unidade de tempo de uma série periódica, obtém-se a sua potência, sendo esta definida por:

$$(22) \quad \text{Potência} = \begin{cases} a_0^2 + \frac{1}{2} \sum_{k=1}^{[(n-1)/2]} (a_k^2 + b_k^2), & \text{se } n \text{ é par} \\ a_0^2 + \frac{1}{2} \sum_{k=1}^{[(n-1)/2]} (a_k^2 + b_k^2) + a_{n/2}^2, & \text{se } n \text{ é ímpar} \end{cases}$$

ou, de forma equivalente,

$$(23) \quad \text{Potência} = \begin{cases} \sum_{k=-(n-1)/2}^{(n-1)/2} |c_k|^2, & \text{se } n \text{ é par} \\ \sum_{k=-n/2+1}^{n/2} |c_k|^2, & \text{se } n \text{ é ímpar} \end{cases}$$

Como já foi referido, os  $j$ -ésimos componentes harmónicos incluem os termos para  $k = j$  e  $k = -j$ , uma vez que correspondem à mesma frequência  $j(2\pi/n)$ . Assim, as quantidades

$$(24) \quad \begin{cases} f_0 = c_0^2 = a_0^2; & f_{n/2} = |c_{n/2}|^2, & \text{quando } n \text{ é par} \\ f_k = |c_{-k}|^2 + |c_k|^2 = 2|c_k|^2 = \frac{1}{2}(a_k^2 + b_k^2) \end{cases}$$

representam a contribuição para a potência total. A quantidade  $f_k$ , representada como função de  $\omega_k$ , chama-se espectro potência e descreve de que forma a potência total se distribui pelas diversas frequências dos componentes de  $Z_t$ .

### 1.5. Representação de Fourier de Funções Temporais Discretas Não Periódicas – A Transformada de Fourier em Tempo Discreto

Considere-se a sequência não periódica  $Z_t$ , de duração finita, tal que  $Z_t = 0$  para  $|t| > M$ . Tomando  $n = (2M + 1)$ , é possível definir a seguinte função,

$$(25) \quad Y_{t+jn} = Z_t, \quad -(n-1)/2 \leq t \leq (n-1)/2, \quad j = 0, \pm 1, \pm 2, \dots$$

que, por construção, é periódica, de período  $n$ .

Usando a representação em série de Fourier, a partir de exponenciais complexas, e sabendo que  $n$  é ímpar temos que:

$$(26) \quad Y_t = \sum_{k=-(n-1)/2}^{(n-1)/2} c_k e^{i2\pi kt/n},$$

$$(27) \quad c_k = \frac{1}{n} \sum_{t=-(n-1)/2}^{(n-1)/2} Y_t e^{-i2\pi kt/n}.$$

Como

$$(28) \quad Z_t = \begin{cases} Y_t, & -(n-1)/2 \leq t \leq (n-1)/2 \\ 0, & \text{c.c.} \end{cases}$$

pode-se escrever,

$$(29) \quad c_k = \frac{1}{n} \sum_{t=-\infty}^{+\infty} Z_t e^{-i2\pi kt/n}.$$

Defina-se agora a função,

$$(30) \quad f(\omega) = \frac{1}{2\pi} \sum_{t=-\infty}^{+\infty} Z_t e^{-i\omega t}.$$

É possível verificar que

$$(31) \quad c_k = \frac{2\pi}{n} f(k\Delta\omega)$$

onde  $\Delta\omega = 2\pi/n$  representa a amplitude entre frequências. Com base nos resultados anteriores, obtém-se a expressão seguinte:

$$(32) \quad Y_t = \sum_{k=-(n-1)/2}^{(n-1)/2} f(k\Delta\omega) e^{ik\Delta\omega t} \Delta\omega.$$

Como cada termo do somatório representa a área de um rectângulo de altura  $f(k\Delta\omega)e^{ik\Delta\omega t}$  e largura  $\Delta\omega = 2\pi/n$ , tendo o somatório  $n$  parcelas, e como, quando  $n \rightarrow \infty$ , tem-se  $Y_t \rightarrow Z_t$  e  $\Delta\omega \rightarrow 0$ , o somatório, em limite, transforma-se num integral com intervalo de integração de amplitude  $2\pi$ , ou seja,

$$(33) \quad Z_t = \lim_{n \rightarrow +\infty} Y_t = \lim_{\Delta\omega \rightarrow 0} \sum_{k=-\infty}^{+\infty} f(k\Delta\omega) e^{ik\Delta\omega t} \Delta\omega = \int_{2\pi} f(\omega) e^{i\omega t} d\omega.$$

Sabendo que a função  $f(\omega)e^{i\omega t}$  é periódica com período  $2\pi$ , pode-se considerar como intervalo de integração qualquer intervalo de amplitude igual ao período, em particular  $[-\pi, \pi]$ . Nestas condições, temos que:

$$(34) \quad Z_t = \int_{-\pi}^{\pi} f(\omega) e^{i\omega t} d\omega, \quad t = 0, \pm 1, \pm 2, \dots$$

e

$$(35) \quad f(\omega) = \frac{1}{2\pi} \sum_{l=-\infty}^{+\infty} Z_l e^{-i\omega l}, \quad -\pi \leq \omega \leq \pi.$$

Nas relações anteriores,  $f(\omega)$  é a transformada de Fourier de  $Z_l$  (em tempo discreto) e  $Z_l$  é a transformada de Fourier inversa de  $f(\omega)$  (em tempo discreto). A transformada de Fourier inversa, tal como definida, decompõe a sequência  $Z_l$  numa soma de sinusóides complexas infinitamente próximas em frequência e com amplitudes  $|f(\omega)|d\omega$ . Assim sendo, a quantidade  $|f(\omega)|$ , com o nome de espectro, fornece informação sobre como as sinusóides em diferentes frequências compõem  $Z_l$ .

A energia associada ao processo  $Z_l$  é dada por  $\sum_{l=-\infty}^{+\infty} |Z_l|^2$ .

A relação de Parseval estabelece a relação entre a energia no domínio tempo e no domínio frequência:

$$(36) \quad \sum_{l=-\infty}^{+\infty} |Z_l|^2 = 2\pi \int_{-\pi}^{\pi} |f(\omega)|^2 d\omega.$$

Assim sendo, a energia total pode ser obtida calculando  $|Z_l|^2$ , ou seja, a energia por unidade de tempo, e somando ao longo de todo o tempo, ou calculando  $2\pi|f(\omega)|^2$ , ou seja, a energia por unidade de frequência, e integrando ao longo de todas as frequências. A função de  $\omega$ ,  $g(\omega) = 2\pi|f(\omega)|^2$  é chamada função densidade espectral de energia.

Na dedução dos resultados anteriores foi considerado que o processo  $Z_l$  tinha uma duração finita. No entanto, a sua representação de Fourier continua válida no caso deste ter uma duração infinita. Nesse caso, podemos-nos questionar sobre a convergência do somatório que define  $f(\omega)$ , a transformada de Fourier de  $Z_l$ . Uma condição sobre  $Z_l$ , que garante a convergência do somatório, é que esta sequência seja absolutamente somável, ou seja:

$$(37) \quad \sum_{l=-\infty}^{+\infty} |Z_l| < \infty.$$

## 1.6. Representação de Fourier de Funções em Tempo Contínuo

Nos parágrafos anteriores, foram apresentadas as representações de Fourier para o caso de processos discretos exibindo ou não um comportamento periódico. Nos parágrafos seguintes considerar-se-á o caso das funções definidas num tempo contínuo.

### 1.6.1. Representação de Fourier de Funções Periódicas

Considere-se a função  $f(t)$ , definida para todo o  $t$  real. Considere-se que  $f(t)$  é periódica de período  $P$ , sendo por isso completamente caracterizada pelo seu comportamento num intervalo de amplitude  $P$ . É suficiente, portanto, estudar a função no intervalo  $[-P/2, P/2]$ .

Como foi referido anteriormente, verifica-se a seguinte relação entre o período fundamental  $P$  e a frequência fundamental  $\omega_0$ :

$$(38) \quad \omega_0 = \frac{2\pi}{P}.$$

Associado à frequência fundamental, pode definir-se o seguinte sistema de exponenciais complexas, harmonicamente relacionadas, que se prova ser um sistema periódico, de período  $P$ , e ser ortogonal em qualquer intervalo de amplitude  $P$ :

$$(39) \quad \{e^{ik\omega_0 t} : k = 0, \pm 1, \pm 2, \dots\}.$$

Com base no sistema ortogonal acima definido, pode-se representar a função  $f(t)$  como uma série de Fourier, da seguinte:

$$(40) \quad f(t) = \sum_{k=-\infty}^{+\infty} c_k e^{ik\omega_0 t}.$$

Se a série anterior convergir uniformemente para  $f(t)$ , o coeficiente de Fourier é dado por:

$$(41) \quad c_k = \frac{1}{P} \int_{-P/2}^{P/2} f(t) e^{-ik\omega_0 t} dt.$$

Para que a representação de Fourier de  $f(t)$  seja válida, ou seja, para que a série de Fourier convirja uniformemente para  $f(t)$  é suficiente que verifique as condições de Dirichelet para funções contínuas. Essas condições dizem que, se  $f(t)$  for uma função periódica limitada com período  $P$ , tiver um número finito de máximos e de mínimos num dado período e um número finito de descontinuidades, então a série de Fourier converge para  $f(t)$  em todos os seus pontos de continuidade e converge para a média dos limites à esquerda e à direita de  $f(t)$  em cada ponto em que esta seja descontínua.

É possível mostrar que a relação de Parseval para esta representação de Fourier é dada por:

$$(42) \quad \int_{-P/2}^{P/2} |f(t)|^2 dt = P \sum_{k=-\infty}^{+\infty} |c_k|^2.$$

## 1.6.2. Representação de Fourier de Funções Não Periódicas – A Transformada de Fourier em Tempo Contínuo

Considere-se a seguinte função  $f(t)$ , não periódica, definida para todo o  $t$  real e com duração finita, ou seja,  $f(t) = 0, \forall |t| > P/2$ . Defina-se a função  $g(t)$ , periódica de período  $P$ , da seguinte forma:

$$(43) \quad \begin{cases} g(t) = f(t), & -P/2 \leq t \leq P/2 \\ g(t + jP) = g(t), & j = \pm 1, \pm 2, \dots \end{cases}$$

É possível escrever:

$$(44) \quad g(t) = \sum_{k=-\infty}^{+\infty} c_k e^{ik\omega_0 t}$$

e

$$(45) \quad c_k = \frac{1}{P} \int_{-P/2}^{P/2} g(t) e^{-ik\omega_0 t} dt$$

onde  $\omega_0 = \frac{2\pi}{P}$ . De acordo com a definição de  $g(t)$  no intervalo  $[-P/2, P/2]$ , pode reescrever-se a equação (45) da seguinte forma:

$$(46) \quad c_k = \frac{\omega_0}{2\pi} \int_{-P/2}^{P/2} f(t) e^{-ik\omega_0 t} dt = \frac{\omega_0}{2\pi} \int_{-P/2}^{P/2} f(u) e^{-ik\omega_0 u} du$$

de onde, após a substituição de  $c_k$  na expressão (44), se tem que:

$$(47) \quad g(t) = \sum_{k=-\infty}^{+\infty} c_k e^{ik\omega_0 t} = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} H(k\omega_0) \omega_0,$$

onde

$$(48) \quad H(k\omega_0) = \int_{-P/2}^{P/2} f(u) e^{ik\omega_0(t-u)} du.$$

Cada termo do somatório que define  $g(t)$  representa a área de um rectângulo de altura  $H(k\omega_0)$  e largura  $\omega_0$ . Quando  $P$  tende para infinito,  $g(t) \rightarrow f(t)$ ,  $\omega_0 \rightarrow 0$  e o somatório torna-se um intergral. Assim sendo,

$$(49) \quad f(t) = \lim_{P \rightarrow +\infty} g(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} H(\omega) d\omega$$

e

$$(50) \quad H(\omega) = \lim_{P \rightarrow +\infty} H(k\omega_0) = \int_{-\infty}^{+\infty} f(u) e^{i\omega(t-u)} du.$$

Substituindo a expressão de  $H(\omega)$  na expressão (49), temos que:

$$(51) \quad f(t) = \int_{-\infty}^{+\infty} F(\omega) e^{i\omega t} d\omega$$

e

$$(52) \quad F(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(t) e^{-i\omega t} dt.$$

À função  $F(\omega)$  chama-se transformada de Fourier ou integral de Fourier de  $f(t)$  e a  $f(t)$  chama-se inversa da transformada de Fourier ou inversa do integral de Fourier de  $F(\omega)$ .

A relação de Parseval para esta representação de Fourier é dada por:

$$(53) \quad \int_{-\infty}^{+\infty} |F(\omega)|^2 d\omega = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |f(t)|^2 dt.$$

Nas deduções anteriores, considerou-se que  $f(t)$  tinha duração finita. No entanto, os resultados continuam válidos para algumas funções de duração infinita. As seguintes condições de Dirichlet para funções não contínuas constituem um conjunto de condições suficientes para a existência de  $F(\omega)$  para o caso de funções de duração infinita:

- a)  $f(t)$  é absolutamente integrável, ou seja,  $\int_{-\infty}^{+\infty} |f(t)| dt < +\infty$ ;
- b)  $f(t)$  tem um número finito de máximos e mínimos e um número finito de descontinuidades em qualquer intervalo finito.

## Capítulo 2: Análise Espectral de Processos Estacionários Univariados

Os conceitos da análise de Fourier expostos no capítulo 1 permitem o estudo das séries temporais no domínio frequência. Na realidade, o espectro de um processo estacionário é a transformada de Fourier da sua função de autocovariância absolutamente somável. No caso da função de autocovariância não ser absolutamente somável, o processo estocástico continua a ter uma representação espectral a partir da função de distribuição espectral. No entanto, no capítulo anterior a análise foi realizada sobre funções fixas, não estocásticas, enquanto que, no presente capítulo, os resultados serão generalizados ao caso dos processos estocásticos. É, portanto, necessário distinguir entre a representação espectral de uma concretização particular de um processo estocástico e a representação espectral de todas as concretizações possíveis de um processo estocástico.

### 2.1. O Espectro

#### 2.1.1. O Espectro e as suas propriedades

Seja  $Z_t$  um processo estacionário de valores reais com uma função de autocovariância absolutamente somável  $\gamma_k$ . Pelos resultados da secção 1.1.5., a transformada de Fourier de  $\gamma_k$  existe e é dada por:

$$(1) \quad f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \gamma_k e^{-iak}$$

Considerando que as funções  $\gamma_k$  e coseno são pares e a função seno é ímpar, temos que:

$$(2) \quad f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \gamma_k \cos \omega k = \frac{1}{2\pi} \gamma_0 + \frac{1}{\pi} \sum_{k=1}^{+\infty} \gamma_k \cos \omega k, \quad -\pi \leq \omega \leq \pi.$$

A função  $\gamma_k$  é recuperável pela inversa da transformada de Fourier,

$$(3) \quad \gamma_k = \int_{-\infty}^{+\infty} f(\omega) e^{iak} d\omega.$$

A transformada de Fourier da função de autocovariância goza das seguintes propriedades:

1.  $f(\omega)$  é uma função contínua, não negativa, de valores reais. É muitas vezes referida como espectro da função de autocovariância ou do processo estacionário.
2.  $f(\omega)$  é uma função periódica de período  $2\pi$ . Além disso, é uma função par, pelo que é usual que a sua representação gráfica se cinja ao intervalo  $[0, \pi]$ .

3. Verifica-se a igualdade  $\text{var}(Z_t) = \gamma_0 = \int_{-\infty}^{+\infty} f(\omega) d\omega$ . Assim sendo,  $f(\omega)$  pode ser interpretada como a decomposição da variância do processo. O termo  $f(\omega) d\omega$  é a contribuição para a variância da componente do processo com frequências no intervalo  $(\omega, \omega + d\omega)$ . Um pico no espectro indica uma importante contribuição para a variância das componentes com frequência no intervalo correspondente.
4. O espectro  $f(\omega)$  e a autocovariância  $\gamma_k$  constituem um par de transformadas de Fourier sendo uma das funções univocamente determinada pela outra. Deste modo, a abordagem no domínio tempo é teoricamente equivalente à abordagem no domínio frequência.

### 2.1.2. A Representação Espectral da Função de Autocovariância – A Função de Distribuição Espectral

Na secção anterior, a representação espectral de  $\gamma_k$  apresentada é válida para o caso em que a função de autocovariância é absolutamente somável. De um modo mais geral, podemos sempre obter a representação espectral de uma função de autocovariância, absolutamente somável ou não, em termos do seguinte integral de Fourier-Stieltjes:

$$(4) \quad \gamma_k = \int_{-\pi}^{\pi} e^{i\omega k} dF(\omega).$$

A este integral chama-se representação espectral da função de autocovariância  $\gamma_k$  e a  $F(\omega)$  chama-se função distribuição espectral.

Como qualquer função de distribuição estatística,  $F(\omega)$  é uma função não decrescente que se pode decompor em três componentes:

1. uma função do tipo degrau que consiste num número contável de saltos finitos;
2. uma função absolutamente contínua;
3. uma função singular, sendo esta última componente insignificante e ignorada na maioria das aplicações.

Assim sendo, pode-se escrever a função de distribuição espectral como

$$(5) \quad F(\omega) = F_s(\omega) + F_c(\omega)$$

onde  $F_s(\omega)$  é uma função do tipo degrau e  $F_c(\omega)$  é uma função absolutamente contínua. Num processo com função de autocovariância absolutamente somável, verifica-se  $F(\omega) = F_c(\omega)$  e  $dF(\omega) = f(\omega) d\omega$ .

Deve notar-se que, embora  $F(\omega)$  seja uma função não decrescente e não negativa, não é uma verdadeira função de distribuição, uma vez que,

$$(6) \quad \int_{-\pi}^{\pi} dF(\omega) = \gamma_0,$$

não sendo este integral, em geral, idêntico a 1. No entanto, é possível definir uma nova função que satisfaz todas as propriedades de uma função de distribuição:

$$(7) \quad G(\omega) = \frac{F(\omega)}{\gamma_0}.$$

Quando  $dF(\omega) = f(\omega)d\omega$ , define-se  $p(\omega)$ , função com as propriedades de uma função densidade de probabilidade no intervalo  $[-\pi, \pi]$  e, por isso, referida como função de densidade espectral (embora o mesmo termo seja usado para referir  $f(\omega)$  em alguma literatura):

$$(8) \quad p(\omega)d\omega = dG(\omega) = \frac{f(\omega)}{\gamma_0}d\omega.$$

É assim possível obter o seguinte par de transformadas de Fourier:

$$(9) \quad p(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \rho_k e^{-i\omega k} = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{k=1}^{+\infty} \rho_k \cos \omega k, \quad -\pi \leq \omega \leq \pi$$

e

$$(10) \quad \rho_k = \int_{-\pi}^{\pi} p(\omega) e^{i\omega k} d\omega, \quad k = 0, \pm 1, \pm 2, \dots$$

### 2.1.3. A Representação Espectral dos Processos Estacionários

Tal como uma função de autocovariância pode ser sempre representada através do integral de Fourier-Stieltjes que envolve a função de distribuição espectral, uma dada concretização de uma série temporal  $Z_t$ , pode ser escrita como,

$$(11) \quad Z_t = \int_{-\pi}^{\pi} e^{i\omega t} dU(\omega)$$

onde  $U(\omega)$  desempenha um papel semelhante ao de  $F(\omega)$  para essa dada concretização da série temporal. No entanto, não é de esperar que, para diferentes concretizações, a função  $U(\omega)$  se mantenha inalterada. De facto, a cada concretização de  $Z_t$ , corresponderá uma concretização de  $U(\omega)$ . Usando o integral acima definido para representar o processo  $Z_t$  em todas as suas possíveis concretizações,  $U(\omega)$  passa a corresponder a um processo estocástico de valores complexos para cada  $\omega$ . Nessas condições, o integral passa a ser estocástico, sendo agora a igualdade dada em termos de média quadrática, ou seja,

$$(12) \quad E \left[ \left| Z_t - \int_{-\pi}^{\pi} e^{i\omega t} dU(\omega) \right|^2 \right] = 0.$$

A relação,

$$(13) \quad Z_t = \int_{-\pi}^{\pi} e^{i\omega t} dU(\omega)$$

é chamada representação espectral do processo estacionário  $Z_t$ .

Para estudar  $U(\omega)$ , considere-se que  $Z_t$  é um processo estacionário de valores complexos e de média nula. Assim,

$$(14) \quad E[dU(\omega)] = 0, \forall \omega,$$

definindo-se a função de autocovariância como

$$(15) \quad \gamma_k = E[Z_t Z_{t+k}^*]$$

onde  $Z_{t+k}^*$  é o complexo conjugado de  $Z_{t+k}$ .

Usando o facto de que  $U(\omega)$  é um processo ortogonal, consequência da estacionaridade de  $Z_t$ , ou seja,

$$(16) \quad E[dU(\omega)dU^*(\lambda)] = 0, \forall \omega \neq \lambda,$$

obtém-se:

$$(17) \quad \gamma_k = \int_{-\pi}^{\pi} e^{-i\omega k} E\{dU(\omega)^2\} = \int_{-\pi}^{\pi} e^{i\omega k} E\{dU(\omega)^2\} = \int_{-\pi}^{\pi} e^{i\omega k} dF(\omega),$$

onde

$$(18) \quad dF(\omega) = E\{dU(\omega)^2\} = E[dU(\omega)dU^*(\omega)], \quad \text{para } -\pi \leq \omega \leq \pi.$$

A  $F(\omega)$  chama-se função de distribuição espectral de  $Z_t$ . Se  $Z_t$  for puramente não determinístico, verificando-se  $F(\omega) = F_c(\omega)$ , vem que:

$$(19) \quad f(\omega)d\omega = dF(\omega) = E\{dU(\omega)^2\} \quad \text{para } -\pi \leq \omega \leq \pi$$

e

$$(20) \quad \gamma_k = \int_{-\pi}^{\pi} e^{i\omega k} f(\omega)d\omega.$$

De acordo com esta representação, pode-se interpretar  $f(\omega)d\omega$  como sendo a contribuição para a variância de  $Z_t$  atribuível à componente  $Z_t(\omega)$  de  $Z_t$ , com frequências no intervalo  $(\omega, \omega + d\omega)$ .

Mostra-se que  $Z_t$  toma valores reais se e só se, para todo o  $\omega$ , se verificar,

$$(21) \quad dU^*(\omega) = dU(\omega).$$

## 2.2.O Espectro de alguns Processos Comuns

### 2.2.1. Representações Autorregressiva e de Médias Móveis de Séries Temporais

Na análise de séries temporais, são utilizadas duas representações básicas dos processos. Uma tem o nome de médias móveis e a outra tem o nome de autorregressiva. A partir destas pode construir-se o modelo autorregressivo de médias móveis.

A representação em médias móveis do processo  $Z_t$ , consiste em escrevê-lo como uma combinação linear de variáveis aleatórias não correlacionadas:

$$(22) \quad Z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \mu + \sum_{j=0}^{+\infty} \psi_j a_{t-j}$$

onde  $\psi_0 = 1$ ,  $\{a_t\}$  é um processo de ruído branco com média nula e  $\sum_{j=0}^{+\infty} \psi_j^2 < +\infty$ .

Definindo  $\dot{Z}_t = Z_t - \mu$  e utilizando o operador de atraso  $B^j x_t = x_{t-j}$ , é possível escrever um processo de médias móveis (MA) da seguinte forma:

$$(23) \quad \dot{Z}_t = \psi(B)a_t,$$

onde  $\psi(B) = \sum_{j=0}^{+\infty} \psi_j B^j$ .

Mostra-se que:

$$(24) \quad E(Z_t) = \mu,$$

$$(25) \quad \text{Var}(Z_t) = \gamma_0 = \sigma_a^2 \sum_{j=0}^{+\infty} \psi_j^2,$$

$$(26) \quad \gamma_k = \sigma_a^2 \sum_{i=0}^{+\infty} \psi_i \psi_{i+k}$$

e

$$(27) \quad \rho_k = \frac{\sum_{i=0}^{+\infty} \psi_i \psi_{i+k}}{\sum_{i=0}^{+\infty} \psi_i^2}.$$

As funções de autocorrelação,  $\rho_k$ , e de autocovariância,  $\gamma_k$ , são funções apenas do intervalo de tempo  $k$ . Assim sendo, a estacionaridade depende do facto de  $\gamma_k$  ser finita ou não para cada  $k$ . Ora,

$$(28) \quad |\gamma_k| = |E(\dot{Z}_t \dot{Z}_{t+k})| \leq \sqrt{\text{Var}(Z_t) \text{Var}(Z_{t+k})} = \sigma_a^2 \sum_{j=0}^{+\infty} \psi_j^2,$$

vindo que o processo será estacionário se  $\sum_{j=0}^{+\infty} \psi_j^2 < +\infty$ .

Um processo que possa ser representado na forma de médias móveis diz-se puramente não determinístico, pois não contém uma componente determinística que possa ser prevista exactamente a partir do seu próprio passado. Este tipo de processo recebe igualmente o nome de processo linear.

Para uma dada função de autocovariância  $\gamma_k$ ,  $k = 0, \pm 1, \pm 2, \dots$ , a função geradora das autocovariâncias é dada por:

$$(29) \quad \gamma(B) = \sum_{k=-\infty}^{+\infty} \gamma_k B^k,$$

onde  $\gamma_0$ , a variância do processo, é o coeficiente de  $B^0$  e a autocovariância do processo de desfasamento  $k$ ,  $\gamma_k$ , é o coeficiente tanto de  $B^k$  como de  $B^{-k}$ .

Para um processo de médias móveis, a função geradora das autocovariâncias é dada por,

$$(30) \quad \gamma(B) = \sigma_a^2 \psi(B) \psi(B^{-1}) = \sigma_a^2 \sum_{j=0}^{+\infty} \psi_j B^j \sum_{i=0}^{+\infty} \psi_i B^{-i},$$

onde  $j = i + k$  e  $\psi_j = 0$  quando  $j < 0$ .

A correspondente função geradora das autocorrelações é dada por

$$(31) \quad \rho(B) = \sum_{k=-\infty}^{+\infty} \rho_k B^k = \frac{\gamma(B)}{\gamma_0}.$$

Ao escrever um processo  $Z_t$  de acordo com uma representação autorregressiva (AR), faz-se a regressão do valor de  $Z_t$  no tempo  $t$  sobre os seus valores passados mais um choque aleatório, ou seja,

$$(32) \quad \dot{Z}_t = \pi_1 \dot{Z}_{t-1} + \pi_2 \dot{Z}_{t-2} + \dots + a_t,$$

ou, usando uma notação abreviada,

$$(33) \quad \pi(B) \dot{Z}_t = a_t,$$

onde  $\pi(B) = 1 - \sum_{j=1}^{+\infty} \pi_j B^j$  e  $1 + \sum_{j=1}^{+\infty} |\pi_j| < +\infty$ . Para que o processo seja estacionário é necessário que seja possível reescrevê-lo como um processo MA, ou seja,

$$(34) \quad \dot{Z}_t = \frac{1}{\pi(B)} a_t = \psi(B) a_t,$$

em que  $\sum_{j=0}^{+\infty} \psi_j^2 < +\infty$ . Tal condição é satisfeita se as raízes  $\delta$  de  $\pi(B)=0$  se encontrarem fora do círculo unitário, ou seja, se  $|\delta| > 1$ .

Se, na representação autorregressiva de um processo, apenas um número finito dos coeficientes  $\pi_j$  for não nulo, isto é, se  $\pi_1 = \phi_1, \pi_2 = \phi_2, \dots, \pi_p = \phi_p$  e  $\pi_k = 0, \forall k > p$ , diz-se que  $Z_t$  é um processo autorregressivo de ordem  $p$ , AR( $p$ ) e que pode ser escrito da seguinte forma:

$$(35) \quad \phi_p(B) \dot{Z}_t = a_t,$$

onde  $\phi_p(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$ .

De forma análoga, se na representação em médias móveis de um processo apenas um número finito dos coeficientes  $\psi_j$  for não nulo, isto é, se  $\psi_1 = -\theta_1, \psi_2 = -\theta_2, \dots, \psi_q = -\theta_q$  e  $\psi_k = 0, \forall k > q$ , diz-se que  $Z_t$  é um processo de médias móveis de ordem  $q$ , MA( $q$ ), que pode ser escrito da seguinte forma:

$$(36) \quad \dot{Z}_t = \theta(B) a_t,$$

onde  $\theta_q(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$ . Este processo é sempre estacionário.

Por vezes torna-se necessário incluir na construção de um modelo tanto os termos autorregressivo como de médias móveis. Dessa forma obtêm-se os processos mistos autorregressivos de médias móveis, ARMA( $p, q$ ), definidos da seguinte forma:

$$(37) \quad \phi_p(B) \dot{Z}_t = \theta_q(B) a_t,$$

onde  $\phi_p(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$  e  $\theta_q(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$ .

### 2.2.2. O Espectro e a Função Geradora das Autocovariâncias

Como foi referido na secção anterior, a função geradora das autocovariâncias é dada por,

$$(38) \quad \gamma(B) = \sum_{k=-\infty}^{+\infty} \gamma_k B^k.$$

Ora, no caso da função de autocovariância ser absolutamente somável, a função de densidade espectral existe e é dada pela expressão,

$$(39) \quad f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \gamma_k e^{-iak}.$$

Comparando as equações (38) e (39), conclui-se que, para um processo com função de autocovariância absolutamente somável, existe a seguinte relação entre a função de densidade espectral e a função geradora das autocovariâncias:

$$(40) \quad f(\omega) = \frac{1}{2\pi} \gamma(e^{-i\omega}).$$

### 2.2.3. O Espectro dos Modelos ARMA

Como já foi visto, um processo linear não determinístico pode ser escrito na forma

$$(41) \quad Z_t = \sum_{j=0}^{+\infty} \psi_j a_{t-j} = \psi(B)a_t,$$

onde  $\psi(B) = \sum_{j=0}^{+\infty} \psi_j B^j$ ,  $\psi_0 = 1$ . Sem perda de generalidade, considere-se que o valor médio do processo  $Z_t$  é nulo. Como já foi igualmente visto, a função geradora das autocovariâncias é dada por,

$$(42) \quad \gamma(B) = \sigma_a^2 \psi(B)\psi(B^{-1}).$$

Um processo estacionário ARMA( $p, q$ ) pode ser escrito na forma

$$(42) \quad Z_t = \psi(B)a_t = \frac{\theta_q(B)}{\phi_p(B)} a_t = \frac{(1 - \theta_1 B - \dots - \theta_q B^q)}{(1 - \phi_1 B - \dots - \phi_p B^p)} a_t.$$

Assim sendo, a função geradora das autocovariâncias de um processo ARMA( $p, q$ ) é dada por,

$$(43) \quad \gamma(B) = \sigma_a^2 \frac{\theta_q(B)\theta_q(B^{-1})}{\phi_p(B)\phi_p(B^{-1})}.$$

Quando o modelo é estacionário, a função de autocovariância é absolutamente somável, pelo que a função de densidade espectral existe e é dada por,

$$(44) \quad f(\omega) = \frac{1}{2\pi} \gamma(e^{-i\omega}) = \frac{\sigma_a^2}{2\pi} \frac{\theta_q(e^{-i\omega})\theta_q(e^{i\omega})}{\phi_p(e^{-i\omega})\phi_p(e^{i\omega})} = \frac{\sigma_a^2}{2\pi} \left| \frac{\theta_q(e^{-i\omega})}{\phi_p(e^{-i\omega})} \right|^2.$$

### 2.2.4. O Espectro da Soma de Dois Processos Independentes

Seja  $Z_t$  um processo resultante da soma de dois processos estacionários independentes  $X_t$  e  $Y_t$ ,

$$(45) \quad Z_t = X_t + Y_t,$$

sendo  $\gamma_Z(B)$ ,  $\gamma_X(B)$  e  $\gamma_Y(B)$  as funções geradoras das autocovariâncias dos processos. Usando o facto da covariância de variáveis independentes ser nula, temos que:

$$(46) \quad \begin{aligned} \gamma_Z(B) &= \sum_{k=-\infty}^{+\infty} \gamma_Z(k)B^k = \sum_{k=-\infty}^{+\infty} \text{Cov}(X_t + Y_t, X_{t+k} + Y_{t+k})B^k = \\ &= \sum_{k=-\infty}^{+\infty} \gamma_X(k)B^k + \sum_{k=-\infty}^{+\infty} \gamma_Y(k)B^k = \gamma_X(B) + \gamma_Y(B) \end{aligned}$$

Da relação anterior vem que o espectro da soma de dois processos estacionários independentes é a soma dos dois espectros independentes, ou seja:

$$(47) \quad f_Z(\omega) = f_X(\omega) + f_Y(\omega).$$

### 2.2.5. O Espectro de Filtros Lineares

Seja  $Z_t$  um processo estacionário de valor médio nulo com função de autocovariância absolutamente somável e espectro dado por,

$$(48) \quad f_Z(\omega) = \frac{1}{2\pi} \sum_{l=-\infty}^{+\infty} \gamma_Z(l) e^{-i\omega l}.$$

Defina-se o seguinte filtro linear de  $Z_t$ :

$$(49) \quad Y_t = \sum_{j=-\infty}^{+\infty} \alpha_j Z_{t-j} = \alpha(B)Z_t,$$

onde  $\alpha(B) = \sum_{j=-\infty}^{+\infty} \alpha_j B^j$  e  $\sum_{j=-\infty}^{+\infty} |\alpha_j| < +\infty$ . São apenas considerados os filtros lineares invariantes no tempo, ou seja, aqueles cujos coeficientes  $\alpha_j$  não dependem de  $t$ .

Prova-se que as funções espectrais de  $Z_t$  e de  $Y_t$  se relacionam da seguinte forma:

$$(50) \quad f_Y(\omega) = |\alpha(e^{i\omega})|^2 f_Z(\omega),$$

onde

$$(51) \quad |\alpha(e^{i\omega})|^2 = \alpha(e^{i\omega}) [\alpha(e^{i\omega})]^* = \alpha(e^{i\omega}) \alpha(e^{-i\omega}).$$

$\alpha(e^{i\omega})$  é a chamada função transferência.

Existem diversos tipos de filtros, tais como os filtros de “passa-alta” – que eliminam as baixas frequências (associadas com a tendência ou com ondas largas) e preservam as altas frequências –, os filtros de “passa-baixa” – que preservam as componentes de baixas frequências e atenuam ou eliminam as componentes de altas frequências. O cálculo das primeiras diferenças é um exemplo de filtro linear do tipo “passa-alta”, enquanto que o cálculo de médias móveis é um filtro do tipo “passa-baixa”.

## 2.3. Estimação do Espectro

### 2.3.1. Análise do Periodograma

#### 2.3.1.1. O Periodograma e as Suas Propriedades Amostrais

A representação de Fourier

$$(52) \quad Z_t = \sum_{k=0}^{\lfloor n/2 \rfloor} (a_k \cos \omega_k t + b_k \sin \omega_k t)$$

onde  $\omega_k = 2\pi k/n$ ,  $k = 0, 1, \dots, \lfloor n/2 \rfloor$  são as frequências de Fourier e

$$(53) \quad a_k = \begin{cases} \frac{1}{n} \sum_{t=1}^n Z_t \cos \omega_k t, & k = 0, \text{ e } k = n/2 \text{ se } n \text{ é par} \\ \frac{2}{n} \sum_{t=1}^n Z_t \cos \omega_k t, & k = 1, 2, \dots, \left[ \frac{n-1}{2} \right] \end{cases},$$

e

$$(54) \quad b_k = \frac{2}{n} \sum_{t=1}^n Z_t \sin \omega_k t, \quad k = 1, 2, \dots, \left[ \frac{n-1}{2} \right],$$

são os coeficientes de Fourier, pode ser utilizada para representar uma série temporal com  $n$  observações. No entanto, esta não será a representação mais adequada, pois pressupõe que o fenômeno estudado seja estritamente periódico, ajustando um modelo com  $n$  parâmetros a uma sucessão com  $n$  termos.

Uma forma de tornar a representação anterior mais adequada e realista consiste em incluir no modelo um erro ou ruído branco, conduzindo ao modelo de tendência cíclica:

$$(55) \quad Z_t = \sum_{k=0}^{\lfloor n/2 \rfloor} (a_k \cos \omega_k t + b_k \sin \omega_k t) + e_t$$

onde  $e_t$  são variáveis i.i.d.  $N(0, \sigma^2)$  e os restantes parâmetros mantêm-se definidos tal como na representação de Fourier para sucessões discretas e periódicas. O modelo de tendência cíclica pode ser encarado como um modelo de regressão linear, sendo os coeficientes de Fourier essencialmente os coeficientes usuais da regressão, estimados, por exemplo, pelo método dos mínimos quadrados.

Chama-se periodograma à quantidade  $I(\omega_k)$ , definida da seguinte forma:

$$(56) \quad I(\omega_k) = \begin{cases} na_0^2, & k = 0 \\ \frac{n}{2}(a_k^2 + b_k^2) & k = 1, \dots, (n-1)/2 \\ na_{n/2}^2 & k = n/2 \text{ se } n \text{ é par} \end{cases} .$$

Assumindo que  $Z_1, Z_2, \dots, Z_n$  são i.i.d.  $N(0, \sigma^2)$ , vem

$$(57) \quad E(a_k) = \frac{2}{n} \sum_{t=1}^n E(Z_t) \cos \omega_k t = 0$$

e

$$(58) \quad Var(a_k) = \frac{4}{n^2} \sum_{t=1}^n \sigma^2 (\cos \omega_k t)^2 = \frac{2\sigma^2}{n},$$

podendo obter-se resultados semelhantes para  $b_k$ .

Assim sendo,

$$(59) \quad a_k \sim N\left(0, \frac{2\sigma^2}{n}\right) \quad \text{e} \quad b_k \sim N\left(0, \frac{2\sigma^2}{n}\right), \quad k = 1, 2, \dots, \frac{n-1}{2},$$

mostrando-se além disso que

$$(60) \quad Cov(a_k, b_j) = 0, \forall k, j,$$

de onde vem que  $na_k^2/2\sigma^2$ ,  $k = 1, 2, \dots, (n-1)/2$ , são i.i.d com distribuição Qui-quadrado, com um grau de liberdade, tal como  $nb_k^2/2\sigma^2$ . Além disso,  $na_k^2/2\sigma^2$  e  $nb_j^2/2\sigma^2$  são independentes para  $k = 1, 2, \dots, (n-1)/2$  e  $j = 1, 2, \dots, (n-1)/2$ .

Em consequência, as ordenadas do periodograma

$$(61) \quad \frac{I(\omega_k)}{\sigma^2} = \frac{n}{2\sigma^2} (a_k^2 + b_k^2),$$

para  $k = 1, 2, \dots, (n-1)/2$ , são independentes e têm distribuição Qui-quadrado com 2 graus de liberdade.

Seguindo o mesmo raciocínio, mostra-se que tanto  $I(0)/\sigma^2$  como  $I(\pi)/\sigma^2$ , no caso de  $n$  ser par, seguem uma distribuição Qui-quadrado com 1 grau de liberdade.

Com base nos resultados anteriores, é possível testar a hipótese de ser  $m$  o número de frequências das sinusóides no modelo dado por:

$$(62) \quad Z_t = a_t + \sum_{i=1}^m (a_{k_i} \cos \omega_{k_i} t + b_{k_i} \sin \omega_{k_i} t) + e_t$$

onde  $e_t$  são variáveis aleatórias i.i.d.  $N(0, \sigma^2)$ ,  $\omega_{k_i} = 2\pi k_i/n$  e o conjunto  $I = \{k_i : i = 1, \dots, m\}$  é um subconjunto de  $\{k : k = 1, 2, \dots, [n/2]\}$ . A hipótese nula do teste é dada por:

$$(63) \quad H_0 : a_{k_i} = b_{k_i} = 0, \quad i = 1, \dots, m$$

sendo a correspondente estatística de teste

$$(64) \quad F = \frac{(n-2m-1) \sum_{i=1}^m (a_{k_i}^2 + b_{k_i}^2)}{2m \sum_{j \in I} (a_j^2 + b_j^2)}$$

que segue uma distribuição  $F$ , com  $2m$  e  $n-2m-1$  graus de liberdade,  $F(2m, n-2m-1)$ .

### 2.3.1.2. Testes para a Detecção de Componentes Periódicas .

A estatística de teste apresentada na equação (64), em geral, não é aplicável na prática pois pressupõe que as frequências  $\omega_{k_i}$  são conhecidas. No entanto, mesmo que se acredite que a série temporal contenha componentes periódicas, as frequências subjacentes são normalmente desconhecidas. Pode então utilizar-se o periodograma para, através da análise dos seus valores máximos, se detectar tais periodicidades escondidas, pois é de prever que o periodograma, na frequência de Fourier mais perto da verdadeira frequência subjacente, apresente um máximo.

Para se proceder à análise da significância estatística dos máximos do periodograma, toma-se como hipótese nula que as  $[n/2]$  estatísticas dadas por

$$(65) \quad \frac{I(\omega_k)}{\sigma^2} = \frac{n}{2\sigma^2} (a_k^2 + b_k^2), \quad k = 1, 2, \dots, \left[ \frac{n}{2} \right]$$

são variáveis aleatórias i.i.d. com distribuição Qui-quadrado com dois graus de liberdade. Esta hipótese resulta de se admitir que o periodograma foi calculado para uma realização de um ruído branco  $N(0, \sigma^2)$ , como foi visto no parágrafo anterior.

A estatística natural para a realização de um teste sobre a significância estatística do máximo das ordenadas do periodograma será,

$$(66) \quad I^{(1)}(\omega_{(1)}) = \max\{I(\omega_k)\}$$

onde  $\omega_{(1)}$  é a frequência de Fourier para a qual se obtém a ordenada máxima do periodograma. Através da distribuição do máximo de uma amostra de  $[n/2]$  variáveis independentes, sabemos que

$$(67) \quad P\left[\frac{I^{(1)}(\omega_{(1)})}{\sigma^2} > g\right] = 1 - (1 - e^{-(g/2)})^{[n/2]},$$

o que nos vai permitir fazer a dedução da estatística de teste proposta em (66). No entanto, a realização deste teste exacto, usando tal estatística e o facto das variáveis  $I(\omega_k)/\sigma^2$  terem distribuição  $\chi^2_{(2)}$ , implica o conhecimento de  $\sigma^2$ . Ora, na maioria das vezes tal não acontece. Torna-se então necessário estimar  $\sigma^2$ , usando o seguinte estimador,

$$(68) \quad \hat{\sigma}^2 = \frac{1}{2[n/2]} \sum_{k=1}^{[n/2]} I(\omega_k),$$

que é centrado, uma vez que

$$(69) \quad E\left[\sum_{k=1}^{[n/2]} I(\omega_k)\right] = \left[\frac{n}{2}\right] 2\sigma^2$$

e é consistente pois, usando o facto que  $I(\omega_k)$ ,  $k = 1, 2, \dots, [n/2]$  serem independentes vem

$$(70) \quad \text{Var}(\hat{\sigma}^2) \rightarrow 0 \text{ quando } n \rightarrow \infty.$$

Deste modo, para amostras de grande dimensão, a estatística

$$(71) \quad V = \frac{I^{(1)}(\omega_{(1)})}{\left(\frac{1}{2[n/2]}\right) \sum_{k=1}^{[n/2]} I(\omega_k)}$$

tem aproximadamente a mesma distribuição do que  $I^{(1)}(\omega_{(1)})/\sigma^2$ , ou seja,

$$(72) \quad P[V > g] \approx 1 - (1 - e^{-(g/2)})^{[n/2]}.$$

O teste assim descrito tem o nome de teste aproximado de Walker.

Fisher (1929) deduziu igualmente um teste para estudar a significância do máximo do periodograma: o teste exacto de Fisher. Este baseia-se na seguinte estatística,

$$(73) \quad T = \frac{I^{(1)}(\omega_{(1)})}{\sum_{k=1}^{\lfloor n/2 \rfloor} I(\omega_k)}$$

que, sob a mesma hipótese nula (do teste aproximado de Walker), Fisher mostrou verificar a seguinte propriedade,

$$(74) \quad P(T > g) = \sum_{j=1}^m (-1)^{(j-1)} \binom{N}{j} (1 - jg)^{N-1},$$

onde  $N = \lfloor n/2 \rfloor$ ,  $g > 0$  e  $m$  é o maior inteiro menor que  $1/g$ . Usando apenas o primeiro termo do somatório já se consegue uma boa aproximação do seu valor exacto, que é usada em muitas aplicações:

$$(75) \quad P(T > g) \approx N(1 - g)^{N-1}.$$

Para valores grandes da estatística de teste, rejeita-se a hipótese nula, o que implica a existência de uma componente periódica, numa dada frequência  $\omega$ . Embora  $\omega$  não seja necessariamente igual a  $\omega_{(1)}$ , esta é uma boa aproximação da verdadeira frequência da componente periódica, pois Hartley (1949) mostrou que a probabilidade de  $|\omega - \omega_{(1)}| > 2\pi/n$  é menor que o nível de significância do teste.

Sendo  $I^{(2)}(\omega_{(2)})$  a segunda maior ordenada do periodograma, ocorrida na frequência de Fourier  $\omega_{(2)}$ , Whittle (1952) propôs a extensão do teste de Fisher ao estudo da significância dos máximos de ordem superior à primeira. Para tal, sugeriu a utilização da estatística de teste,

$$(76) \quad T_2 = \frac{I^{(2)}(\omega_{(2)})}{\left\{ \sum_{k=1}^{\lfloor n/2 \rfloor} I(\omega_k) \right\} - I^{(1)}(\omega_{(1)})}.$$

Determina-se  $P(T_2 > g)$  substituindo  $N$  por  $N - 1$  na expressão de  $P(T > g)$ . Se o máximo de segunda ordem for significativo, passa-se ao estudo do de terceira ordem, substituindo  $N - 1$  por  $N - 2$ , e assim sucessivamente, até se obter um resultado não significativo. É-se assim conduzido à estimação do valor  $m$ , o número de componentes periódicas presentes na série temporal.

### 2.3.2. O Espectro Amostral

O espectro discreto, associado a um modelo de tendência cíclica que, por sua vez, corresponde ao caso em que a função de distribuição espectral é em forma de escada, pode ser estudado de forma satisfatória pela análise do periodograma. No entanto, os movimentos oscilatórios irregulares, tal como costumam ocorrer no domínio da economia, são representados de forma mais adequada por modelos de espectro contínuo, onde não é uma dada frequência que contribui para a variância do processo, mas um intervalo de frequências.

Ir-se-á estudar, nesta secção, a estimação do espectro para séries temporais com função de autocovariância absolutamente somável, existindo, portanto, função de densidade espectral, como foi visto, dada por,

$$(77) \quad f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \gamma_k e^{-i\omega k} = \frac{1}{2\pi} \gamma_0 + \frac{1}{\pi} \sum_{k=1}^{+\infty} \gamma_k \cos \omega k, \quad -\pi \leq \omega \leq \pi.$$

Notando que, para uma série temporal com  $n$  observações, apenas é possível calcular  $\hat{\gamma}_k$  para  $k=0,1,\dots,n-1$ , a substituição de  $\gamma_k$  por  $\hat{\gamma}_k$  na expressão anterior resulta no seguinte estimador do espectro,

$$(78) \quad \begin{aligned} \hat{f}(\omega) &= \frac{1}{2\pi} \sum_{k=-(n-1)}^{(n-1)} \hat{\gamma}_k e^{-i\omega k} = \frac{1}{2\pi} \hat{\gamma}_0 + \frac{1}{\pi} \sum_{k=1}^{(n-1)} \hat{\gamma}_k \cos \omega k = \\ &= \frac{1}{2\pi} \left( \hat{\gamma}_0 + 2 \sum_{k=1}^{(n-1)} \hat{\gamma}_k \cos \omega k \right) \end{aligned}$$

chamado espectro amostral.

Uma vez que o periodograma pode ser reescrito na seguinte forma,

$$(79) \quad I(\omega_k) = 2 \left( \hat{\gamma}_0 + 2 \sum_{j=1}^{n-1} \hat{\gamma}_j \cos \omega_k j \right),$$

verifica-se a relação que se segue entre o espectro amostral e o periodograma, nas frequências de Fourier,

$$(80) \quad \hat{f}(\omega_k) = \frac{1}{4\pi} I(\omega_k), \quad k = 1, 2, \dots, [n/2]$$

excepto no caso em que  $n$  é par, vindo então  $\hat{f}(\omega_{n/2}) = I(\omega_{n/2})/2\pi = n\alpha_{n/2}^2/2\pi$ .

Se  $Z_t$  for um ruído branco  $N(0, \sigma^2)$ , verifica-se que,

$$(81) \quad \hat{f}(\omega_k) \sim \frac{\sigma^2}{2\pi} \chi^2_{(2)}$$

sendo  $\sigma^2/2\pi$  o espectro  $Z_i$ .

Mostra-se que o estimador do espectro,  $\hat{f}(\omega)$ , goza das seguintes propriedades, quando  $\omega$  pertence à sucessão de Fourier:

- É um estimador centrado;
- Não é um estimador consistente, pois a sua variância não tende para zero quando  $n$  tende para infinito;
- Apresenta um comportamento irregular e errático, pois  $Cov[\hat{f}(\omega_k), \hat{f}(\omega_j)] = 0$ ,  $\forall \omega_k \neq \omega_j$ .

### 2.3.3. O Espectro Alisado

Numa tentativa de encontrar um estimador com melhores propriedades estatísticas, surge a idéia de alisamento, ou combinação linear convenientemente ponderada dos estimadores  $\hat{f}(\omega)$ .

#### 2.3.3.1. Alisamento no Domínio Frequência – A Janela Espectral

Uma forma de reduzir a variância e a irregularidade do espectro amostral consiste em calcular uma média ponderada de  $m$  valores do espectro, à esquerda e à direita da frequência alvo  $\omega_k$ , ou seja,

$$(82) \quad \hat{f}_W(\omega_k) = \sum_{j=-m_n}^{m_n} W_n(\omega_j) \hat{f}(\omega_k - \omega_j)$$

onde  $\omega_k = 2k\pi/n$  são as frequências de Fourier, sendo  $n$  a dimensão da amostra,  $k = 0, \pm 1, \pm 2, \dots, \pm [n/2]$  e  $W_n(\omega_j)$  é a função de ponderação que goza das seguintes propriedades:

$$(83) \quad \sum_{j=-m_n}^{m_n} W_n(\omega_j) = 1$$

$$(84) \quad W_n(\omega_j) = W_n(-\omega_j)$$

e

$$(85) \quad \lim_{n \rightarrow \infty} \sum_{j=-m_n}^{m_n} W_n^2(\omega_j) = 0.$$

À função  $W_n(\omega_j)$  assim definida chama-se janela espectral. As frequências incluídas na expressão de alisamento,  $\omega_k - \omega_j, -m_n \leq j \leq m_n$ , constituem a banda de frequências da janela espectral e o número de frequências,  $2m_n + 1$ , está directamente relacionado com a largura da janela espectral, conhecida como largura da banda.

Se os valores do “verdadeiro” espectro  $f(\omega)$  forem aproximadamente constantes dentro da banda, temos que,

$$(86) \quad \begin{aligned} E[\hat{f}_W(\omega_k)] &= \sum_{j=-m_n}^{m_n} W_n(\omega_j) E[\hat{f}(\omega_k - \omega_j)] \approx \\ &\approx f(\omega_k) \sum_{j=-m_n}^{m_n} W_n(\omega_j) = f(\omega_k) \end{aligned}$$

e

$$(87) \quad \text{Var}[\hat{f}_W(\omega_k)] \approx [f(\omega_k)]^2 \sum_{j=-m_n}^{m_n} W_n^2(\omega_j).$$

Assim sendo, a variância do espectro alisado diminui quando  $m_n$  aumenta, resultando também desse aumento um estimador mais estável. Por outro lado, à medida que a largura da banda aumenta, aumenta o número de ordenadas espectrais a ser incluídas na média ponderada, o que, exceptuando o caso em que  $f(\omega)$  é constante, resulta num aumento do enviesamento. Torna-se necessário, portanto, encontrar uma solução de compromisso entre a diminuição da variância e o aumento do enviesamento.

Considere-se agora o caso mais geral em que o espectro amostral está definido no intervalo contínuo  $[-\pi, \pi]$  e não apenas nas frequências de Fourier. O espectro alisado é agora dado por:

$$(88) \quad \hat{f}_W(\omega) = \int_{-\pi}^{\pi} W_n(\lambda) \hat{f}(\omega - \lambda) d\lambda = \int_{-\pi}^{\pi} W_n(\omega - \lambda) \hat{f}(\lambda) d\lambda$$

onde  $W_n(\lambda)$  é a janela espectral, satisfazendo as seguintes condições:

$$(89) \quad \int_{-\pi}^{\pi} W_n(\lambda) d\lambda = 1,$$

$$(90) \quad W_n(\lambda) = W_n(-\lambda)$$

e

$$(91) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \int_{-\pi}^{\pi} W_n^2(\lambda) d\lambda = 0.$$

Se o espectro for aproximadamente constante dentro da banda considerada, temos que,

$$(92) \quad E[\hat{f}_W(\omega)] = f(\omega)$$

e

$$(93) \quad \text{Var}[\hat{f}_W(\omega)] \approx \frac{2\pi}{n} f^2(\omega) \int_{-\pi}^{\pi} W_n^2(\lambda) d\lambda.$$

Coloca-se desta forma a mesma questão do potencial enviesamento do estimador quando a largura da banda aumenta de forma a diminuir a variância do mesmo.

### 2.3.3.2. Alisamento no Domínio Tempo – A Janela Retardada

Como foi mostrado, o espectro  $f(\omega)$  e função de autocovariância  $\gamma_k$  constituem um par de transformadas de Fourier. Assim sendo, em alternativa ao alisamento no domínio frequência pode proceder-se ao alisamento no domínio tempo, aplicando uma função de ponderação  $W(k)$  à função de autocovariância amostral, ou seja,

$$(94) \quad \hat{f}(\omega) = \frac{1}{2\pi} \sum_{k=-(n-1)}^{(n-1)} W(k) \hat{\gamma}_k e^{-i\omega k}.$$

Uma vez que  $\hat{\gamma}_k$  é uma função par e que para valores grandes de  $k$  os estimadores  $\hat{\gamma}_k$  são muito imprecisos pois envolvem um número reduzido de pares de observações,  $W(k)$  deve ser simétrica e com pesos inversamente proporcionais à magnitude de  $k$ .

Obtém-se então:

$$(95) \quad \hat{f}(\omega) = \frac{1}{2\pi} \sum_{k=-M}^M W_n(k) \hat{\gamma}_k e^{-i\omega k},$$

onde  $W_n(k)$  é uma sucessão absolutamente somável, derivada a partir de uma função par, contínua e limitada  $W(x)$ , definida por  $W_n(k) = W(k/M)$ . A função  $W(x)$  verifica as seguintes condições:

$$(96) \quad \begin{aligned} |W(x)| &\leq 1, \\ W(0) &= 1, \\ W(x) &= W(-x), \\ W(x) &= 0, \quad |x| > 1. \end{aligned}$$

O valor  $M$  é o chamado ponto de truncagem, dependente da dimensão da amostra,  $n$ . A função  $W_n(k)$  tem o nome de janela retardada. O estimador resultante da aplicação de uma janela diz-se um estimador alisado do espectro.

Assumindo que  $M = M_n$ , o ponto de truncagem, é uma função de  $n$  tal que  $M \rightarrow \infty$  e  $M/n \rightarrow \infty$  quando  $n \rightarrow \infty$  e admitindo que se verifica a relação  $Z_t = \sum_{k=-\infty}^{\infty} \psi_k X_{t-k}$ ,

onde  $\{X_t\} \sim IID(0, \sigma^2)$ ,  $\sum_{k=-\infty}^{\infty} |\psi_k| |k|^{1/2} < \infty$  e  $EX_t^4 < \infty$ , é possível mostrar que o estimador alisado do espectro, já definido, é consistente no sentido da média quadrática.

Mostra-se que as janelas espectral e retardada verificam as seguintes relações:

$$(97) \quad W_n(\omega) = \frac{1}{2\pi} \sum_{k=-M}^M W_n(k) e^{-i\omega k}$$

e

$$(98) \quad W_n(k) = \int_{-\pi}^{\pi} W_n(\omega) e^{i\omega k} d\omega, \quad k = 0, \pm 1, \dots, \pm M.$$

Ou seja, a janela espectral é a transformada de Fourier da janela retardada e a janela retardada é a transformada de Fourier inversa da janela espectral, constituindo ambas um par de transformadas de Fourier.

### 2.3.3.3. Principais Janelas Utilizadas

A dificuldade em decidir qual a melhor janela a utilizar em cada caso concreto é testemunhada pela existência de várias propostas de janelas para o alisamento do estimador do espectro. Como foi dito, é necessário encontrar uma solução de compromisso entre uma maior estabilidade (menor variância) e uma maior resolução (menor enviesamento).

De seguida, são apresentadas as principais janelas apresentadas na prática.

#### Janela Rectangular

Esta janela é derivada da seguinte função:

$$(99) \quad W(x) = \begin{cases} 1, & |x| \leq 1 \\ 0, & |x| > 1 \end{cases}$$

A janela retardada é definida da seguinte forma:

$$W_n^R(k) = \begin{cases} 1, & |k| \leq M \\ 0, & |k| > M \end{cases}$$

E a correspondente janela espectral toma a forma:

$$(101) \quad W_n^R(\omega) = \frac{1}{2\pi} \sum_{k=-M}^M W_n^R(k) e^{-i\omega k} = \frac{1}{2\pi} \sum_{k=-M}^M e^{-i\omega k} = \frac{1}{2\pi} \frac{\text{sen}(\omega(M+1/2))}{\text{sen}(\omega/2)}$$

O uso desta janela pode conduzir à obtenção de valores negativos para algumas frequências, resultado indesejável visto que o espectro é uma função não negativa. Esta janela tem também o inconveniente de apresentar lóbulos subsidiários que levam a uma ponderação mais elevada do que o desejável de ordenadas correspondentes a frequências algo afastadas da frequência alvo, fenómeno este conhecido por infiltração.

Para uma dada janela espectral, com máximo atingido para  $\omega = 0$ , pode considerar-se que a largura de banda é dada pela distância entre as frequências pertencentes ao lóbulo principal correspondentes aos pontos onde a janela atinge metade do seu valor máximo. Ou seja, a largura da banda é dada por  $2\omega_1$ , onde  $\omega_1$  é tal que

$$(102) \quad W_n(\pm \omega_1) = \frac{1}{2} W_n(0).$$

O valor máximo de  $W_n^R(\omega)$ , atingido para  $\omega = 0$ , é  $(2M+1)/2\pi$ . Os seus zeros ocorrem para  $\omega = \pm 2j\pi/(2M+1)$ ,  $j = 1, 2, \dots$ . Considerando que o valor de  $\omega_1$  pode ser aproximado pela metade do valor do primeiro zero da janela, vem que a largura de banda pode ser aproximada pelo valor do primeiro zero, ou seja,

$$(103) \quad 2\omega_1 \approx \frac{2\pi}{2M+1}.$$

Conclui-se, portanto, que a largura de banda varia inversamente com o ponto de truncagem. Assim sendo, quando  $M$  aumenta, a largura de banda diminui, bem como o enviesamento do estimador, aumentando, no entanto, a sua variância.

Nas próximas janelas apresentadas, continua a ser válida esta relação inversa entre a largura de banda e o ponto de truncagem.

### Janela de Bartlett

Esta janela foi proposta por Bartlett (1950) e baseia-se na seguinte função:

$$(104) \quad W(x) = \begin{cases} 1 - |x|, & |x| \leq 1 \\ 0, & |x| > 1 \end{cases}$$

A janela retardada é definida da seguinte forma,

$$(105) \quad W_n^B(k) = \begin{cases} 1 - \frac{|k|}{M}, & |k| \leq M \\ 0, & |k| > M \end{cases}$$

e a correspondente janela espectral toma a forma

$$(106) \quad W_n^B(\omega) = \frac{1}{2\pi} \sum_{k=-M}^M \left(1 - \frac{|k|}{M}\right) e^{-iak} = \frac{1}{2\pi M} \left[ \frac{\text{sen}(\omega M/2)}{\text{sen}(\omega/2)} \right]^2.$$

A correspondente largura de banda pode tomar-se como sendo igual a  $3\pi/M$ .

A janela espectral de Bartlett é não negativa, vindo o estimador alisado resultante é também não negativo. Desta forma, é ultrapassado um dos inconvenientes da janela rectangular. Além disso, os lóbulos subsidiários da janela de Bartlett são menores que os da janela rectangular, continuando, no entanto, a existir o efeito de infiltração.

Os lóbulos laterais da janela espectral são resultado da transformada de Fourier de pontos angulosos da janela retardada. Consequentemente, a forma de eliminar o efeito de infiltração consiste em especificar janelas retardadas sem pontos angulosos.

### Janela de Blackman-Tuckey

Esta janela foi proposta por Blackman e Tuckey (1958) e deriva da seguinte função:

$$(107) \quad W(x) = \begin{cases} 1 - 2a + 2a \cos(\pi x), & |x| \leq 1 \\ 0, & |x| > 1 \end{cases}$$

sendo a janela retardada dada por:

$$(108) \quad W_n^T(k) = \begin{cases} 1 - 2a + 2a \cos(\pi k/M), & |k| \leq M \\ 0, & |k| > M \end{cases}$$

e a correspondente janela espectral dada por:

$$(109) \quad \begin{aligned} W_n^T(\omega) &= \frac{1}{2\pi} \sum_{k=-M}^M [1 - 2a + 2a \cos(\pi k/M)] e^{-iak} = \\ &= \frac{a}{2\pi} \frac{\text{sen}[(\omega - \pi/M)(M + 1/2)]}{\text{sen}[(\omega - \pi/M)/2]} + \frac{(1 - 2a)}{2\pi} \frac{\text{sen}[\omega(M + 1/2)]}{\text{sen}(\omega/2)} + \\ &= \frac{a}{2\pi} \frac{\text{sen}[(\omega + \pi/M)(M + 1/2)]}{\text{sen}[(\omega + \pi/M)/2]} \end{aligned}$$

A constante  $a$  é escolhida tal que  $0 < a \leq 0,25$  e  $W_n^T(k) \geq 0$  para todo o  $k$ . A largura de banda desta janela é, aproximadamente,  $1,33 \cdot 2\pi/M$ .

Comparando a expressão da janela espectral de Blackman-Tuckey com a expressão da janela espectral rectangular, deduz-se a seguinte relação,

$$(110) \quad W_n^T(\omega) = aW_n^R(\omega - \pi/M) + (1 - 2a)W_n^R(\omega) + aW_n^R(\omega + \pi/M),$$

a partir da qual se pode concluir que a janela de Blackman-Tuckey pode conduzir a estimativas do espectro negativas.

Para  $a = 0,23$ , a janela também é conhecida como janela de Tuckey-Hamming e, para  $a = 0,25$ , a janela é conhecida como janela de Tuckey-Hanning.

### Janela de Parzen

A janela sugerida por Parzen (1961) é baseada na função

$$(111) \quad W(x) = \begin{cases} 1 - 6x^2 + 6|x|^3, & |x| \leq 1/2 \\ 2(1 - |x|)^3, & 1/2 < |x| \leq 1. \\ 0, & |x| > 1 \end{cases}$$

A janela retardada é dada por,

$$(112) \quad W_n^P(k) = \begin{cases} 1 - 6(k/M)^2 + 6(|k|/M)^3, & |k| \leq M/2 \\ 2(1 - |k|/M)^3, & M/2 < |k| \leq M. \\ 0, & |k| > M \end{cases}$$

e a janela espectral correspondente toma a forma,

$$(113) \quad W_n^P(\omega) = \frac{1}{2\pi} \left\{ \sum_{k=-M/2}^{M/2} \left[ 1 - 6\left(\frac{k}{M}\right)^2 + 6\left(\frac{|k|}{M}\right)^3 \right] e^{-iak} + 2 \sum_{M/2 < |k| \leq M} \left( 1 - \frac{|k|}{M} \right)^3 e^{-iak} \right\} =$$

$$= \frac{3}{8\pi M^3} \left[ \frac{\text{sen}(\omega M/4)}{(1/2)\text{sen}(\omega/2)} \right]^4 \left\{ 1 - \frac{2}{3} \left[ \text{sen}\left(\frac{\omega}{2}\right) \right]^2 \right\}$$

e, para  $M$  grande,

$$(114) \quad W_n^P(\omega) \approx \frac{3}{8\pi M^3} \left[ \frac{\text{sen}(\omega M/4)}{(1/2)\text{sen}(\omega/2)} \right]^4.$$

Esta janela conduz a estimativas do espectro não negativas. A sua largura de banda é, aproximadamente,  $1,86 \cdot 2\pi/M$ .

#### 2.3.3.4. Escolha da Janela

A qualidade dos estimadores do espectro alisado depende tanto da forma da janela como do ponto de truncagem escolhido, sendo, no entanto, este último um factor mais preponderante nos resultados conseguidos. Como já referido, a largura da banda está inversamente relacionada com o ponto de truncagem utilizado na janela retardada e, se por um lado a resolução do estimador é função decrescente da largura de banda, por outro lado, a estabilidade é função crescente da largura de banda.

Não existindo qualquer critério de optimização do ponto de truncagem, existem no entanto algumas regras de procedimentos a seguir.

Jenkins e Watts (1968) sugerem que se empregue a técnica do sucessivo “fecho da janela”, consistindo esta na estimação do espectro com uma janela com um ponto de truncagem  $M$  relativamente pequeno e proceder, em seguida, a novas estimativas aumentando progressivamente o valor de  $M$ . Deve parar-se o estreitamento da banda (correspondente ao aumento do valor de truncagem) quando se conseguiu o desejado aumento de resolução sem grandes prejuízos de estabilidade. Para tal, pode seguir-se as seguintes regras práticas:

- a) Escolhem-se três valores para o ponto de truncagem,  $M_1$ ,  $M_2$  e  $M_3 = 4M_1$ . Para cada um destes valores, estima-se o espectro para averiguar se é necessário alargar o estudo para algum  $M < M_1$ , para algum  $M_1 < M < M_3$  ou para algum  $M > M_3$ ;
- b) Pode escolher-se  $M$ , tal que  $\hat{y}(k) \approx 0$  para  $k > M$ ;
- c) Se o número de termos da sucessão,  $n$ , não for muito elevado, pode tomar-se  $M = \lceil n/10 \rceil$ ;
- d) Se, *a priori*, se tiver conhecimento da existência de “picos” do espectro nas frequências  $\lambda_1, \lambda_2, \lambda_3, \dots$ , pretendendo distingui-los, é necessário que a largura da banda não exceda o intervalo mínimo entre “picos” adjacentes. É comum que se escolha a largura de banda igual a  $\min_i |\lambda_{i+1} - \lambda_i|$ .

### Capítulo 3: Análise Espectral de Processos Estocásticos Multivariados

Até agora, os resultados apresentados aplicavam-se a séries temporais univariadas, no entanto, em muitos casos, os dados em estudo provêm de observações de várias variáveis, existindo uma forte interdependência entre diversas séries cronológicas. Nestes casos, torna-se mais indicado considerar que as séries temporais em estudo são componentes de uma série temporal multivariada  $\mathbf{X}_t$ , em que não existe apenas dependência sequencial em cada componente  $X_{it}$  mas também existe interdependência entre diferentes componentes  $X_{it}$  e  $X_{jt}$ .

#### 3.1. Conceitos Gerais

Seja  $\mathbf{X}_t = [X_{t1}, X_{t2}, \dots, X_{tm}]$ ,  $t = 0, \pm 1, \pm 2, \dots$  uma série temporal  $m$  dimensional, verificando-se  $E(X_{it}^2) < \infty$  para todo o  $t$  e todo o  $i$ .

As propriedades de segunda ordem de  $\mathbf{X}_t$  são especificadas pelos vectores valor médio e pelas matrizes de covariância:

$$(1) \quad \boldsymbol{\mu}_t = E(\mathbf{X}_t) = [\mu_{t1}, \dots, \mu_{tm}]'$$

$$(2) \quad \Gamma(t+k, t) = E[(\mathbf{X}_{t+k} - \boldsymbol{\mu}_{t+k})(\mathbf{X}_t - \boldsymbol{\mu}_t)'] = [\gamma_{ij}(t+k, t)]_{i,j=1}^m.$$

Se  $\mathbf{X}_t$  tivesse componentes de valor complexo, então as matrizes de covariância seriam dadas por:

$$(3) \quad \Gamma(t+k, t) = E[(\mathbf{X}_{t+k} - \boldsymbol{\mu}_{t+k})(\mathbf{X}_t - \boldsymbol{\mu}_t)^*]$$

onde  $*$  representa o transposto do complexo conjugado. No entanto, nos próximos resultados, salvo onde indicado, considerar-se-á que as componentes de  $\mathbf{X}_t$  tomam valores reais.

Diz-se que uma série  $\mathbf{X}_t$ , com valor médio  $\boldsymbol{\mu}_t$  e matrizes de covariância  $\Gamma(t+k, t)$ ,  $k = 0, \pm 1, \pm 2, \dots$ , é estacionária se  $\boldsymbol{\mu}_t$  e  $\Gamma(t+k, t)$  são funções independentes de  $t$ . No caso das sucessões estacionárias, usa-se a seguinte notação:

$$(4) \quad \boldsymbol{\mu} = E(\mathbf{X}_t) = [\mu_{11}, \dots, \mu_{1m}]'$$

e

$$(5) \quad \Gamma(k) = E[(\mathbf{X}_{t+k} - \boldsymbol{\mu})(\mathbf{X}_t - \boldsymbol{\mu})'] = [\gamma_{ij}(k)]_{i,j=1}^m, \quad k = 0, \pm 1, \pm 2, \dots$$

Como função de  $k$ ,  $\Gamma(k)$  é chamada a função matriz de covariância. Para cada  $i = j$ ,  $\gamma_{ii}(k)$  é a função de autocovariância de  $X_{ii}$ ; para  $i \neq j$ ,  $\gamma_{ij}(k)$  é a função de covariância cruzada entre  $X_{ii}$  e  $X_{jj}$ . Cada componente da série estacionária multivariada é também estacionário.

A função matriz de correlações é definida como:

$$(6) \quad \rho(k) = \mathbf{D}^{-1/2} \Gamma(k) \mathbf{D}^{-1/2} = [\rho_{ij}(k)]_{i,j=1}^m = \left[ \frac{\gamma_{ij}(k)}{[\gamma_{ii}(0)\gamma_{jj}(0)]^{1/2}} \right]_{i,j=1}^m,$$

onde  $\mathbf{D} = \text{diag}[\gamma_{11}(0), \gamma_{22}(0), \dots, \gamma_{mm}(0)]$ . Para  $i = j$ ,  $\rho_{ii}(k)$  é a função de autocorrelação de  $X_{ii}$ ; para  $i \neq j$ ,  $\rho_{ij}(k)$  é a função de correlação cruzada entre  $X_{ii}$  e  $X_{jj}$ .

A função matriz de covariâncias de uma série temporal estacionária goza das seguintes propriedades:

- (i)  $\Gamma(k) = \Gamma'(-k)$
- (ii)  $|\gamma_{ij}(k)| \leq [\gamma_{ii}(0)\gamma_{jj}(0)]^{1/2}$ ,  $i, j = 1, \dots, m$
- (iii)  $\gamma_{ii}(k)$  é uma função de autocovariância  $i = 1, \dots, m$
- (iv)  $\sum_{j,k=1}^n \mathbf{a}'_j \Gamma(j-k) \mathbf{a}_k \geq 0$ ,  $\forall n \in \{1, 2, \dots\}$  e  $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$ .

As propriedades anteriores são igualmente válidas para a função matriz de autocorrelações que goza ainda da seguinte propriedade:

$$(v) \quad \rho_{ii}(0) = 1.$$

Desde que a função matriz de covariâncias verifique a condição  $\sum_{k=-\infty}^{+\infty} |\gamma_{ij}(k)| < \infty$ ,  $i, j = 1, \dots, m$ , esta tem uma transformada de Fourier, chamada função matriz de densidade espectral ou espectro, dada por:

$$(7) \quad \mathbf{f}(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \Gamma(k) e^{-iak}, \quad -\pi \leq \omega \leq \pi.$$

Nestas condições pode escrever-se:

$$(8) \quad \Gamma(k) = \int_{-\pi}^{\pi} e^{iak} \mathbf{f}(\omega) d\omega$$

e  $\mathbf{X}_t$  também tem uma representação espectral, dada por:

$$(9) \quad X_t = \int_{-\pi}^{\pi} e^{i\omega t} dU(\omega),$$

onde  $\{U(\omega), -\pi \leq \omega \leq \pi\}$  é um processo cujos componentes são processos de incremento ortogonais satisfazendo

$$(10) \quad E[dU_j(\omega)dU_k^*(\lambda)] = \begin{cases} f_{jk}(\omega)d\omega & \text{se } \omega = \lambda \\ 0 & \text{se } \omega \neq \lambda \end{cases}$$

Para terminar este parágrafo, indica-se como uma matriz complexa pode ser escrita às custas de uma matriz real e de que forma estas se relacionam. Seja  $Z = [Z_{jk}]$  onde  $Z_{jk} = \text{Re} Z_{jk} + i \text{Im} Z_{jk}$ . Vem portanto que  $\text{Re} Z = [\text{Re} Z_{jk}]$  e  $\text{Im} Z = [\text{Im} Z_{jk}]$ . É possível mostrar que a cada matriz complexa  $Z_{(J \times K)}$  corresponde uma matriz real  $Z^R_{(2J \times 2K)}$ , dada por

$$(11) \quad Z^R = \begin{bmatrix} \text{Re} Z & \text{Im} Z \\ -\text{Im} Z & \text{Re} Z \end{bmatrix}$$

e satisfazendo as seguintes condições:

- i) se  $Z = X + Y$ , então  $Z^R = X^R + Y^R$
- ii) se  $Z = XY$ , então  $Z^R = X^R Y^R$
- iii) se  $Y = Z^{-1}$ , então  $Y^R = (Z^R)^{-1}$
- iv)  $\text{Det} Z^R = |\text{Det} Z|^2$
- v) se  $Z$  é Hermitiana,  $Z^R$  é simétrica
- vi) se  $Z$  é unitária,  $Z^R$  é ortogonal
- vii) se os valores e vectores próprios de  $Z$  são  $\mu_j, \alpha_j, j = 1, \dots, J$ , então os valores e vectores próprios de  $Z^R$  são

$$(12) \quad \mu_j, \begin{bmatrix} \text{Re} \alpha_j \\ \text{Im} \alpha_j \end{bmatrix}; \quad \mu_j, \begin{bmatrix} -\text{Im} \alpha_j \\ \text{Re} \alpha_j \end{bmatrix} \quad j = 1, \dots, J.$$

### 3.2.O Caso Bivariado

Nesta secção, estendem-se os resultados apresentados no capítulo 2 ao caso de uma série temporal bivariada estacionária. As ideias apresentadas para o caso bivariado são estendidas ao caso multivariado na secção 3.3..

### 3.2.1. O Espectro Cruzado

Seja  $\mathbf{X}_t = [X_{t1}, X_{t2}]'$  uma série temporal bivariada estacionária, com média nula,  $\mathbf{0}$ , e função matriz de covariâncias  $\Gamma(k) = [\gamma_{ij}(k)]_{i,j=1,2}$ , onde  $\gamma_{ij}(k) = E(X_{t+k,i} X_{t,j})$ , satisfazendo  $\sum_{k=-\infty}^{+\infty} |\gamma_{ij}(k)| < \infty$ ,  $i, j = 1, 2$ , então, a função

$$(13) \quad f_{12}(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \gamma_{12}(k) e^{-i\omega k}, \quad -\pi \leq \omega \leq \pi,$$

chama-se espectro cruzado ou densidade espectral cruzada de  $X_{t1}$  e  $X_{t2}$ . A matriz

$$(14) \quad \mathbf{f}(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \Gamma(k) e^{-i\omega k} = \begin{bmatrix} f_{11}(\omega) & f_{12}(\omega) \\ f_{21}(\omega) & f_{22}(\omega) \end{bmatrix}$$

chama-se matriz de densidade espectral ou espectro de  $\mathbf{X}_t$ .

Deste modo, as representações espectrais de  $\gamma_{ij}(k)$  e de  $\Gamma(k)$  são:

$$(15) \quad \gamma_{ij}(k) = \int_{-\pi}^{\pi} e^{i\omega k} f_{ij}(\omega) d\omega, \quad i, j = 1, 2$$

e

$$(16) \quad \Gamma(k) = \int_{-\pi}^{\pi} e^{i\omega k} \mathbf{f}(\omega) d\omega.$$

A função  $f_{ii}(\omega)$  é o espectro da série univariada  $X_{ti}$ , tal como definido no capítulo 2, sendo, portanto, real e par. No entanto, como  $\gamma_{ij}(k)$ , para  $i \neq j$ , não é uma função par, o espectro cruzado toma normalmente valores complexos. As matrizes  $\mathbf{f}(\omega)$  são Hermitianas, ou seja,  $\mathbf{f}(\omega) = \mathbf{f}^*(\omega)$ , e definidas não negativas, ou seja,  $\mathbf{a}^* \mathbf{f}(\omega) \mathbf{a} \geq 0$ ,  $\forall \mathbf{a} \in \mathbb{C}^2$ .

É possível, então, escrever:

$$(17) \quad f_{12}(\omega) = c_{12}(\omega) - iq_{12}(\omega)$$

onde

$$(18) \quad c_{12}(\omega) = \text{Re}\{f_{12}(\omega)\} = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \gamma_{12}(k) \cos \omega k$$

e

$$(19) \quad -q_{12}(\omega) = \text{Im}\{f_{12}(\omega)\} = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \gamma_{12}(k) \text{sen} \omega k.$$

A função  $c_{12}(\omega)$  chama-se coespectro de  $X_{i1}$  e  $X_{i2}$ ; a função  $q_{12}(\omega)$  chama-se espectro quadratura.

Expressando  $f_{12}(\omega)$  em coordenadas polares vem:

$$(20) \quad f_{12}(\omega) = \alpha_{12}(\omega) \exp\{i\phi_{12}(\omega)\}$$

onde a função dada por:

$$(21) \quad \alpha_{12}(\omega) = |f_{12}(\omega)| = \sqrt{c_{12}^2(\omega) + q_{12}^2(\omega)}$$

é chamada espectro amplitude e a função correspondente a:

$$(22) \quad \phi_{12}(\omega) = \arg[c_{12}(\omega) - iq_{12}(\omega)] = \tan^{-1} \left[ -\frac{q_{12}(\omega)}{c_{12}(\omega)} \right]$$

é chamada espectro fase de  $X_{i1}$  e  $X_{i2}$ .

Outras duas funções, úteis para a análise bivariada no domínio frequência, são a função de coerência e a função ganho. A função de coerência (coherence ou squared coherency) é definida por:

$$(23) \quad K_{12}^2(\omega) = \frac{|f_{12}(\omega)|^2}{f_{11}(\omega)f_{22}(\omega)},$$

que no fundo é o espectro amplitude estandardizado.

A função ganho é definida como:

$$(24) \quad G_{12}(\omega) = \frac{|f_{12}(\omega)|}{f_{11}(\omega)} = \frac{\alpha_{12}(\omega)}{f_{11}(\omega)}.$$

### 3.2.2. Interpretação das Funções Espectrais Cruzadas

Considere-se que  $X_i$  e  $U(\omega)$  são os dois processos definidos no parágrafo 3.1., verificando as condições neste indicadas e verificando (9) e (10). Assim sendo, é possível escrever:

$$(25) \quad f_i(\omega)d\omega = f_{ii}(\omega)d\omega = E[dU_i(\omega)]^2, \quad i = 1, 2.$$

e

$$(26) \quad f_{12}(\omega)d\omega = E[dU_1(\omega)dU_2^*(\omega)].$$

Ora,  $E[dU_1(\omega)dU_2^*(\omega)] = \text{cov}(dU_1(\omega), dU_2(\omega))$ . Assim sendo, o espectro amplitude,  $\alpha_{12}(\omega) = |f_{12}(\omega)|$ , mede a covariância entre a componente de frequência  $\omega$  de  $X_{i1}$  e a componente de frequência  $\omega$  de  $X_{i2}$ .

Da mesma forma, notando que  $E[dU_i(\omega)]^2 = \text{var}[dU_i(\omega)]$ ,  $i = 1, 2$ , vê-se que

$$(27) \quad K_{12}^2(\omega) = \frac{[f_{12}(\omega)]^2}{f_{11}(\omega)f_{22}(\omega)} = \frac{\{\text{cov}(dU_1(\omega), dU_2(\omega))\}}{\text{var}[dU_1(\omega)]\text{var}[dU_2(\omega)]}.$$

Conseqüentemente,  $K_{12}^2(\omega)$  é o quadrado do coeficiente de correlação entre a componente de frequência  $\omega$  de  $X_{i1}$  e a componente de frequência  $\omega$  de  $X_{i2}$ , sendo óbvio que  $0 \leq K_{12}^2(\omega) \leq 1$ . Um valor perto de 1 implica que essas componentes de  $X_{i1}$  e  $X_{i2}$  têm uma forte relação linear.

De forma semelhante, vem:

$$(28) \quad G_{12}(\omega) = \frac{|f_{12}(\omega)|}{f_{11}(\omega)} = \frac{|E[dU_1(\omega)dU_2^*(\omega)]|}{E[dU_1(\omega)]^2} = \frac{|\text{cov}(dU_1(\omega), dU_2(\omega))|}{\text{var}[dU_1(\omega)]},$$

o que implica que a função ganho é o valor absoluto do coeficiente de  $X_{i1}(\omega)$  na regressão linear de  $X_{i2}(\omega)$  sobre  $X_{i1}(\omega)$ . Esta função representa o *ganho* de  $X_{i2}$  sobre  $X_{i1}$  na componente de frequência  $\omega$ .

Para interpretar o espectro fase, note-se que para todo o  $\omega$ ,  $dU_j(\omega)$ ,  $j = 1, 2$ , são variáveis aleatórias que tomam valores complexos, sendo possível escrever:

$$(29) \quad dU_j(\omega) = \alpha_j(\omega) \exp\{i\phi_j(\omega)\},$$

onde  $\alpha_j(\omega)$  e  $\phi_j(\omega)$  são o espectro amplitude e o espectro fase da série  $X_{ij}$ ,  $j = 1, 2$ . Então,

$$(30) \quad \alpha_{12}(\omega) \exp\{i\phi_{12}(\omega)\} = f_{12}(\omega) = E[dU_1(\omega)dU_2(\omega)] = \\ = E[\alpha_1(\omega)\alpha_2(\omega)]E[\exp\{i(\phi_1(\omega) - \phi_2(\omega))\}],$$

assumindo a independência do espectro amplitude e do espectro fase.

Pode então considerar-se que o espectro amplitude  $\alpha_{12}(\omega)$  é o valor médio do produto das amplitudes das componentes de frequência  $\omega$  de  $X_{i1}$  e  $X_{i2}$ .

O espectro fase representa o desfasamento médio,  $[\phi_1(\omega) - \phi_2(\omega)]$ , entre as componentes de amplitude  $\omega$  de  $X_{t1}$  e  $X_{t2}$ . Indica em que medida uma série é avançada, atrasada ou coincidente em relação a outra. Diz-se que a componente de frequência  $\omega$  de  $X_{t1}$  é avançada em relação à componente de frequência  $\omega$  de  $X_{t2}$  se a fase  $\phi_{12}(\omega)$  é negativa. Diz-se que a componente de frequência  $\omega$  de  $X_{t1}$  é atrasada em relação à componente de frequência  $\omega$  de  $X_{t2}$  se a fase  $\phi_{12}(\omega)$  é positiva. O desfasamento correspondente, no domínio tempo, é dado por:

$$(31) \quad \tau = -\frac{\phi_{12}(\omega)}{\omega}.$$

A interpretação do coespectro e do espectro quadratura é realizada no parágrafo 3.3., relativo ao caso multivariado.

### 3.2.3. Filtros Lineares Invariantes no Tempo

Considere-se que se aplicam filtros lineares invariantes no tempo, tal como definidos na secção 2.2.5., às componentes de  $X_t$ , vindo:

$$(32) \quad Y_{t1} = \sum_{k=-\infty}^{+\infty} \alpha_{k1} X_{t-k,1} = \alpha_1(B)X_{t1}$$

e

$$(33) \quad Y_{t2} = \sum_{k=-\infty}^{+\infty} \alpha_{k2} X_{t-k,2} = \alpha_2(B)X_{t2}.$$

Assim sendo, de acordo com os resultados enunciados para o caso univariado,

$$(34) \quad f_{Y_j}(\omega) = |\alpha_j(e^{-i\omega})|^2 f_{X_j}(\omega), \quad j = 1,2,$$

mostrando-se também que

$$(35) \quad f_{Y_1 Y_2}(\omega) = \alpha_1(e^{-i\omega}) \alpha_2(e^{i\omega}) f_{X_1 X_2}(\omega),$$

onde

$$(36) \quad \alpha_j(e^{-i\omega}) = \sum_{k=-\infty}^{+\infty} \alpha_{k,j} e^{-iak}, \quad j = 1,2,$$

é a função de transferência associada a cada filtro.

Verifica-se que a coerência é invariante para as transformações lineares invariantes no tempo.

Considere-se agora que as componentes de  $X_t$ ,  $X_{t1}$  e  $X_{t2}$ , se encontram relacionadas através de um filtro linear invariante no tempo,

$$(37) \quad X_{t2} = \sum_{k=-\infty}^{+\infty} \alpha_k X_{t-k,1},$$

onde  $\sum_k \alpha_k < \infty$ .

Neste caso, mostra-se que

$$(38) \quad f_{21}(\omega) = \sum_{k=-\infty}^{+\infty} \alpha_k e^{-i\omega k} f_{11}(\omega) = \alpha(e^{-i\omega}) f_{11}(\omega)$$

e que, assim sendo, usando o seguinte resultado para o caso univariado,

$$(39) \quad f_{22}(\omega) = |\alpha(e^{-i\omega})|^2 f_{11}(\omega),$$

vem,

$$(40) \quad K_{21}^2(\omega) = \frac{|f_{21}(\omega)|^2}{f_{11}(\omega)f_{22}(\omega)} = \frac{|\alpha(e^{-i\omega})|^2 [f_{11}(\omega)]^2}{f_{11}(\omega) |\alpha(e^{-i\omega})|^2 f_{11}(\omega)} = 1.$$

Ou seja, como seria de esperar, o valor da coerência entre  $X_{t1}$  e  $X_{t2}$  é 1.

### 3.3.O Caso Multivariado

Os resultados apresentados no parágrafo anterior são facilmente generalizados para o caso multivariado.

Seja  $X_t$ ,  $t = 0, \pm 1, \pm 2, \dots$ , uma série  $r$  dimensional, estacionária, com séries componentes dadas por  $X_{tj}$ ,  $t = 0, \pm 1, \pm 2, \dots$ ,  $j = 1, \dots, r$ .

Considere-se que  $E(X_t) = \mu$ ,  $E(X_{tj}) = \mu_j$ , e que  $\text{cov}[X_{t+k}, X_t] = \Gamma(k)$ ,  $\text{cov}(X_{t+k,i}, X_{tj}) = \gamma_{ij}(k)$ , onde  $t, k = 0, \pm 1, \pm 2, \dots$  e  $i, j = 1, \dots, r$ .

A representação espectral de  $X_t$  é dada por

$$(41) \quad X_t = \int_{-\pi}^{\pi} e^{i\omega t} dU(\omega),$$

onde  $d\mathbf{U}(\omega) = [dU_1(\omega), \dots, dU_r(\omega)]'$  e onde os processos  $dU_j(\omega)$ ,  $j = 1, \dots, r$ , são ortogonais e ortogonais cruzados.

A representação espectral da função matriz de covariâncias é dada por

$$(42) \quad \Gamma(k) = \int_{-\pi}^{\pi} e^{iak} d\mathbf{F}_{XX}(\omega),$$

verificando-se

$$(43) \quad d\mathbf{F}_{XX}(\omega) = E\{d\mathbf{U}(\omega)d\mathbf{U}^*(\omega)\},$$

onde  $\mathbf{F}_{XX}(\omega)$  é a função matriz de distribuição espectral de  $\mathbf{X}_t$  na frequência  $\omega$ . Os elementos diagonais desta matriz são as funções de distribuição espectral de cada elemento de  $\mathbf{X}_t$ . Os elementos da matriz que estão fora da diagonal principal correspondem às funções de distribuição espectral cruzada entre cada par de elementos de  $\mathbf{X}_t$ .

Quando cada elemento da matriz função de covariâncias for absolutamente somável, ou seja, quando  $\sum_{k=-\infty}^{\infty} |\gamma_{ij}(k)| < \infty$ ,  $i, j = 1, \dots, r$ , verifica-se que:

$$(44) \quad d\mathbf{F}_{XX}(\omega) = \mathbf{f}_{XX}(\omega)d\omega.$$

É possível definir  $\mathbf{f}_{XX}(\omega)$ , a matriz de densidade espectral na frequência  $\omega$  da série  $\mathbf{X}_t$ , como

$$(45) \quad \mathbf{f}_{XX}(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \Gamma(k)e^{-iak},$$

onde os elementos da diagonal principal da matriz são os já definidos espectros de cada série componente de  $\mathbf{X}_t$  e os elementos fora da diagonal principal são os igualmente já definidos espectros cruzados. Como função de  $\omega$ , a matriz  $\mathbf{f}_{XX}(\omega)$  é periódica com período  $2\pi$ , Hermitiana e definida não negativa. Devido a estas propriedades, o domínio básico de definição de  $\mathbf{f}_{XX}(\omega)$  é  $[0, \pi]$ .

Para cada par de processos componentes de  $\mathbf{X}_t$ , é possível definir as funções espectrais definidas para o caso bivariado: coespectro, quadratura, espectro fase, função ganho e coerência.

Considere-se agora que é aplicado um filtro linear invariante no tempo ao processo  $\mathbf{X}_t$ ,  $r$  dimensional, resultando num processo  $\mathbf{Y}_t$ ,  $s$  dimensional, ou seja,

$$(46) \quad Y_t = \sum_{k=-\infty}^{\infty} a_k X_{t-k} = \sum_{k=-\infty}^{\infty} a_{t-k} X_t, \quad t = 0, \pm 1, \pm 2, \dots,$$

onde  $a_t$  é um filtro linear invariante no tempo,  $(s \times r)$ , absolutamente somável, com função de transferência

$$(47) \quad A(\omega) = \sum_{k=-\infty}^{\infty} a_k e^{-iak}.$$

Neste contexto, a função matriz de densidade espectral de  $Y_t$  é dada por:

$$(48) \quad f_{YY}(\omega) = A(\omega) f_{XX}(\omega) A^*(\omega).$$

De seguida, por forma a prosseguir a interpretação das funções espectrais iniciada na secção 3.2.2., considera-se o efeito da aplicação do filtro  $(2r \times 1)$ , com a seguinte função de transferência:

$$(49) \quad A(\omega) = \begin{bmatrix} 1 \\ \vdots \\ 1 \\ -i \operatorname{sgn} \omega \\ \vdots \\ -i \operatorname{sgn} \omega \end{bmatrix}.$$

Da aplicação deste filtro, resulta o processo  $2r$  dimensional dado por:

$$(50) \quad \begin{bmatrix} X_t(\omega) \\ X_t^H(\omega) \end{bmatrix},$$

onde  $X_t(\omega)$  é a componente de frequência  $\omega$  de  $X_t$ ,  $X_t^H(\omega)$  é a componente de frequência  $\omega$  de  $X_t^H$  (matriz transformada de Hilbert<sup>1</sup> de  $X_t$ ), e cuja matriz de covariância é, aproximadamente, proporcional a

---

<sup>1</sup> Mostra-se que, sendo a representação espectral de  $X_t$  dada por

$$X_t = \int_{-\pi}^{\pi} e^{iat} dZ(\omega) = \int_{-\pi}^{\pi} [\cos(\omega t) dU(\omega) + \operatorname{sen}(\omega t) dV(\omega)]$$

onde  $U(\omega) = \operatorname{Re}\{Z(\omega)\}$  e  $V(\omega) = -\operatorname{Im}\{Z(\omega)\}$ , então

$$(51) \quad \begin{bmatrix} \text{Re}f_{XX}(\omega) & \text{Im}f_{XX}(\omega) \\ -\text{Im}f_{XX}(\omega) & \text{Re}f_{XX}(\omega) \end{bmatrix}.$$

Assim sendo, a parte real da matriz de densidade espectral na frequência  $\omega$  de  $\mathbf{X}_t$ ,  $\text{Re}\{f_{XX}(\omega)\}$  é proporcional à matriz de covariância da componente de frequência  $\omega$  de  $\mathbf{X}_t$ ,  $\mathbf{X}_t(\omega)$ , e que a parte imaginária,  $\text{Im}\{f_{XX}(\omega)\}$ , é proporcional à matriz de covariância cruzada da componente de frequência  $\omega$  de  $\mathbf{X}_t$ ,  $\mathbf{X}_t(\omega)$ , com a sua transformada de Hilbert,  $\mathbf{X}_t^H(\omega)$ . Deste modo, para qualquer par  $ij$  de processos componentes de  $\mathbf{X}_t$ , o coespectro de  $X_{it}$  com  $X_{jt}$  é proporcional à covariância da componente de frequência  $\omega$  da série  $X_{it}$  com a correspondente componente da série  $X_{jt}$ ; o espectro quadratura é proporcional à covariância da transformada de Hilbert da componente de frequência  $\omega$  da série  $X_{it}$  com a componente de frequência  $\omega$  da série  $X_{jt}$ . Assim sendo, o coespectro e o espectro quadratura são medidas de grau de associação linear.

### 3.4. Estimação do Espectro Cruzado

Considere-se novamente  $\mathbf{X}_t = [X_{t1}, X_{t2}]'$  uma série temporal bivariada estacionária.

Como foi visto no capítulo 2, para uma dada série temporal  $X_{jt}$ , com  $t = 1, \dots, n$ ,  $j = 1, 2$ , o seu espectro pode ser estimado pelo espectro alisado dado por

$$(52) \quad \hat{f}_{jj}(\omega) = \hat{f}_j(\omega) = \frac{1}{2\pi} \sum_{k=-M_j}^{M_j} W_j(k) \hat{\gamma}_j(k) e^{-i\omega k}, \quad j = 1, 2,$$

onde  $\hat{\gamma}_j(k)$  é a função de autocovariância amostral,  $W_j(k)$  é a janela retardada e  $M_j$  é o ponto de truncagem.

De modo semelhante, é possível estimar o espectro cruzado por:

---


$$\mathbf{X}_t^H = \int_{-\pi}^{\pi} [\text{sen}(\omega t) d\mathbf{U}(\omega) - \text{cos}(\omega t) d\mathbf{V}(\omega)]$$

e

$$\mathbf{X}_t^H(\omega) = 2[\text{sen}(\omega t) d\mathbf{U}(\omega) - \text{cos}(\omega t) d\mathbf{V}(\omega)],$$

verificando-se

$$\mathbf{X}_t(\omega) + i\mathbf{X}_t^H(\omega) = 2e^{i\omega t} d\mathbf{Z}(\omega)$$

$$(53) \quad \hat{f}_{12}(\omega) = \frac{1}{2\pi} \sum_{k=-M_{12}}^{M_{12}} W_{12}(k) \hat{\gamma}_{12}(k) e^{-i\omega k},$$

onde  $\hat{\gamma}_{12}(k)$  é a função de covariância cruzada amostral,  $W_{12}(k)$  é a janela retardada e  $M_{12}$  é o ponto de truncagem.

De forma a estimar-se o coespectro e o espectro quadratura, note-se que, escrevendo

$$(54) \quad \hat{f}_{12}(\omega) = \hat{c}_{12}(\omega) - i\hat{q}_{12}(\omega),$$

vem:

$$(55) \quad \begin{aligned} \hat{c}_{12}(\omega) &= \frac{1}{2\pi} \sum_{k=-M_{12}}^{M_{12}} W_{12}(k) \hat{\gamma}_{12}(k) \cos \omega k = \\ &= \frac{1}{2\pi} \left\{ W_{12}(0) \hat{\gamma}_{12}(0) + \sum_{k=1}^{M_{12}} W_{12}(k) [\hat{\gamma}_{12}(k) + \hat{\gamma}_{12}(-k)] \cos \omega k \right\} \end{aligned}$$

e

$$(56) \quad \begin{aligned} \hat{q}_{12}(\omega) &= \frac{1}{2\pi} \sum_{k=-M_{12}}^{M_{12}} W_{12}(k) \hat{\gamma}_{12}(k) \sin \omega k = \\ &= \frac{1}{2\pi} \left\{ \sum_{k=1}^{M_{12}} W_{12}(k) [\hat{\gamma}_{12}(k) - \hat{\gamma}_{12}(-k)] \sin \omega k \right\}. \end{aligned}$$

O ponto de truncagem e a janela retardada são escolhidos de forma idêntica ao caso univariado.

Para uma dada série temporal vectorial, os pontos de truncagem e respectivas janelas não têm de ser idênticas para a estimação de cada espectro e de cada espectro cruzado, no entanto, de um modo geral, dada a simplificação das propriedades amostrais das estimativas, é isso que acontece.

A estimação das restantes funções espectrais é feita da seguinte forma:

$$(57) \quad \hat{\alpha}_{12}(\omega) = [\hat{c}_{12}^2(\omega) + \hat{q}_{12}^2(\omega)]^{1/2},$$

$$(58) \quad \hat{\phi}_{12}(\omega) = \tan^{-1} [-\hat{q}_{12}(\omega) / \hat{c}_{12}(\omega)],$$

$$(59) \quad \hat{G}_{12}(\omega) = \hat{\alpha}_{12}(\omega) / \hat{f}_{11}(\omega)$$

e

$$(60) \quad \hat{K}_{12}^2(\omega) = \frac{|\hat{f}_{12}(\omega)|^2}{\hat{f}_{11}(\omega) \hat{f}_{22}(\omega)}.$$

Note-se que a janela retardada utilizada para o alisamento do estimador do espectro coloca maior ênfase nas covariâncias cruzadas com desfasamento 0. Assim sendo, se uma série for adiantada em relação à outra – o que resulta no facto da função de covariância cruzada não apresentar um pico no desfasamento 0 – o espectro cruzado estimado será enviesado. O enviesamento torna-se especialmente grave no caso da coerência. No entanto, como esta é invariável a transformações lineares, o enviesamento pode ser corrigido através de um correcto “alinhamento” das séries, de forma a que o pico da função de covariância cruzada ocorra para no desfasamento 0. É de referir que, quando a coerência é baixa, a estimação da função ganho, espectro amplitude e espectro fase não é feita com um nível de confiança aceitável.

## Capítulo 4: Análise em Componentes Principais Dinâmicas

### 4.1. Introdução

A técnica de Análise em Componentes Principais é um método que transforma variáveis correlacionadas em novas variáveis, que são combinações lineares das variáveis originais. Um dos seus principais objectivos é a redução do número de variáveis em estudo. Assim sendo, são retidas as combinações lineares das variáveis originais que permitem que a redução da dimensão dos dados origine a menor perda possível de informação sobre a variabilidade dos mesmos.

Neste capítulo, considera-se o caso da aproximação de uma série por uma versão filtrada de si mesma, sendo que esta versão filtrada é de dimensão inferior à série original. Imagine-se que se pretendem transmitir  $r$  sinais, correspondentes a uma concretização de uma série  $r$  dimensional  $X_t$ , existindo, no entanto, apenas  $q \leq r$  canais de transmissão disponíveis. Imagine-se então que, para transmitir o máximo de informação possível sobre  $X_t$ , se constrói a série dada por

$$(1) \quad Z_t = \sum_k b_k X_{t-k}$$

onde  $\{b_t\}_{(q \times r)}$  é um filtro linear. Assim sendo, transmite-se os valores da série  $q$  dimensional  $Z_t$ , sendo necessário, na recepção destes sinais, reconstruir, da forma mais aproximada possível, a série original  $X_t$ , formando a série

$$(2) \quad \tilde{X}_t = \mu + \sum_k c_k Z_{t-k},$$

para um vector  $\mu_{(r \times 1)}$  e uma matriz filtro  $\{c_t\}_{(r \times q)}$ . O problema que se coloca neste capítulo é a determinação do vector  $\mu_{(r \times 1)}$  e dos filtros  $\{b_t\}_{(q \times r)}$  e  $\{c_t\}_{(r \times q)}$  tais que  $\tilde{X}_t$  aproxime da melhor forma possível a série original  $X_t$ .

Começar-se-á por relembrar os conceitos da análise em Componentes Principais clássica, na secção 4.2.. As componentes principais obtidas por este processo são referidas como Componentes Principais Estáticas. Em seguida, na secção 4.3., é abordada a análise em Componentes Principais Dinâmicas. Por fim, na secção 4.4., enunciam-se algumas relações existentes entre as Componentes Principais Estáticas e as Componentes Principais Dinâmicas.

Neste capítulo, considere-se a série  $r$  dimensional  $X_t$ ,  $t = 0, \pm 1, \dots$ , estacionária, com média

$$(3) \quad E(X_t) = \mu_X$$

com função de autocovariância absolutamente somável:

$$(4) \quad E\left[(\mathbf{X}_{t+k} - \boldsymbol{\mu}_X)(\mathbf{X}_t - \boldsymbol{\mu}_X)'\right] = \Gamma_{XX}(k), \quad k = 0, \pm 1, \pm 2, \dots$$

e com matriz espectro

$$(5) \quad \mathbf{f}_{XX}(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \Gamma_{XX}(k) e^{-i\omega k}, \quad -\infty < \omega < +\infty.$$

## 4.2. Componentes Principais Estáticas

Seja  $\mathbf{X}_t$  o vector definido na secção anterior. Pretende-se determinar o vector  $r$  dimensional  $\boldsymbol{\mu}$ , a matriz  $(q \times r)$   $\mathbf{B}$  e a matriz  $(r \times q)$   $\mathbf{C}$  que minimizam simultaneamente todos os valores próprios de

$$(6) \quad E\left\{(\mathbf{X}_t - \boldsymbol{\mu} - \mathbf{CBX}_t)(\mathbf{X}_t - \boldsymbol{\mu} - \mathbf{CBX}_t)'\right\}^1.$$

É possível mostrar que estes são dados por:

$$(7) \quad \mathbf{B} = \begin{bmatrix} \mathbf{V}'_1 \\ \vdots \\ \mathbf{V}'_q \end{bmatrix},$$

$$(8) \quad \mathbf{C} = [\mathbf{V}_1 \quad \dots \quad \mathbf{V}_q] = \mathbf{B}'$$

e

$$(9) \quad \boldsymbol{\mu} = \boldsymbol{\mu}_X - \mathbf{CB}\boldsymbol{\mu}_X$$

onde  $\mathbf{V}_j$  é o  $j$ -ésimo vector próprio da matriz de covariâncias  $\Gamma_{XX}(0)$ ,  $j = 1, \dots, r$ , sendo  $\lambda_j$  o correspondente valor próprio. Verifica-se então:

$$(10) \quad \min_{\boldsymbol{\mu}, \mathbf{C}, \mathbf{B}} E\left\{(\mathbf{X}_t - \boldsymbol{\mu} - \mathbf{CBX}_t)(\mathbf{X}_t - \boldsymbol{\mu} - \mathbf{CBX}_t)'\right\} = \sum_{j>q} \lambda_j \mathbf{V}_j \mathbf{V}'_j.$$

A variável  $Z_{jt} = \mathbf{V}'_j \mathbf{X}_t$  é chamada a  $j$ -ésima componente principal de  $\mathbf{X}_t$ ,  $j = 1, \dots, r$ .

<sup>1</sup> Minimizar uma função de matrizes hermitianas  $\mathbf{A}(\boldsymbol{\theta})$ , em relação a  $\boldsymbol{\theta}$ , corresponde a determinar o valor  $\boldsymbol{\theta}_0$  tal que  $\mathbf{A}(\boldsymbol{\theta}) \geq \mathbf{A}(\boldsymbol{\theta}_0)$  para todo o  $\boldsymbol{\theta}$ . É possível mostrar que o mesmo valor  $\boldsymbol{\theta}_0$ , que minimiza  $\mathbf{A}(\boldsymbol{\theta})$ , minimiza também o seu determinante, o seu traço, os valores da sua diagonal principal e os seus valores próprios.

Verifica-se facilmente que,

$$(11) \quad \text{cov}(\mathbf{V}'_j \mathbf{X}_t, \mathbf{V}'_k \mathbf{X}_t) = \begin{cases} 0, & j \neq k \\ \lambda_j, & j = k \end{cases}$$

Assim sendo, as componentes principais de  $\mathbf{X}_t$  são combinações lineares não correlacionadas de  $\mathbf{X}_t$ . Também se pode caracterizar a  $j$ -ésima componente principal de  $\mathbf{X}_t$  como sendo a combinação linear  $Z_{jt} = \alpha'_j \mathbf{X}_t$ , onde  $\alpha'_j \alpha = 1$ , com variância máxima e que é não correlacionada com  $Z_{kt}$ ,  $k < j$ , ou seja, é a combinação linear  $Z_{jt} = \alpha'_j \mathbf{X}_t$  resultante da solução do seguinte problema de maximização:

$$(12) \quad \begin{aligned} & \max \text{var}(\alpha'_j \mathbf{X}_t) \\ & \text{s.a.: } \alpha'_j \alpha = 1 \\ & \text{cov}(\alpha'_j \mathbf{X}_t, \alpha'_k \mathbf{X}_t) = 0, \quad k < j \end{aligned}$$

Generalizando os resultados anteriores para o caso complexo, vem que o vector  $r$  dimensional  $\mu$ , a matriz  $(q \times r)$   $\mathbf{B}$  e a matriz  $(r \times q)$   $\mathbf{C}$  que minimizam simultaneamente todos os valores próprios de

$$(13) \quad E\{(\mathbf{X}_t - \mu - \mathbf{CBX}_t)(\mathbf{X}_t - \mu - \mathbf{CBX}_t)^*\},$$

são dados por:

$$(14) \quad \mathbf{B} = \begin{bmatrix} \mathbf{V}_1^* \\ \vdots \\ \mathbf{V}_q^* \end{bmatrix},$$

$$(15) \quad \mathbf{C} = [\mathbf{V}_1 \quad \dots \quad \mathbf{V}_q] = \mathbf{B}^*$$

e

$$(16) \quad \mu = \mu_X - \mathbf{CB}\mu_X$$

onde  $\mathbf{V}_j$  é o  $j$ -ésimo vector próprio da matriz de covariâncias  $\Gamma_{XX}(0)$ ,  $j = 1, \dots, r$ , sendo  $\lambda_j$  o correspondente valor próprio. Nestas condições, verifica-se que:

$$(17) \quad \min_{\mu, \mathbf{C}, \mathbf{B}} E\{(\mathbf{X}_t - \mu - \mathbf{CBX}_t)(\mathbf{X}_t - \mu - \mathbf{CBX}_t)^*\} = \sum_{j>q} \lambda_j \mathbf{V}_j \mathbf{V}_j^*.$$

A variável  $Z_{jt} = \mathbf{V}_j^* \mathbf{X}_t$  é chamada a  $j$ -ésima componente principal de  $\mathbf{X}_t$ ,  $j = 1, \dots, r$ .

Verifica-se que

$$(18) \quad \text{cov}(\mathbf{V}_j^* \mathbf{X}_t, \mathbf{V}_k^* \mathbf{X}_t) = \begin{cases} 0, & j \neq k \\ \lambda_j, & j = k \end{cases}.$$

Note-se que os resultados expostos, tanto para o caso real como complexo, conduzem a uma aproximação de  $\mathbf{X}_t$  dada por:

$$(19) \quad \boldsymbol{\mu}_X + \mathbf{A}(\mathbf{X}_t - \boldsymbol{\mu}_X),$$

onde  $\mathbf{A} = \mathbf{C}\mathbf{B}$ , sendo, no caso real, igual a

$$(20) \quad \mathbf{A} = \mathbf{V}_1 \mathbf{V}_1' + \mathbf{V}_2 \mathbf{V}_2' + \dots + \mathbf{V}_q \mathbf{V}_q'$$

e, no caso complexo, igual a

$$(21) \quad \mathbf{A} = \mathbf{V}_1 \mathbf{V}_1^* + \mathbf{V}_2 \mathbf{V}_2^* + \dots + \mathbf{V}_q \mathbf{V}_q^*.$$

O grau de aproximação conseguido depende directamente de quão próximos de zero estão os valores próprios  $\lambda_j, j > q$ .

### 4.3. Componentes Principais Dinâmicas

Considere-se o problema em que se pretende determinar o vector  $\boldsymbol{\mu}_{(r \times 1)}$ , o filtro  $\{\mathbf{b}_t\}_{(q \times r)}$  e o filtro  $\{\mathbf{c}_t\}_{(r \times q)}$  tais que, se

$$(22) \quad \mathbf{Z}_t = \sum_k \mathbf{b}_k \mathbf{X}_{t-k},$$

a diferença

$$(23) \quad \mathbf{X}_t - \boldsymbol{\mu} - \sum_k \mathbf{c}_k \mathbf{Z}_{t-k}$$

é minimizada.

A determinação dos filtros e do vector acima referidos, a partir de uma análise no domínio temporal, implicaria a determinação de um número infinito de coeficientes. Assim sendo, a análise em componentes principais é feita no domínio das frequências.

Seja  $\mathbf{X}_t, t = 0, \pm 1, \dots$  um vector  $r$  dimensional, estacionário de segunda ordem com valor médio  $\boldsymbol{\mu}_X$ , função de autocovariância absolutamente somável  $\Gamma_{XX}(k)$  e matriz de densidade espectral  $\mathbf{f}_{XX}(\omega)$ .

Assim sendo, os valores de  $\mu_{(r \times 1)}$ ,  $\{b_t\}_{(q \times r)}$  e  $\{c_t\}_{(r \times q)}$  que minimizam,

$$(24) \quad E \left\{ \left[ \mathbf{X}_t - \mu - \sum_k \mathbf{c}_k \mathbf{Z}_{t-k} \right] \left[ \mathbf{X}_t - \mu - \sum_k \mathbf{c}_k \mathbf{Z}_{t-k} \right]^* \right\}$$

são dados por:

$$(25) \quad \mu = \mu_X - \left( \sum_k \mathbf{c}_k \right) \left( \sum_k \mathbf{b}_k \right) \mu_X,$$

$$(26) \quad \mathbf{b}_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{B}(\omega) e^{iak} d\omega$$

e

$$(27) \quad \mathbf{c}_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{C}(\omega) e^{iak} d\omega,$$

onde

$$(28) \quad \mathbf{B}(\omega) = \begin{bmatrix} \mathbf{V}_1(\omega)^* \\ \vdots \\ \mathbf{V}_q(\omega)^* \end{bmatrix}$$

e

$$(29) \quad \mathbf{C}(\omega) = \mathbf{B}(\omega)^* = [\mathbf{V}_1(\omega) \quad \dots \quad \mathbf{V}_q(\omega)].$$

O vector  $\mathbf{V}_j(\omega)$  representa o  $j$ -ésimo vector próprio de  $\mathbf{f}_{XX}(\omega)$ ,  $j = 1, \dots, r$ , sendo  $\lambda_j$  o correspondente valor próprio. Assim sendo, o mínimo obtido é dado por:

$$(30) \quad \int_{-\pi}^{\pi} \left\{ \sum_{j>q} \lambda_j(\omega) \right\} d\omega.$$

Assim, a série  $\mathbf{X}_t$ ,  $t = 0, \pm 1, \dots$ , tem a seguinte representação espectral:

$$(31) \quad \mathbf{X}_t = \int_{-\pi}^{\pi} e^{iat} d\mathbf{U}_X(\omega),$$

e a série  $\mathbf{Z}_t$ , correspondente aos valores que minimizam a expressão,

$$(32) \quad E \left\{ \left[ \mathbf{X}_t - \mu - \sum_k \mathbf{c}_k \mathbf{Z}_{t-k} \right] \left[ \mathbf{X}_t - \mu - \sum_k \mathbf{c}_k \mathbf{Z}_{t-k} \right]^* \right\}$$

toma a forma,

$$(33) \quad \mathbf{Z}_t = \int_{-\pi}^{\pi} \mathbf{B}(\omega) e^{iat} d\mathbf{U}_X(\omega).$$

A sua  $j$ -ésima componente,

$$(34) \quad Z_{j,t} = \int_{-\pi}^{\pi} \mathbf{V}_j(\omega)^* e^{i\omega t} d\mathbf{U}_X(\omega)$$

é a  $j$ -ésima componente principal dinâmica de  $\mathbf{X}_t$ .

A matriz de densidade espectral de  $\mathbf{Z}_t$  é dada por

$$(35) \quad \begin{bmatrix} \lambda_1(\omega) & & 0 \\ & \ddots & \\ 0 & & \lambda_q(\omega) \end{bmatrix},$$

ou seja, o espectro de  $Z_{j,t}$  é dado por  $\lambda_j(\omega)$  e, cada par  $Z_{j,t}$ ,  $Z_{k,t}$ ,  $j \neq k$ , tem coerência igual a 0 em todas as frequências.

Seja  $\tilde{\mathbf{X}}_t = \boldsymbol{\mu} + \sum_k \mathbf{c}_k \mathbf{Z}_{t-k}$  a aproximação de  $\mathbf{X}_t$  construída a partir dos vectores e matrizes indicados anteriormente; seja  $\varepsilon_t$  a série do erro dada por  $\mathbf{X}_t - \tilde{\mathbf{X}}_t$ ; seja  $\mathbf{A}(\omega) = \mathbf{C}(\omega)\mathbf{B}(\omega)$ , a matriz com característica menor ou igual a  $q$ . A representação espectral de  $\varepsilon_t$  é dada por:

$$(36) \quad \int_{-\pi}^{\pi} [\mathbf{I} - \mathbf{A}(\omega)] e^{i\omega t} d\mathbf{U}_X(\omega) = \int_{-\pi}^{\pi} \left\{ \sum_{j>q} \mathbf{V}_j(\omega) \mathbf{V}_j(\omega)^* \right\} e^{i\omega t} d\mathbf{U}_X(\omega).$$

A série  $\varepsilon_t$  tem valor médio 0 e matriz de densidade espectral dada por:

$$(37) \quad \mathbf{f}_{\varepsilon\varepsilon}(\omega) = \sum_{j>q} \lambda_j(\omega) \mathbf{V}_j(\omega) \mathbf{V}_j(\omega)^*.$$

Assim sendo, o grau de aproximação de  $\mathbf{X}_t$  por  $\tilde{\mathbf{X}}_t$  é directamente relacionada com quão próximos estão os valores próprios  $\lambda_j(\omega)$ ,  $j > q$ , de zero. Note-se também que as matrizes espectrais cruzadas entre  $\varepsilon_t$  e  $\mathbf{Z}_t$  e entre  $\varepsilon_t$  e  $\tilde{\mathbf{X}}_t$  são identicamente iguais a 0.

Note-se que as componentes principais dinâmicas podem ser definidas como sendo o resultado de problemas de maximização, tal como as componentes principais estáticas. Assim, a  $j$ -ésima componente principal dinâmica é a série real  $Z_{j,t}$  dada por,

$$(38) \quad \int_{-\pi}^{\pi} \mathbf{B}_j(\omega) e^{i\omega t} d\mathbf{U}_X(\omega)$$

que tem variância máxima, sujeita à condição  $\mathbf{B}_j(\omega)\mathbf{B}_j(\omega)^* = 1$  ( $\mathbf{B}_j(\omega)_{(1 \times r)}$ ) e à condição de que tenha coerência zero com as séries  $Z_{k,t}$ ,  $k < j$ ,  $j = 1, \dots, r$ . A variância máxima atingida por  $Z_{j,t}$  é:

$$(39) \quad \int_{-\pi}^{\pi} \lambda_j(\omega) d\omega.$$

Denominando este valor por  $\lambda_j$ , a variância explicada pelas primeiras  $q$  componentes principais dinâmicas, passando do domínio frequência para o domínio tempo, é dada por:

$$(40) \quad \frac{\lambda_1 + \dots + \lambda_q}{\lambda_1 + \dots + \lambda_n}$$

Refira-se por último, que os valores e vectores próprios de uma matriz de densidade espectral não são invariantes sob a aplicação de um filtro à série em estudo. Assim sendo, séries com maior variabilidade têm maior peso nas componentes principais obtidas, surgindo problemas quando as séries estão medidas em escalas não comparáveis. Uma forma de ultrapassar estas dificuldades consiste em realizar os cálculos sobre a matriz de coerência e na sobre a matriz de densidade espectral.

#### 4.4. Relação entre a Análise em Componentes Principais Dinâmicas e em Componentes Principais Estáticas

A análise em componentes principais estáticas é realizada a partir dos valores e vectores próprios da matriz de covariâncias  $\Gamma(0)$ , ou seja, sem considerar todas as possíveis relações entre as diferentes componentes desfasadas de  $\mathbf{X}_t$ . A análise em componentes principais dinâmicas permite explorar as relações de precedência ou antecedência existentes entre as componentes de  $\mathbf{X}_t$ . Ora, considerar todos os avanços e atrasos no tempo implica a estimação de um número infinito de parâmetros. Este problema é ultrapassado realizando a análise em componentes principais no domínio frequência, através dos valores e vectores próprios da matriz de densidade espectral, e convertendo posteriormente os parâmetros definidos no domínio frequência para o domínio tempo. Note-se, que o estudo realizado para todas as frequências é idêntico à realização de um estudo para todos os atrasos e avanços no tempo. Note-se também que o caso estático assume implicitamente que  $\Gamma(k) = 0, \forall k > 0$ , pelo que só faz sentido realizar uma análise em componentes principais dinâmica quando existirem componentes de  $\mathbf{X}_t$  correlacionadas em diferentes desfasamentos no tempo.

Verifica-se que a análise em componentes principais de uma série estacionária  $X_t$ , realizada no domínio frequência, é idêntica à análise em componentes principais clássica ou estática realizada sobre as componentes de  $X_t$  em cada frequência e as suas transformadas de Hilbert. Para tal, considere-se que  $X_t(\omega)$  representa a componente de frequência  $\omega$  de  $X_t$  e que  $f_{XX}(\omega)$  é a matriz de densidade espectral de  $X_t$ . Assim sendo, o vector de dimensão  $2r$ , dado por:

$$(41) \quad \begin{bmatrix} X_t(\omega) \\ X_t(\omega)^H \end{bmatrix}$$

tem uma matriz de covariâncias que é dada por:

$$(42) \quad \begin{bmatrix} \text{Re} f_{XX}(\omega) & \text{Im} f_{XX}(\omega) \\ -\text{Im} f_{XX}(\omega) & \text{Re} f_{XX}(\omega) \end{bmatrix}.$$

A análise em componentes principais estáticas conduziria à determinação dos valores e vectores próprios desta matriz de covariâncias:

$$(43) \quad \lambda_j(\omega), \begin{bmatrix} \text{Re} V_j(\omega) \\ \text{Im} V_j(\omega) \end{bmatrix} \quad \lambda_j(\omega), \begin{bmatrix} -\text{Im} V_j(\omega) \\ \text{Re} V_j(\omega) \end{bmatrix}$$

$j = 1, \dots, r$ , onde  $\lambda_j(\omega)$  e  $V_j(\omega)$  são os valores e vectores próprios da matriz de densidade espectral de  $X_t$ , tal como determinados no caso da análise em componentes principais dinâmicas.

## Capítulo 5: Análise Factorial Dinâmica Generalizada

### 5.1. Introdução

O modelo que se segue, o modelo factorial dinâmico generalizado, delineado nos artigos Forni, Hallin, Lippi e Reichlin (2000), Forni e Lippi (2001) e Forni, Hallin, Lippi e Reichlin (2002), surge numa tentativa de extrair a máxima informação possível de grandes painéis de dados.

Muitas vezes, em particular no caso de variáveis macroeconómicas, as séries de um painel de dados apresentam fortes co-movimentos. Nestes casos, a modelação univariada de tais séries resultaria num grande desperdício da informação disponível. Se o número de séries temporais em estudo,  $n$ , for pequeno, os modelos VAR ou VARMA são adequados para o seu estudo. No entanto, quando  $n$  é grande (e nomeadamente quando é maior que o número de observações por série,  $T$ ) estes modelos não são indicados, pois implicam a estimação de demasiados parâmetros. Os modelos factoriais surgem como uma alternativa, onde cada variável é representada como a soma de uma componente comum – dependente de um pequeno número de factores não observáveis comuns a todas as variáveis – e de uma componente idiossincrática (ou específica), que são mutuamente ortogonais. Vários modelos foram propostos em Geweke (1977), Sargent e Sims (1977), Quah e Sargent (1993), Stock e Watson (1989), Chamberlain e Rothschild (1983). A metodologia proposta por Forni, Hallin, Lippi e Reichlin difere das anteriores, uma vez que estes analisam um modelo que é dinâmico e permite, simultaneamente, a existência de correlação cruzada entre as componentes idiossincráticas. Uma característica importante destes modelos é permitirem que a componente comum tenha uma representação em médias móveis infinitas, de forma a possibilitar que as respostas aos factores comuns apresentem tanto a forma de médias móveis como a forma autorregressiva.

### 5.2. Modelos Factoriais para a Análise de Painéis de Dados de Séries Temporais

#### 5.2.1. Notação e Definições

Nas secções seguintes, será adoptada a notação de seguida indicada, bem como as hipóteses iniciais formuladas.

Seja  $P = (\Omega, \mathcal{F}, P)$  um espaço de probabilidade e seja  $L_2(P, \mathbb{C})$  o espaço linear de todas as variáveis aleatórias de valores complexos, com média nula, de quadrado integrável, definidas em  $\Omega$ .  $L_2(P, \mathbb{C})$  é um espaço de Hilbert, quando munido do produto interno dado por  $\langle x, y \rangle = E(x\bar{y}) = \text{cov}(x, y)$  e da norma  $\|x\| = \sqrt{\text{var}(x)}$ .

Sendo  $Q$  um subconjunto de  $L_2(P, \mathbb{C})$ ,  $\overline{\text{span}}(Q)$  representa o subespaço linear gerado por  $Q$ . Se  $V$  for um subespaço linear fechado de  $L_2(P, \mathbb{C})$  e  $x \in L_2(P, \mathbb{C})$ , a projecção ortogonal de  $x$  em  $V$  é representada por  $\text{proj}(x | V)$ .

Considere-se a sequência  $\mathbf{x} = \{x_{it}, i \in \mathbb{N}, t \in \mathbb{Z}\}$ , onde  $x_{it} \in L_2(\mathbb{P}, \mathbb{C})$ , pelo que os momentos de primeira e segunda ordem são finitos. Defina-se então,

- $\mathbf{X} = \overline{\text{span}}(\mathbf{x})$ ;
- $\mathbf{x}_t$  é o vector coluna de dimensão infinita dado por  $(x_{1t}, x_{2t}, \dots, x_{it}, \dots)'$ ;
- $\mathbf{x}_{nt}$  é o vector coluna  $n$ -dimensional  $(x_{1t}, x_{2t}, \dots, x_{nt})'$ ;
- $\mathbf{X}_n = \overline{\text{span}}(\{x_{st}, s = 1, 2, \dots, n, t \in \mathbb{Z}\})$ .

Assuma-se que, para todo o  $n \in \mathbb{N}$ , o processo  $\mathbf{x}_{nt}$  é estacionário e possui densidade espectral. A sua matriz de densidade espectral é representada por  $\Sigma_n(\theta)$  ou por  $\Sigma_n^x(\theta)$ ,  $\theta \in [-\pi, \pi]$ . Os seus elementos são representados por  $\sigma_{ij}(\theta)$ . Recorde-se que

$E(\mathbf{x}_{nt} \mathbf{x}_{n,t-k}^*) = \Gamma_{nk} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\theta} \Sigma_n(\theta) d\theta$ . Seja  $\Sigma^x$  a matriz de dimensão infinita cuja submatriz superior esquerda, de dimensão  $n \times n$ , é  $\Sigma_n^x$ .

Se  $\mathbf{a}$  for um vector linha infinito  $(a_1, a_2, \dots, a_n, a_{n+1}, \dots)$ ,  $\mathbf{a}^{[n]}$  representa o vector linha infinito  $(a_1, \dots, a_n, 0, 0, \dots)$  e  $\mathbf{a}^{(n)}$  representa o vector linha  $n$ -dimensional  $(a_1, \dots, a_n)$ .

$L_2^\infty(\Theta, \mathbb{C}, \Sigma^x)$ ,  $\Theta = [-\pi, \pi]$ , representa o espaço linear complexo de todos os vectores linha infinitos  $\mathbf{f} = (f_1, \dots, f_n, \dots)$ , tais que:  $f_i$  é uma função complexa mensurável definida em  $\Theta$ ;  $\int_{-\pi}^{\pi} \mathbf{f}(\theta) \Sigma^x(\theta) \mathbf{f}^*(\theta) d\theta = \lim_n \int_{-\pi}^{\pi} \mathbf{f}^{(n)}(\theta) \Sigma_n^x(\theta) \mathbf{f}^{(n)*}(\theta) d\theta < \infty$ . Os vectores  $\mathbf{f}$  e  $\mathbf{g}$  são considerados equivalentes se  $\int_{-\pi}^{\pi} (\mathbf{f}(\theta) - \mathbf{g}(\theta)) \Sigma^x(\theta) (\mathbf{f}^*(\theta) - \mathbf{g}^*(\theta)) d\theta = 0$ .  $L_2^\infty(\Theta, \mathbb{C}, \Sigma^x)$  é um espaço de Hilbert quando munido do produto interno  $\langle \mathbf{f}, \mathbf{g} \rangle_{\Sigma^x} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{f}(\theta) \Sigma^x(\theta) \mathbf{g}^*(\theta) d\theta$  e da norma  $\|\mathbf{f}\|_{\Sigma^x} = \sqrt{\langle \mathbf{f}, \mathbf{f} \rangle_{\Sigma^x}}$ .

O espaço  $L_2^\infty(\Theta, \mathbb{C})$  é definido tal como  $L_2^\infty(\Theta, \mathbb{C}, \Sigma^x)$ , substituindo  $\Sigma^x$  pela matriz de identidade infinita. O produto interno e a norma em  $L_2^\infty(\Theta, \mathbb{C})$  são representados, respectivamente, por  $\langle \mathbf{f}, \mathbf{g} \rangle$  e por  $\|\mathbf{f}\|$ .

Os espaços  $L_2^n(\Theta, \mathbb{C}, \Sigma_n^x)$  e  $L_2^n(\Theta, \mathbb{C})$  são definidos de forma equivalente, substituindo os vectores de dimensão infinita por vectores  $n$ -dimensionais.

Seja  $\lambda_{nj}^x$  a função que associa a cada  $\theta \in [-\pi, \pi]$  o  $j$ -ésimo valor próprio (em ordem descendente), real, não negativo de  $\Sigma_n^x(\theta)$ . Estas funções são os valores próprios dinâmicos de  $\Sigma_n^x$ . Defina-se a função  $\lambda_i^x$ , para cada  $i$ , como sendo  $\lambda_i^x(\theta) = \sup_n \lambda_{ni}^x(\theta)$ .

Recorde-se que, para uma dada matriz de densidade espectral,  $\Sigma_n^x(\theta)$ , existem  $n$  vectores de funções de valores complexos,

$$(1) \quad \mathbf{p}_{nj}^x(\theta) = (p_{nj,1}^x(\theta) \ p_{nj,2}^x(\theta) \ \cdots \ p_{nj,n}^x(\theta)), \quad j = 1, \dots, n,$$

tais que,

- (i)  $\mathbf{p}_{nj}^x(\theta)$  é um vector próprio linha de  $\Sigma_n^x(\theta)$  correspondente a  $\lambda_{nj}^x(\theta)$ ; isto é,  $\mathbf{p}_{nj}^x(\theta)\Sigma_n^x(\theta) = \lambda_{nj}^x(\theta)\mathbf{p}_{nj}^x(\theta)$ ,  $\forall \theta \in [-\pi, \pi]$ ;
- (2) (ii)  $|\mathbf{p}_{nj}^x(\theta)|^2 = 1$ ,  $\forall j$ ,  $\forall \theta \in [-\pi, \pi]$ ;
- (iii)  $\mathbf{p}_{nj}^x(\theta)\mathbf{p}_{ns}^x(\theta) = 0$ ,  $\forall j \neq s$ ,  $\forall \theta \in [-\pi, \pi]$ ;
- (iv)  $\mathbf{p}_{nj}^x(\theta)$  é  $\theta$  mensurável em  $[-\pi, \pi]$ .

Os vectores (1) assim definidos são os vectores próprios dinâmicos de  $\Sigma_n^x$ .

Os vectores próprios dinâmicos admitem um desenvolvimento em séries de Fourier, dado por:

$$(3) \quad \mathbf{p}_{nj}^x(\theta) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \left[ \int_{-\pi}^{\pi} \mathbf{p}_{nj}^x(\theta) e^{ik\theta} d\theta \right] e^{-ik\theta}.$$

Defina-se o seguinte filtro, de quadrado somável:

$$(4) \quad \underline{\mathbf{p}}_{nj}^x(L) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \left[ \int_{-\pi}^{\pi} \mathbf{p}_{nj}^x(\theta) e^{ik\theta} d\theta \right] L^k.$$

A  $j$ -ésima componente principal dinâmica de  $\mathbf{x}_{nt}$ , cuja densidade espectral é

$$(5) \quad \mathbf{p}_{nj}^x(\theta)\Sigma_n^x(\theta)\mathbf{p}_{nj}^{x*}(\theta) = \lambda_{nj}^x(\theta),$$

é dada pelo processo  $\{\underline{\mathbf{p}}_{nj}^x(L)\mathbf{x}_{nt}, t \in \mathbb{Z}\}$ .

Por fim, defina-se  $\mathbf{x}_n^T = (x_{it})_{i=1, \dots, n; t=1, \dots, T}$ , uma matriz rectangular,  $n \times T$ , que se assume ser uma realização finita do processo estocástico de valores reais  $\mathbf{x} = \{x_{it}, i \in \mathbb{N}, t \in \mathbb{Z}\}$ .

Ou seja, onde os processos vectoriais  $n$  dimensionais  $\{\mathbf{x}_{nt} = (x_{1t} \cdots x_{nt})', t \in \mathbb{Z}\}$ ,  $n \in \mathbb{N}$ , são estacionários, com valor médio  $\mathbf{0}_n$  e momentos de segunda ordem finitos.

## 5.2.2. Breve Descrição de Alguns Modelos Factoriais

### 5.2.2.1. O Modelo Factorial Estático

Os modelos factoriais tradicionais decompõem um número finito  $n$  de processos  $\{x_{it}\}$  em duas componentes não observáveis: uma componente comum  $\{\mathcal{X}_{it}\}$  explicada por  $q$  factores comuns  $\{F_{jt}\}$ ,  $j=1, \dots, q$ , onde  $q$  é pequeno quando comparado com  $n$ , e uma componente idiossincrática ou específica,  $\{\xi_{it}\}$ . Quando os factores têm impacto apenas contemporaneamente, o modelo delineado é o modelo factorial estático:

$$(6) \quad x_{it} = \mathcal{X}_{it} + \xi_{it} = a_{i1}F_{1t} + a_{i2}F_{2t} + \dots + a_{iq}F_{qt} + \xi_{it} = \mathbf{a}_i \mathbf{F}_t + \xi_{it},$$

onde  $\mathbf{a}_i = (a_{i1}, \dots, a_{iq})$  e  $\mathbf{F}_t = (F_{1t}, \dots, F_{qt})'$ . Para que o modelo seja identificável e estimável, são assumidas as seguintes hipóteses:

- Os processos  $\{F_{jt}, j=1, \dots, q, t \in Z\}$  e  $\{\xi_{it}, i=1, \dots, n, t \in Z\}$  são mutuamente ortogonais, donde resulta que também o são os processos  $\{\mathcal{X}_{it}, i=1, \dots, n, t \in Z\}$  e  $\{\xi_{it}, i=1, \dots, n, t \in Z\}$ ;
- Os processos idiossincráticos  $\{\xi_{i't}, t \in Z\}$  e  $\{\xi_{i''t}, t \in Z\}$  são mutuamente ortogonais para todo  $i' \neq i''$ .

### 5.2.2.2. O Modelo Factorial Dinâmico

Sargent e Sims (1977) e Geweke (1977) estenderam o modelo factorial estático ao caso em que os factores têm um impacto dinâmico e não apenas contemporâneo. Dessa forma, definiram o modelo factorial dinâmico,

$$(7) \quad x_{it} = \mathcal{X}_{it} + \xi_{it} = b_{i1}(L)u_{1t} + b_{i2}(L)u_{2t} + \dots + b_{iq}(L)u_{qt} + \xi_{it} = \mathbf{b}_i(L)\mathbf{u}_t + \xi_{it}$$

onde  $\{\mathbf{u}_t = (u_{1t}, \dots, u_{qt})', t \in Z\}$  é o vector de choques comuns, um processo de ruído branco ortonormal de dimensão  $q$  (com  $q$  muito pequeno em relação a  $n$ ), não observável, e  $b_{ij}(L) = \sum_{k=0}^{\infty} b_{ij,k}L^k$ ,  $i=1, \dots, n$ ,  $j=1, \dots, q$ , são  $nq$  filtros de quadrado somável. As condições a impor ao modelo são:

- Os processos  $\{\mathbf{u}_t, t \in Z\}$  e  $\{\xi_{it}, i=1, \dots, n, t \in Z\}$  são mutuamente ortogonais, pelo que também o são os processos  $\{\mathcal{X}_{it}, i=1, \dots, n, t \in Z\}$  e  $\{\xi_{it}, i=1, \dots, n, t \in Z\}$ ;
- Os processos idiossincráticos  $\{\xi_{i't}, t \in Z\}$  e  $\{\xi_{i''t}, t \in Z\}$  são mutuamente ortogonais para todo  $i' \neq i''$ .

### 5.2.3. O Modelo Factorial Dinâmico Generalizado

Em algumas aplicações, a condição que impõe a ortogonalidade dos processos idiossincráticos é demasiado restritiva. No entanto, não se admitindo tal hipótese, para um  $n$  finito, os modelos factoriais estático e dinâmico não são identificáveis.

O modelo proposto por Chamberlain e Rothschild (1983) generaliza o modelo factorial estático de forma a admitir uma definição mais geral da componente idiossincrática, para o caso em que existe um número infinito de séries.

Os modelos propostos por Forni et al. (2000) e Forni et al. (2002), generalizam o modelo factorial dinâmico de forma a que a identificação do modelo não dependa da hipótese de ortogonalidade dos processos idiossincráticos. A componente comum é estimada de forma consistente e não paramétrica quando  $n$  e  $T$  tendem para infinito.

#### 5.2.3.1. Definição do Modelo e Fundamentos Teóricos

Forni e Lippi (1999) analisam o modelo factorial dinâmico generalizado de um ponto de vista puramente teórico. Assim, os autores determinam as condições que a população deve verificar para que exista uma estrutura factorial dinâmica nos dados.

Como foi referido, não se requer a ortogonalidade das componentes idiossincráticas para que o modelo seja identificável. No entanto, é necessário que a variância das médias ponderadas das sequências idiossincráticas tenda para zero, quando a soma do quadrado dos ponderadores tende para zero. Como tal, antes de se apresentar o modelo propriamente dito, convém especificar a noção de componente idiossincrática adoptada pelos referidos autores, sendo, para tal, necessário expor outros conceitos relacionados.

Seja  $\mathbf{f} \in L_2^\infty(\Theta, \mathbb{C}, \Sigma^x)$ . Defina-se  $\underline{\mathbf{f}}(L)\mathbf{x}_t$ , da seguinte forma:

$$(8) \quad \underline{\mathbf{f}}(L)\mathbf{x}_t = \Omega^{-1}(\mathbf{f}e^{it}),$$

onde

$$(9) \quad \Omega(x_{ht}) = e^{it}(\delta_{h1} \quad \delta_{h2} \quad \dots \quad \delta_{hk} \quad \dots)$$

e  $\delta_{hk} = 1$  se  $h = k$  e  $\delta_{hk} = 0$  se  $h \neq k$ .

A densidade espectral do processo  $\{\underline{\mathbf{f}}(L)\mathbf{x}_t, t \in \mathbb{Z}\}$  é dada por  $\mathbf{f}\Sigma^x\mathbf{f}^*$ . Quando a sua expansão em série de Fourier,  $\mathbf{f}(\theta) = \sum_{k=-\infty}^{\infty} \mathbf{F}_k^f e^{-ik\theta}$ , onde  $\mathbf{F}_k^f = (1/2\pi) \int_{-\pi}^{\pi} \mathbf{f}(\theta) e^{ik\theta} d\theta$ , converge em  $L_2^\infty(\Theta, \mathbb{C}, \Sigma^x)$ , é possível definir  $\underline{\mathbf{f}}(L)$  como sendo o filtro linear  $\sum_{k=-\infty}^{\infty} \mathbf{F}_k^f L^k$ . Nesta situação temos que:

$$(10) \quad \underline{f}(L)x_t = \lim_s \sum_{k=-s}^s F_k^f x_{t-k}$$

tem variância finita e por isso faz sentido em  $X$ . No entanto, em geral,  $\underline{f}(L)x_t$  não pode ser representado como a soma de uma série, conforme apresentado em (10), e deve ser definido como em (9).

Seja  $a_n \in L_2^\infty(\Theta, \mathbb{C}) \cap L_2^\infty(\Theta, \mathbb{C}, \Sigma^x)$ ,  $\forall n \in \mathbb{N}$ . Diz-se que a sequência  $\{a_n, n \in \mathbb{N}\}$  é uma sequência ponderadora dinâmica (DAS - *dynamic averaging sequence*) se  $\lim_n \|a_n\| = 0$ .

Seja  $y_t \in X$ . Diz-se que  $y_t$  é um agregado se existir uma sequência ponderadora dinâmica  $\{a_n, n \in \mathbb{N}\}$  tal que  $\lim_n \underline{a}_n(L)x_t = y_t$ . O conjunto de todos os agregados será representado por  $\mathcal{G}(x)$ .

Diz-se que  $x$  é idiossincrático se  $\lim_n \underline{a}_n(L)x_t = 0$ , para qualquer sequência ponderadora dinâmica  $\{a_n, n \in \mathbb{N}\}$ . Nesse caso temos que  $\mathcal{G}(x) = \{0\}$ . É possível mostrar que afirmar que  $x$  é idiossincrático é equivalente a afirmar que  $\lambda_1^x$  é essencialmente limitada<sup>1</sup>. Mostra-se também que se  $x$  for idiossincrático, então,

$$(11) \quad \sup_n \int_{-\pi}^{\pi} \lambda_{n1}^x(\theta) d\theta = \lim_n \int_{-\pi}^{\pi} \lambda_{n1}^x(\theta) d\theta < \infty.$$

Após definição destes conceitos base, é possível apresentar o modelo factorial dinâmico generalizado. Seja  $q$  um inteiro não negativo. A sequência  $x$  diz-se uma sequência de  $q$  factores dinâmicos, se  $L_2(\mathbb{P}, \mathbb{C})$  contiver um processo vectorial de ruído branco  $q$ -dimensional ortonormal,  $u = \{u_{1t}, u_{2t}, \dots, u_{qt}\}'$ ,  $t \in Z$ , e a sequência  $\xi = \{\xi_{it}, i \in \mathbb{N}, t \in Z\}$ , que satisfazem a hipótese de estacionaridade e de existência de densidade espectral, tais que:

$$(i) \quad \forall i \in \mathbb{N},$$

$$(12) \quad x_{it} = \chi_{it} + \xi_{it}$$

$$(13) \quad \chi_{it} = \underline{b}_{i1}(L)u_{1t} + \underline{b}_{i2}(L)u_{2t} + \dots + \underline{b}_{iq}(L)u_{qt} = \underline{b}_i(L)u_t,$$

onde  $\underline{b}_i \in L_2^q([- \pi, \pi], \mathbb{C})$ .

(ii) Os processos  $\{u_t, t \in Z\}$  e  $\{\xi_{it}, i = 1, \dots, n, t \in Z\}$  são mutuamente ortogonais, donde resulta que também o são os processos  $\{\chi_{it}, i = 1, \dots, n, t \in Z\}$  e

<sup>1</sup> Diz-se que  $f$  (mensurável) é essencialmente limitada se existir  $c > 0$  tal que o conjunto  $\{x : |f(x)| > c\}$  tem medida nula; e  $c$  será então um *majorante essencial* de  $f$ .

$\{\xi_{it}, i=1, \dots, n, t \in Z\}$ , ou seja,  $\forall i \in N, j=1, 2, \dots, q$  e  $k \in Z$ , verifica-se  $\xi_{it} \perp u_{jt-k}$ , donde vem, consequentemente, que  $\xi_{it} \perp \chi_{st-k}, \forall i \in N, s \in N$  e  $k \in Z$ .

- (iii) O primeiro valor próprio dinâmico idiossincrático  $\lambda_{n1}^\xi$  é uniformemente limitado. Isto é, existe um real  $\Lambda$  tal que  $\lambda_{n1}^\xi(\theta) \leq \Lambda$  para qualquer  $\theta \in [-\pi, \pi]$  e qualquer  $n \in N$ . De outra forma,  $\lambda_1^\xi$  é essencialmente limitado. Isto é,  $\xi$  é idiossincrático.
- (iv) Definindo  $\chi = \{\chi_{it}, i \in N, t \in Z\}$ , os  $q$  primeiros valores próprios dinâmicos comuns divergem quase por toda a parte (q.p.t.p.)<sup>2</sup> em  $[-\pi, \pi]$ . Isto é,  $\lambda_q^x(\theta) = \infty$ , q.p.t.p. em  $[-\pi, \pi]$  ou, de outra forma,  $\lim_{n \rightarrow \infty} \lambda_{nj}^x(\theta) = \infty$  para  $j \leq q$ , q.p.t.p. em  $[-\pi, \pi]$ .

Tal como anteriormente,  $\chi$  e  $\xi$  são, respectivamente, as componentes comum e idiossincrática do modelo.

A condição (iii) da definição do modelo factorial dinâmico generalizado garante que a componente idiossincrática não sobrevive à aplicação de uma sequência ponderadora dinâmica, mas permite, por outro lado, a existência de alguma correlação dinâmica cruzada entre as componentes idiossincráticas. Pelo contrário, a condição (iv) garante a sobrevivência da componente comum, ou seja, garante a existência de uma correlação cruzada mínima entre as componentes comuns, estando cada choque comum presente em infinitas séries, com importância não decrescente.

No artigo referido, mostra-se que  $x$  é uma sequência de  $q$  factores dinâmicos se e só se

- (i)  $\lambda_{q+1}^x$  é essencialmente limitada;
- (ii)  $\lambda_q^x(\theta) = \infty$ , q.p.t.p. em  $[-\pi, \pi]$ .

Mostra-se também que, se  $x$  for uma sequência de  $q$  factores dinâmicos, então,

$$(14) \quad \overline{\text{span}}(\chi) = \overline{\text{span}}(\mathbf{u}) = \mathcal{G}(x),$$

e, além disso,

$$(15) \quad \chi_{it} = \text{proj}(x_{it} | \mathcal{G}(x)).$$

A demonstração destes resultados baseia-se na construção de um ruído branco ortonormal  $q$ -dimensional, denominado por  $\psi_t$ , cujos elementos são combinações lineares das componentes principais dinâmicas de ordem  $m$ ,  $\underline{p}_{mj}^x(L)x_{mt}$ , para  $j=1, 2, \dots, q$ ,  $t \in Z$ . Projecta-se  $\psi_t$  no espaço gerado pelas componentes principais

<sup>2</sup> Diz-se que uma propriedade vale *quase por toda a parte* ou *em quase todos os pontos* de um certo conjunto quando só deixa de se verificar nalgum subconjunto de medida nula

dinâmicas de ordem  $n$ ,  $\underline{p}_{nj}^x(L)\mathbf{x}_n$ ,  $j = 1, 2, \dots, q$ ,  $t \in \mathbb{Z}$ , com  $n > m$ . Seja  $\mathbf{y}_t$  essa projecção. Mostra-se então que, à medida que  $n$  e  $m$  aumentam, a distância entre  $\psi_t$  e  $\mathbf{y}_t$  diminui. É construída uma sequência de vectores de ruído branco  $q$ -dimensionais, cujas componentes são sequências de Cauchy e convergem para  $\mathcal{G}(\mathbf{x})$ .

**Mostra-se então que o espaço gerado pelas primeiras  $q$  componentes principais dinâmicas converge para o espaço gerado pelos  $q$  factores comuns.**

Note-se que, como consequência da definição do modelo e dos resultados acima apresentados, as componentes comum e idiosincrática são determinadas de forma única, o mesmo não acontecendo com o processo  $\mathbf{u}_t$ , ou com os filtros  $\underline{b}_i(L)$ . Além disso, os filtros  $\underline{b}_i(L)$ , de acordo com a definição do modelo, admitem um representação como soma de uma série, tal como em (10),  $b_{ij}(L) = \sum_{k=0}^{\infty} b_{ij,k} L^k$ , que são  $nq$  filtros cujos coeficientes são somáveis ao quadrado:  $\sum_{k=0}^{\infty} b_{ij,k}^2 < \infty$ . Refira-se também que, se  $\mathbf{x}$  for uma sequência de  $q$  factores dinâmicos,  $q$  é minimal, ou seja, não é possível uma representação de acordo com um modelo factorial dinâmico generalizado com um menor número de factores.

### 5.2.3.2. Identificação e Estimação Bilateral : Forni e al (2000)

Na secção anterior, o modelo factorial dinâmico generalizado foi exposto na sua vertente puramente teórica, identificando-se as condições a verificar na população para que esta apresente uma estrutura de acordo com este modelo.

Nesta secção, apresenta-se a primeira proposta de estimação das componentes comum e idiosincrática. Assume-se, portanto, que existe uma estrutura factorial dinâmica generalizada subjacente à população, verificando-se todas as hipóteses definidas no modelo.

A identificação da componente comum é realizada em dois passos. No primeiro passo, devido à Lei dos Grandes Números,  $q$  combinações lineares, escolhidas de modo “apropriado”, das observações tornam-se crescentemente colineares com os factores comuns quando  $n$  tende para infinito, “eliminando-se” deste modo as componentes idiosincráticas. Isto implica que o espaço gerado por essas  $q$  combinações lineares das observações converge para o espaço gerado pelos  $q$  factores comuns. Numa segunda etapa, projectam-se as observações no passado, presente e futuro de tais agregados, de forma a convergir para a componente comum, identificando-a. Em Forni et al. (2000), mostra-se que os agregados “apropriados” são as  $q$  primeiras componentes principais dinâmicas das observações. Como visto no Capítulo 4, as componentes principais dinâmicas são combinações lineares nos avanços e atrasos das variáveis  $\mathbf{X}_t$ . Brillinger (1981) mostra que a projecção de  $\mathbf{X}_t$  no presente, passado e futuro das primeiras  $q$  primeiras componentes principais dinâmicas (populacionais) constituem, para um dado

$n$ , a melhor aproximação em média quadrática de  $\mathbf{X}$ , através de  $q$  combinações lineares nos avanços e atrasos das observações.

Considere-se o espaço gerado pelas primeiras  $q$  componentes principais dinâmicas,

$$(16) \quad U_n = \overline{\text{span}}\left(\underline{\mathbf{p}}_{nj}(L)\mathbf{x}_{nt}, j = 1, \dots, q, t \in Z\right)$$

e as projecções ortogonais

$$(17) \quad \mathcal{X}_{it,n} = \text{proj}(x_{it} | U_n).$$

A expressão explícita de  $\mathcal{X}_{it,n}$  pode ser obtida considerando que:

$$(18) \quad \mathbf{I}_n = \mathbf{p}_{n1}^*(\theta)\mathbf{p}_{n1}(\theta) + \dots + \mathbf{p}_{nq}^*(\theta)\mathbf{p}_{nq}(\theta),$$

donde,

$$(19) \quad \mathbf{x}_{nt} = \underline{\mathbf{p}}_{n1}^*(L)\underline{\mathbf{p}}_{n1}(L)\mathbf{x}_{nt} + \dots + \underline{\mathbf{p}}_{nq}^*(L)\underline{\mathbf{p}}_{nq}(L)\mathbf{x}_{nt}.$$

Uma vez que as componentes principais dinâmicas são mutuamente ortogonais para qualquer avanço ou atraso, é possível escrever,

$$(20) \quad \mathcal{X}_{it,n} = \underline{\mathbf{K}}_{ni}(L)\mathbf{x}_{nt},$$

com,

$$(21) \quad \mathbf{K}_{ni}(\theta) = p_{n1,i}^*(\theta)\mathbf{p}_{n1}(\theta) + \dots + p_{nq,i}^*(\theta)\mathbf{p}_{nq}(\theta).$$

Ou seja,  $\mathcal{X}_{it,n}$  é a soma das projecções ortogonais de  $x_{it}$  nos avanços e atrasos de cada uma das primeiras  $q$  componentes principais dinâmicas. Os coeficientes da  $j$ -ésima projecção ortogonal são os coeficientes do filtro  $\underline{\mathbf{p}}_{nj}^*(L)$ . As observações e fórmulas definidas para a componente  $\mathcal{X}_{it,n}$ , são válidas para  $\xi_{it,n}$  e para as componentes principais dinâmicas  $q+1$  a  $n$ .

Sob as condições definidas no modelo factorial dinâmico generalizado, é possível demonstrar que,

$$(22) \quad \lim_{n \rightarrow \infty} \mathcal{X}_{it,n} = \mathcal{X}_{it},$$

em média quadrática, para quaisquer  $i$  e  $t$ .

Os filtros  $\underline{p}_{nj,i}^*(L)\underline{p}_{nj}(L)$ , são seqüências ponderadoras dinâmicas. A condição (iii) do modelo factorial dinâmico generalizado, garante que  $\underline{p}_{nj,i}^*(L)\underline{p}_{nj}(L)\xi_{nt}$  tende para zero quando  $n$  tende para infinito. Vindo, portanto, que na expressão

$$(23) \quad \underline{p}_{nj,i}^*(L)\underline{p}_{nj}(L)x_{nt} = \underline{p}_{nj,i}^*(L)\underline{p}_{nj}(L)\chi_{nt} + \underline{p}_{nj,i}^*(L)\underline{p}_{nj}(L)\xi_{nt},$$

quando  $n$  tende para infinito, apenas o termo  $\underline{p}_{nj,i}^*(L)\underline{p}_{nj}(L)\chi_{nt}$  não é eliminado (sobrevive).

Como já foi referido, se existir uma estrutura factorial dinâmica generalizada de  $q$  choques comuns subjacente aos dados, não é possível uma representação de acordo com o modelo definido com um menor número de factores. Ora, em situações empíricas, não se conhece o valor real de  $q$ . É possível, no entanto, demonstrar que a escolha de um valor para  $q$  superior ao verdadeiro valor não tem consequências graves.

Os resultados anteriores permitem determinar a componente comum de forma assintótica. Os filtros  $\underline{K}_{ni}(L)$ , utilizados na recuperação da componente comum, são determinados como funções da matriz de densidade espectral populacional,  $\Sigma_n(\theta) = (\sigma_{ij}(\theta))$ , de  $x_{nt}$ . Ora, na prática, esta matriz não é conhecida, sendo necessário estimá-la, a partir de concretização finita de  $x_{nt}$ , ou seja  $x_n^T$ .

Para garantir a consistência do estimador da matriz de densidade espectral, é necessário que se verifique ainda a seguinte condição:

(v) O vector  $x_{nt}$  admite uma representação dada por

$$(24) \quad x_{nt} = \sum_{k=-\infty}^{\infty} C_k Z_{t-k},$$

onde  $Z_t$  é um ruído branco  $n$  dimensional, com matriz de covariância não singular, momentos de quarta ordem finitos, verificando-se ainda

$$(25) \quad \sum_{k=-\infty}^{\infty} |C_{ij,k}| |k|^{1/2} < \infty$$

para  $i,j=1,\dots,n$ , onde  $C_{ij,k}$  é o elemento  $i,j$  de  $C_k$ .

Seja  $\Sigma_n^T(\theta) = (\sigma_{ij}^T(\theta))$ ,  $\theta \in [-\pi, \pi]$  um estimador arbitrário da matriz de densidade espectral  $\Sigma_n(\theta) = (\sigma_{ij}(\theta))$  de  $x_{nt}$ , baseado na matriz de covariâncias empírica  $\Gamma_{nk}^T = (n-k)^{-1} \sum_{t=k+1}^T x_{nt} x_{n,t-k}^T$ . Se  $x_{nt}$  satisfizer a condição anterior, este estimador é consistente, ou seja, converge em probabilidade para a matriz de densidade espectral, quando  $T$  tende para infinito.

Seja  $\lambda_{nj}^T(\theta)$  o  $j$ -ésimo maior valor próprio de  $\Sigma_n^T(\theta)$  e seja  $\mathbf{p}_{nj}^T(\theta) = (p_{nj,1}^T(\theta), \dots, p_{nj,n}^T(\theta))$  o correspondente vector próprio (vector linha). Os valores e vectores próprios de  $\Sigma_n^T(\theta)$  convergem em probabilidade para os valores e vectores próprios da matriz de densidade espectral populacional, pois são funções contínuas das entradas de  $\Sigma_n^T(\theta)$ . Assim sendo,  $\mathbf{K}_{ni}^T(\theta)$ , com  $i \leq q$ , é dado por:

$$(26) \quad \mathbf{K}_{ni}^T(\theta) = p_{n1,i}^T(\theta) \mathbf{p}_{n1}^T(\theta) + \dots + p_{nq,i}^T(\theta) \mathbf{p}_{nq}^T(\theta),$$

e tende para  $\mathbf{K}_{ni}(\theta)$  em probabilidade e uniformemente, para  $\theta \in [-\pi, \pi]$ , quando  $T$  tende para infinito.

Assim sendo, teoricamente, após a determinação de  $\Sigma_n^T(\theta)$ , seria possível determinar  $\mathbf{K}_{ni}^T(\theta)$  e, conseqüentemente, os coeficientes do correspondente filtro bilateral:

$$(27) \quad \underline{\mathbf{K}}_{ni}^T(L) = \sum_{k=-\infty}^{\infty} \mathbf{K}_{ni,k}^T L^k = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \left[ \int_{-\pi}^{\pi} \mathbf{K}_{ni}^T(\theta) e^{-ik\theta} d\theta \right] L^k.$$

Na prática, no entanto, não é possível obter a projecção de  $x_{it}$  no espaço gerado pelas primeiras  $q$  componentes principais empíricas,

$$(28) \quad \chi_{it,n}^T = \underline{\mathbf{K}}_{ni}^T(L) \mathbf{x}_{nt},$$

pois, para  $t \leq 0$  e  $t > T$ ,  $\mathbf{x}_{nt}$  não é conhecido.

Torna-se, pois, necessário considerar uma versão truncada dos filtros  $\underline{\mathbf{K}}_{ni}^T(L)$ , da forma  $\sum_{k=-M(T)}^{M(T)} \mathbf{K}_{ni,k}^T L^k$ , onde  $M(T) \rightarrow \infty$  é tal que  $\limsup_{T \rightarrow \infty} M^3(T)/T < \infty$ , quando  $T$  tende para infinito. No entanto, mesmo esta versão truncada deve ser ainda alterada, no que toca a  $\mathbf{x}_{nt}$ , quando  $t \leq M(T)$  ou  $t > T - M(T)$ , resultando no filtro de ordem finita,

$$(29) \quad \underline{\mathbf{K}}_{ni}^{Tt}(L) = \sum_{k=\max(-T-t, -M(T))}^{\min(t-1, M(T))} \mathbf{K}_{ni,k}^T L^k.$$

Ora, devido a esta necessidade de truncagem, para um dado  $t$ , a componente comum não poderá ser recuperada, mesmo quando  $n$  e  $T$  tendem para infinito. Assim sendo, é possível recuperar a componente comum apenas na parte central da amostra. De facto, é possível mostrar que quando se verificam as hipóteses do modelo factorial dinâmico generalizado e a condição (v), dada pelas expressões (24) e (25), então,

$$(30) \quad \forall \varepsilon > 0, \eta > 0, \exists N_0(\varepsilon, \eta): P \left[ \left| \underline{\mathbf{K}}_{ni}^{Tt}(L) \mathbf{x}_{nt} - \chi_{it}^T \right| > \varepsilon \right] \leq \eta,$$

para todo o  $n \geq N_0$ , para todo o  $T$  maior que algum  $T_0(n, \varepsilon, \eta)$  e para todo o  $t = t^*(T)$  satisfazendo a condição

$$(31) \quad 0 < a \leq \liminf_{T \rightarrow \infty} \frac{t^*(T)}{T} \leq \limsup_{T \rightarrow \infty} \frac{t^*(T)}{T} \leq b < 1.$$

Face aos resultados apresentados, é proposto o seguinte estimador bilateral. Para um dado  $M = M(T)$ , calcula-se as matrizes de covariâncias  $\Gamma_{nk}^T$ ,  $k = 0, 1, \dots, M$ , de  $\mathbf{x}_{nt}$  e  $\mathbf{x}_{n,t-k}$ , e as transformadas de Fourier discretas, para  $2M+1$  pontos, dadas por

$$(32) \quad \Sigma_n^T(\theta_h) = \sum_{k=-M}^M \Gamma_{nk}^T \omega_k e^{-ik\theta_h},$$

onde  $\Gamma_{n,-k}^T = \Gamma_{nk}^T$ ,  $\theta_h = 2\pi h / (2M+1)$ ,  $h = 0, 1, \dots, 2M$  e  $\omega_k = 1 - \lfloor |k| / (M+1) \rfloor$  são os ponderadores correspondentes a uma janela retardada de Bartlett de dimensão  $M$ . A estimação da matriz de densidade espectral é garantida se  $M(T) \rightarrow \infty$  e  $M(T)/T \rightarrow 0$ , quando  $T \rightarrow \infty$ .

De seguida, obtêm-se os  $q$  primeiros vectores próprios,  $\mathbf{p}_{nj}^T(\theta_h)$ ,  $j = 1, 2, \dots, q$ , de  $\Sigma_n^T(\theta_h)$ ,  $h = 0, 1, \dots, 2M$ . Note-se que, quando  $M = 0$ , estes valores próprios correspondem aos valores próprios da matriz de covariância de  $\mathbf{x}_{nt}$ , reduzindo-se as componentes principais dinâmicas às componentes principais estáticas. Calcula-se, então,  $\mathbf{K}_{ni}^T(\theta_h) = p_{ni,i}^T(\theta_h) \mathbf{p}_{ni}^T(\theta_h) + \dots + p_{ni,q}^T(\theta_h) \mathbf{p}_{ni,q}^T(\theta_h)$ ,  $h = 0, 1, \dots, 2M$ . Estima-se  $\underline{\mathbf{K}}_{ni}^T(L)$ ,  $i = 1, \dots, q$ , através da inversa das transformadas de Fourier discretas:

$$(33) \quad \underline{\mathbf{K}}_{ni,k}^T = \frac{1}{2M+1} \sum_{h=0}^{2M} \mathbf{K}_{ni}^T(\theta_h) e^{ik\theta_h}, \quad k = -M, \dots, M,$$

vindo

$$(34) \quad \underline{\mathbf{K}}_{ni}^T(L) = \sum_{k=-M}^M \underline{\mathbf{K}}_{ni,k}^T L^k.$$

A escolha de  $M$ , tal que  $M(T) = o(T^{1/3})$ , garante a consistência da matriz de densidade espectral assim estimada e a consistência da componente comum estimada. Embora existam outras alternativas, tomar  $M(T) = \text{round}\left(\frac{2}{3}T^{1/3}\right)$  funciona bastante bem em casos de simulação.

Ora, até ao momento, considerou-se  $q$  conhecido. Na prática, porém é necessário determiná-lo. Para tal, toma-se em consideração o resultado que afirma que quando a população apresenta uma estrutura factorial dinâmica com  $q$  factores comuns, então os primeiros  $q$  valores próprios de  $\Sigma_n^x(\theta)$  divergem q.p.t.p. em  $[-\pi, \pi]$ , enquanto que o

$(q+1)$ -ésimo está uniformemente limitado. Assim sendo, se existirem  $T$  observações para um número  $n$  grande de observações  $x_{it}$ , é possível estimar as matrizes de densidade espectral,  $\Sigma_r^T, r \leq n$ , e os respectivos valores próprios dinâmicos,  $\lambda_{rj}^T$ , para um dado conjunto de frequências. Se os valores próprios calculados apresentarem o comportamento que se segue, pode considerar-se que os dados foram gerados de acordo com um modelo factorial dinâmico generalizado:

1. a média em  $\theta$  dos primeiros  $q$  valores próprios diverge, enquanto que a média do  $(q+1)$ -ésimo é relativamente estável;
2. tomando-se  $r = n$ , existe uma diferença substancial entre a variância explicada pela  $q$ -ésima componente principal e a variância explicada pela  $(q+1)$ -ésima.

Como já referido, a escolha de um número de factores comuns superior ao verdadeiro, não tem consequências importantes para a estimação do modelo.

### 5.2.3.3. Identificação e Estimação Unilateral (2002)

O primeiro método de estimação da componente comum, do modelo factorial dinâmico generalizado proposto, apresenta, no entanto, o seguinte problema: o estimador definido para a componente comum é baseado no uso de um filtro bilateral sobre as observações, ou seja, de uma combinação linear de observações passadas, presentes e futuras. Embora a utilização de observações futuras não seja problemática, quando o que se pretende é a estimação da componente comum em pontos que não pertençam aos extremos da amostra, esta torna-se problemática quando se pretende fazer previsões para o futuro, uma vez que, para o final da amostra, não existem, obviamente, observações futuras disponíveis.

De forma a ultrapassar este problema, foi proposto este segundo método de estimação da componente comum, baseado no método anterior, de forma a tirar vanfagem da informação sobre a estrutura de covariância dinâmica existente em todo o painel, e visando a obtenção de um estimador unilateral, consistente.

Assim sendo, num primeiro passo, obtêm-se as matrizes de covariâncias empíricas, em todos os avanços e atrasos, das componentes comum e idiossincrática, a partir das transformadas de Fourier inversas das correspondentes matrizes de densidade espectral estimadas. Em seguida, as estimativas das matrizes de covariâncias são usadas na construção de combinações lineares do passado e presente das observações que apresentem menor rácio de variância idiossincrática - variância comum. Os agregados desejados são obtidos através da solução de um problema de componentes principais generalizado.

### 5.2.3.3.1. Hipóteses Adicionais

Para além das hipóteses necessárias para a obtenção de um estimador bilateral consistente da componente comum, a determinação de um estimador unilateral consistente implica a imposição de novas condições e alteração de outras:

- (vi) O modelo factorial dinâmico é válido com uma função de transferência VARMA( $S,s$ ), ou seja,

$$(35) \quad \mathbf{x}_{nt} = \boldsymbol{\chi}_{nt} + \boldsymbol{\xi}_{nt} = \mathbf{B}_n(L)[\mathbf{A}(L)]^{-1} \mathbf{u}_t + \boldsymbol{\xi}_{nt}$$

onde  $\mathbf{B}_n(L) = \mathbf{B}_0^n + \mathbf{B}_1^n L + \dots + \mathbf{B}_s^n L^s$  é um polinómio  $n \times q$  de ordem  $s$  e  $\mathbf{A}(L) = \mathbf{I} - \mathbf{A}_1 L - \dots - \mathbf{A}_s L^s$  é um polinómio  $q \times q$  de ordem  $S$ ; mais, todas as soluções de  $\det[\mathbf{A}(z)] = 0$ ,  $z \in \mathbb{C}$ , estão fora do círculo unitário e existe um valor  $m$  tal que  $\mathbf{B}_s^n \neq 0$  para  $n \geq m$  e  $S \leq s+1$ .

- (vii) Sendo  $\Gamma_{nk}^{\boldsymbol{\chi}}$  a matriz de covariância com desfasamento  $k$  de  $\boldsymbol{\chi}_{nt}$  e sendo  $\mu_{nj}^{\boldsymbol{\chi}}$  o  $j$ -ésimo valor próprio de  $\Gamma_{n0}^{\boldsymbol{\chi}}$ ,  $\lim_{n \rightarrow \infty} \mu_{nr}^{\boldsymbol{\chi}} = \infty$ , onde  $r = q(s+1)$ .

- (viii) Sendo  $\Gamma_{nk}^{\boldsymbol{\xi}}$  a matriz de covariância com desfasamento  $k$  de  $\boldsymbol{\xi}_{nt}$  e sendo  $\mu_{nj}^{\boldsymbol{\xi}}$ ,  $j = 1, \dots, n$ , os valores próprios de  $\Gamma_{n0}^{\boldsymbol{\xi}}$ , o menor deles,  $\mu_{nn}^{\boldsymbol{\xi}}$ , é limitado inferiormente por um valor superior a zero, quando  $n$  tende para infinito.

A hipótese (vii) substitui a hipótese (iv) imposta na definição do modelo factorial dinâmico generalizado. Tem a mesma função de garantir que os factores comuns não influenciam apenas um número finito de séries.

Note-se que, na hipótese (vi), escrevendo  $\mathbf{f}_t = (f_{1,t}, \dots, f_{q,t})'$  para representar  $[\mathbf{A}(L)]^{-1} \mathbf{u}_t$  e  $\mathbf{F}_t$  para representar  $(\mathbf{f}_t' \mathbf{f}_{t-1}' \dots \mathbf{f}_{t-s}')'$ , o modelo reduz-se ao modelo factorial estático:

$$(36) \quad \mathbf{x}_{nt} = \mathbf{B}_n(L) \mathbf{f}_t + \boldsymbol{\xi}_{nt} = \mathbf{C}_n \mathbf{F}_t + \boldsymbol{\xi}_{nt}$$

com  $r = q(s+1)$  factores e  $\mathbf{C}_n = (\mathbf{B}_0^n \mathbf{B}_1^n \dots \mathbf{B}_s^n)$ .

### 5.2.3.3.2. Estimação e Previsão

Tal como foi definido, o modelo assume a ortogonalidade, para todos os avanços e atrasos, de  $\chi_{it}$  e  $\xi_{it}$ . Assim sendo, o problema da estimação e previsão de  $x_{it}$  pode ser decomposto na estimação e previsão da componente comum e da componente idiossincrática em separado. Uma vez que o modelo pressupõe igualmente que as componentes idiossincráticas das diferentes séries sejam ortogonais ou apresentem uma pequena correlação cruzada,  $\xi_{it}$  pode ser previsto com base nos métodos usuais das séries temporais, univariados ou multivariados de pequena dimensão.

Como ocorria no primeiro método de estimação, a questão principal consiste na estimação e previsão da componente comum. Seja  $\mathcal{H}(\chi, \tau)$  o espaço de Hilbert gerado pelas variáveis  $\{\chi_{jt}, t \leq \tau, j \in N\}$ . Desejando-se prever  $\chi_{i,T+h}$ , o seu estimador linear óptimo, no sentido de erro quadrático médio, conhecendo-se  $\mathcal{H}(\chi, T)$ , seria a projecção de  $\chi_{i,T+h}$  em  $\mathcal{H}(\chi, T)$ :

$$(37) \quad \phi_{i,T+h|T} = \text{proj}(\chi_{i,T+h} | \mathcal{H}(\chi, T)) = \sum_{j=1}^q \sum_{k=0}^{\infty} b_{ij,k} \text{proj}(f_{j,T+h} | \mathcal{H}(\chi, T)),$$

onde  $b_{ij,k}$  são os coeficientes do filtro  $b_{ij}(L)$  definido na formulação do modelo factorial dinâmico generalizado.

Sendo  $\mathcal{G}(\mathbf{F}, \tau)$  o espaço  $r$ -dimensional gerado pelas  $r$  componentes de  $\mathbf{F}_\tau$ ,  $k = 1, 2, \dots, r$ , e, tomando em conta a hipótese a) formulada, a projecção anterior toma a forma:

$$(38) \quad \phi_{i,T+h|T} = \text{proj}(\chi_{i,T+h} | \mathcal{G}(\mathbf{F}, T)) = \mathbf{M}_{ih} \mathbf{F}_T,$$

onde  $\mathbf{M}_{ih}$  é um vector  $r$ -dimensional, determinável a partir dos coeficientes de  $A(L)$  e dos coeficientes  $b_{ij,k}$ . Ora, as quantidades envolvidas nesta expressão não são observáveis. Pretende-se então construir um estimador que convirja para este estimador óptimo, quando  $n$  e  $T$  tendem para infinito. Isto é conseguido em duas etapas: na primeira, constrói-se um estimador consistente de  $\mathcal{G}(\mathbf{F}, T)$ , denominado por  $\mathcal{G}_n^T(\mathbf{F}, T)$ ; na segunda, projecta-se  $\chi_{i,T+h}$  em  $\mathcal{G}_n^T(\mathbf{F}, T)$ .

A determinação de  $\mathcal{G}_n^T(\mathbf{F}, T)$ , como se irá mostrar, depende das matrizes de covariância das componentes comum e idiossincrática. De forma a que se possa beneficiar do conhecimento da estrutura de covariância dinâmica presente nos dados, obtêm-se  $\Gamma_{nk}^z$  e  $\Gamma_{nk}^\xi$  usando a metodologia desenvolvida para a construção do estimador bilateral da componente comum.

Assim sendo, começa-se por calcular as matrizes de densidade espectral empíricas,  $\Sigma_n^T(\theta)$ , tal como indicado para o estimador bilateral.

Sejam

$$(39) \quad \mathbf{P}_n^T(\theta) = \left( \mathbf{p}_{n1}^T(\theta) \quad \dots \quad \mathbf{p}_{nq}^T(\theta) \right)' \quad \text{e} \quad \mathbf{Q}_n^T(\theta) = \left( \mathbf{p}_{n,q+1}^T(\theta) \quad \dots \quad \mathbf{p}_{nn}^T(\theta) \right)'.$$

Defina-se:

$$(40) \quad \Sigma_n^{\chi T}(\theta) = \mathbf{P}_n^{T*}(\theta) \text{Diag}(\lambda_{n1}^T(\theta), \dots, \lambda_{nq}^T(\theta)) \mathbf{P}_n^T(\theta),$$

$$(41) \quad \Sigma_n^{\xi T}(\theta) = \mathbf{Q}_n^{T*}(\theta) \text{Diag}(\lambda_{n,q+1}^T(\theta), \dots, \lambda_{nn}^T(\theta)) \mathbf{Q}_n^T(\theta),$$

$$(42) \quad \tilde{\Sigma}_n^{\chi}(\theta) = \mathbf{P}_n^*(\theta) \text{Diag}(\lambda_{n1}(\theta), \dots, \lambda_{nq}(\theta)) \mathbf{P}_n(\theta)$$

e

$$(43) \quad \tilde{\Sigma}_n^{\xi}(\theta) = \mathbf{Q}_n^*(\theta) \text{Diag}(\lambda_{n,q+1}(\theta), \dots, \lambda_{nn}(\theta)) \mathbf{Q}_n(\theta),$$

onde  $\mathbf{P}_n(\theta)$  e  $\mathbf{Q}_n(\theta)$  são as correspondentes populacionais de  $\mathbf{P}_n^T(\theta)$  e  $\mathbf{Q}_n^T(\theta)$ , respectivamente. Sendo satisfeitas as condições previstas na construção do estimador bilateral, as matrizes

$$(44) \quad \Gamma_{nk}^{\chi T} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\theta} \Sigma_n^{\chi T}(\theta) d\theta \quad \text{e} \quad \Gamma_{nk}^{\xi T} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\theta} \Sigma_n^{\xi T}(\theta) d\theta$$

são estimadores consistentes, quando  $T$  tende para infinito, de

$$(45) \quad \tilde{\Gamma}_{nk}^{\chi} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\theta} \tilde{\Sigma}_n^{\chi}(\theta) d\theta \quad \text{e} \quad \tilde{\Gamma}_{nk}^{\xi} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\theta} \tilde{\Sigma}_n^{\xi}(\theta) d\theta$$

### 5.2.3.3.2.1. Problema de Valores Próprios Generalizados

Antes de se passar propriamente à determinação de um estimador consistente do espaço  $r$ -dimensional  $\mathcal{G}(\mathbf{F}, T)$ , é necessária a familiarização com o problema de valores próprios generalizados.

Sendo  $\Gamma$  e  $\mathbf{D}$  duas matrizes  $n \times n$ , reais, simétricas e semidefinidas positivas, considerem-se os valores complexos  $v_j$  e os vectores complexos  $\mathbf{v}_j$ , tais que:

$$(46) \quad \begin{aligned} \Gamma \mathbf{v}_j^* &= v_j \mathbf{D} \mathbf{v}_j^*, \\ \mathbf{v}_j \mathbf{D} \mathbf{v}_j^* &= 1, \\ \mathbf{v}_j \mathbf{D} \mathbf{v}_m^* &= 0, \quad j \neq m \end{aligned}$$

As soluções  $v_j$  e  $\mathbf{v}_j$ , do problema anterior, são os valores próprios generalizados e os vectores próprios generalizados, respectivamente, do par de matrizes  $(\Gamma, \mathbf{D})$ .

Usando os valores e vectores próprios generalizados é possível enunciar o seguinte resultado.

Sejam  $\Gamma$  e  $\mathbf{D}$  duas matrizes  $n \times n$ , reais, simétricas e semidefinidas positivas. Considerem-se os  $n$  problemas de maximização:

$$(47) \quad \begin{aligned} \mathbf{a}_j &= \text{Arg max}_{\mathbf{a} \in R^n} \mathbf{a} \Gamma \mathbf{a}^* \\ \text{s.a.:} \quad \mathbf{a} \mathbf{D} \mathbf{a}^* &= 1, \\ \mathbf{a} \mathbf{D} \mathbf{a}_m^* &= 0 \quad \text{para } 1 \leq m \leq j-1 \end{aligned}$$

para  $j = 1, 2, \dots, n$ . Sendo  $v_j$  e  $\mathbf{v}_j$ , respectivamente, os valores e vectores próprios generalizados do par de matrizes  $(\Gamma, \mathbf{D})$ , as soluções dos problemas de maximização anteriores são  $\mathbf{a}_j = \mathbf{v}_j$  e  $\mathbf{a}_j \Gamma \mathbf{a}_j^* = v_j$ .

### 5.2.3.3.2.2. 1ª Etapa: Estimação Consistente do Espaço de Factores Comuns

Para se estimar o espaço gerado pelos factores comuns, pretende-se determinar  $r$  agregados lineares de  $\mathbf{x}_{nt}$ , da forma

$$(48) \quad W_{nt}^{kT} = \mathbf{Z}_{nk}^T \mathbf{x}_{nt}, \quad k = 1, \dots, r,$$

e mostrar que o espaço  $\mathcal{G}_n^T(\mathbf{F}, T)$ , gerado por  $W_{nT}^{1T}, \dots, W_{nT}^{rT}$ , converge para  $\mathcal{G}(\mathbf{F}, T)$ , quando  $n$  e  $T$  tendem para infinito.

De forma a que estes agregados maximizem os quocientes das variâncias das componentes comum e idiosincrática, escolhem-se os ponderadores  $\mathbf{Z}_{nk}^T$  como sendo as soluções, para  $1 \leq l \leq r$ , dos seguintes problemas de optimização:

$$(49) \quad \begin{aligned} \mathbf{Z}_{nl}^T &= \text{Arg max}_{\mathbf{a} \in R^n} \text{var}(\mathbf{a} \chi_{nl}^T) \\ \text{s.a.:} \quad \text{var}(\mathbf{a} \xi_{nl}^T) &= 1 \\ \mathbf{a} \xi_{nl}^T \perp \mathbf{Z}_{nm}^T \xi_{nl}^T, \quad &1 \leq m < l, \quad 1 \leq l \leq n \end{aligned}$$

Ora, estes problemas de optimização podem ser reescritos da seguinte forma:

$$(50) \quad \begin{aligned} \mathbf{Z}_{nl}^T &= \text{Arg max}_{\mathbf{a} \in R^n} \mathbf{a} \Gamma_{n0}^{\chi T} \mathbf{a}^* \\ \text{s.a.:} \quad \mathbf{a} \Gamma_{n0}^{\xi T} \mathbf{a}^* &= 1 \\ \mathbf{a} \Gamma_{n0}^{\xi T} \mathbf{Z}_{nm}^T &= 0, \quad 1 \leq m < l, \quad 1 \leq l \leq n \end{aligned}$$

reduzindo-se a problemas de valores próprios generalizados. Assim, como foi visto, os vectores  $\mathbf{Z}_{nj}^T$ ,  $j = 1, \dots, n$ , são os vectores próprios generalizados associados aos valores próprios generalizados  $v_{nj}^T$  do par de matrizes  $(\Gamma_{n0}^{\chi T}, \Gamma_{n0}^{\varepsilon T})$ , ou seja,

$$(51) \quad \mathbf{Z}_{nj}^T \Gamma_{n0}^{\chi T} = v_{nj}^T \mathbf{Z}_{nj}^T \Gamma_{n0}^{\varepsilon T}, \quad j = 1, \dots, n.$$

Assim sendo, os agregados pretendidos são as componentes principais generalizadas.

Seja  $w_{nt}^{kT}$  a versão estandardizada de  $W_{nt}^{kT}$ ,  $w_{nt}^{kT} = \mathbf{z}_{nk}^T \mathbf{x}_{nt}$ , onde  $\mathbf{z}_{nk}^T = \mathbf{Z}_{nk}^T / \sqrt{1 + v_{nk}^T}$ . Os vectores  $w_{nt}^{kT}$ ,  $k = 1, \dots, r$ , formam um sistema ortonormal que gera um espaço da mesma dimensão que  $\mathcal{G}(\mathbf{F}, T)$ . O seguinte resultado estabelece, formalmente, que o espaço gerado pelos agregados assim determinados aproxima o espaço não observável  $\mathcal{G}(\mathbf{F}, T)$ .

Sob as hipóteses formuladas, verifica-se:

$$(52) \quad \forall \varepsilon > 0, \eta > 0, \exists N_0 = N_0(\varepsilon, \eta), T_0 = T_0(n, \varepsilon, \eta): P\left[\left|w_{nt}^{kT} - \text{proj}\left(w_{nt}^{kT} \mid \mathcal{G}(\mathbf{F}, t)\right)\right| > \varepsilon\right] \leq \eta$$

para todo  $1 \leq t \leq T, n \geq N_0, T \geq T_0$  e  $1 \leq k \leq r$ .

### 5.2.3.3.2.3. 2ª Etapa: Construção do Estimador Ótimo

Para determinar  $\chi_{i,T+h}$  propõe-se a estimativa da projecção de  $x_{i,T+h}$  no espaço gerado pelos  $r$  agregados  $W_{nT}^{kT}$ ,  $k = 1, \dots, r$ , isto é,

$$(53) \quad \hat{\chi}_{i,T+h}^{nT} = \phi_{i,T+h|T}^{nT} = \text{proj}\left(x_{i,T+h} \mid W_{nT}^{1T}, \dots, W_{nT}^{rT}\right) = \left(\Gamma_{nh}^{\chi T} \mathbf{Z}_n^T \left(\mathbf{Z}_n^T \Gamma_{n0}^T \mathbf{Z}_n^T\right)^{-1} \mathbf{Z}_n^T \mathbf{x}_{nt}\right)_i$$

onde  $\mathbf{Z}_n^T = \left(\mathbf{Z}_{n1}^T \quad \dots \quad \mathbf{Z}_{nr}^T\right)'$ .

O seguinte resultado mostra que  $\phi_{i,T+h|T}^{nT}$  é assintoticamente equivalente, (quando  $n$  e  $T$  tendem para infinito), à melhor estimação no sentido da média quadrática dada por  $\phi_{i,T+h|T}$ .

Sob as hipóteses formuladas, verifica-se que:

$$(54) \quad \forall \varepsilon > 0, \eta > 0, \exists N_0 = N_0(\varepsilon, \eta), T_0 = T_0(n, \varepsilon, \eta): P\left[\left|\phi_{i,T+h|T}^{nT} - \phi_{i,T+h|T}\right| > \varepsilon\right] \leq \eta,$$

para todo  $n \geq N_0, T \geq T_0$  e  $1 \leq i \leq N_0$ .

De forma análoga, é possível estimar a componente comum, na parte central da amostra, projectando  $x_{it}$ ,  $t \leq T$ , no espaço gerado pelos agregados  $W_{nt}^{kT}$ ,  $k = 1, \dots, r$ :

$$(55) \quad \hat{\chi}_{i,t}^{nT} = \hat{\phi}_{i,t}^{nT} = \text{proj}(x_{i,t} | W_{nt}^{1T}, \dots, W_{nt}^{rT}) = \left( \Gamma_{n0}^{xT} \mathbf{Z}_n^T (\mathbf{Z}_n^T \Gamma_{n0}^T \mathbf{Z}_n^T)^{-1} \mathbf{Z}_n^T x_{nt} \right)_t$$

Nas mesmas condições que as previstas para o caso da previsão, este estimador tende em probabilidade para  $\chi_{it}$ , sendo, desta forma, um estimador unilateral consistente da componente comum. Deste modo, são evitados os problemas do primeiro estimador proposto (bilateral), sendo ainda possível explorar o conhecimento da estrutura de covariâncias dinâmicas.

### 5.2.3.3.3. Variáveis Não Estacionárias

O caso das variáveis que são estacionárias apenas após ser retirada a tendência ou calculada a primeira diferença é igualmente acomodável ao modelo factorial dinâmico generalizado.

No primeiro caso, viria  $x_{it} = T_t + z_{it}$ , onde  $T_t$  é uma tendência determinística. Assim sendo, os resultados enunciados seriam aplicados à série  $z_{it}$ , estacionária.

No segundo caso, considere-se que  $x_{it}$  é I(1). Assim sendo, as séries  $y_{it} = (1-L)x_{it}$  são estacionárias. Estando estas novas variáveis nas condições do modelo, é possível a representação:

$$(56) \quad \begin{aligned} (1-L)x_{it} &= \chi_{it} + \xi_{it} \\ \chi_{it} &= \mathbf{b}_i(L)\mathbf{u}_t \end{aligned}$$

onde  $\mathbf{u}_t$  é  $q$ -dimensional e  $\xi$  é idiossincrático. Como já mencionado, os vectores  $\chi_{it}$  e  $\xi_{it}$  são únicos, bem como as suas correspondentes matrizes de densidade espectral. Assim sendo, partindo da informação fornecida pelas matrizes de densidade espectral de  $x$ , é possível determinar se as componentes comuns e idiossincráticas são I(1) ou I(0) ou se existem relações de cointegração. É possível, depois, aplicar os resultados enunciados à série resultante da primeira diferença e após a subtracção da média.

Por vezes, a estacionarização das séries pode envolver a aplicação de logaritmos. E por vezes, é necessário normalizar as séries, subtraindo a média e dividindo pelo desvio-padrão. Também nestes casos é possível aplicar o modelo factorial dinâmico generalizado às séries estacionárias.

## Referências

- Brillinger, D. R. (1981), *Time Series Analysis: Data Analysis and Theory*, Holden-Day, San Francisco.
- Brockwell, P. J.; Davis, R. A. (1991), *Time Series: Theory and Methods, Second Edition*, Springer-Verlag, New York.
- Chamberlain, G.; Rothschild, M. (1983), "Arbitrage, factor structure and mean-variance analysis in large asset markets", in *Econometrica*, **51**, 1305-1324.
- Cristadoro, R.; Forni, M.; Reichlin, L.; Veronese, G. (2001), "A core inflation index for the Euro area", *CEPR working paper*.
- Dias Agudo, F. R. (1990), *Análise Real Vol. II*, Escolar Editora, Lisboa.
- Forni, M.; Hallin, M.; Lippi, M.; Reichlin, L. (1999), "Reference cycles: the NBER methodology revisited", *manuscrito*, Dezembro.
- Forni, M.; Hallin, M.; Lippi, M.; Reichlin, L. (2000), "The generalized dynamic factor model: identification and estimation", *The Review of Economics and Statistics*, Novembro, **82(4)**, 540-554.
- Forni, M.; Hallin, M.; Lippi, M.; Reichlin, L. (2001a), "Coincident and leading indicators for the Euro area", *The Economic Journal*, **111**, 62-85.
- Forni, M.; Hallin, M.; Lippi, M.; Reichlin, L. (2001b), "Do Financial variables help forecasting inflation and real activity in the Euro area?", *CEPR working paper*
- Forni, M.; Hallin, M.; Lippi, M.; Reichlin, L. (2002), "The generalized dynamic factor model: one-sided estimation and forecasting", *manuscrito*, Junho.
- Forni, M.; Lippi, M. (2001), "The generalized dynamic factor model: representation theory", *Econometric Theory*, **17**, 1113-1141.
- Forni, M.; Reichlin, L. (1996), "Dynamic common factors in large cross-sections", *Empirical Economics*, **21**, 27-42.
- Forni, M.; Reichlin, L. (1998), "Let's get real: a factor analytic approach to disaggregated business cycle dynamics", *Review of Economic Studies*, **65**, 453-473.
- Forni, M.; Reichlin, L. (2001), "National policies and local economies", *European Economic Review*, Janeiro.
- Geweke, J. (1977), "The dynamic factor analysis of economic time series", in Dennis J. Aigner and Arthur S. Goldberger (Eds.), *Latent Variables in Sócio-Economic Models*, North-Holland, Amsterdam.

- Müller, D. A. (1996), *Análise Espectral e Modelos não Lineares em Séries Temporais*, Série Textos de Apoio, n.º 10/TA, CEMAPRE, ISEG, UTL, Novembro.
- Müller, D. A.; Murteira, B. J. F.; Turkman, K. F. (1993), *Análise de Sucessões Cronológicas*, McGraw-Hill, Lisboa.
- Quah, D.; Sargent, T.J. (1993), "A dynamic index model for large cross sections", in J. H. Stock and M. W. Watson (Eds.), *Business Cycles, Indicators and Forecasting*, NBER, 285-309.
- Reichlin, L. (2002), "Factor models in large cross-sections of time series", manuscrito, Fevereiro.
- Sargent, T. J.; Sims, C. A. (1977), "Business cycle modelling without pretending to have too much *a priori* economic theory", in Christopher A. Sims (Ed.), *New Methods in Business Research*, Federal Reserve Bank of Minneapolis, Minneapolis.
- Stock, J. H.; Watson, M. W. (1989), "New indexes of coincident and leading economic indicators", NBER *Macroeconomic Annual*, 351-394.
- Wei, W. W. S. (1990), *Time Series Analysis: Univariate and Multivariate Methods*, Addison-Wesley.