

ANÁLISE ESPECTRAL DE
UMA CLASSE DE PROCESSOS
ESTOCÁSTICOS NÃO ESTACIONÁRIOS

ELISABETE CORRÊA LEME

DISSERTAÇÃO APRESENTADA AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA OBTENÇÃO DO GRAU DE MESTRE
EM
ESTATÍSTICA

Orientador:

Prof. Dr. PEDRO ALBERTO MORETTIN

- SÃO PAULO, novembro de 1975 -

AGRADECIMENTOS

Ao Prof. Dr. Pedro Alberto Morettin, pela paciência e amizade com que orientou este trabalho.

Ao Prof. Dr. Odelar Leite Linhares, pelo incentivo constante e empenho em nos proporcionar excelentes condições de trabalho.

Aos colegas e amigos do Instituto de Ciências Matemáticas, a colaboração e o apoio durante nossos estudos e em especial ao Flávio Grisi, cujo auxílio na parte de programação foi indispensável.

Ao Sr. Mário Ulysses Franchini pelo cuidadoso trabalho de datilografia.

Este trabalho dependeu parcialmente de auxílios fornecidos pela CAPES e FAPESP.

Í N D I C E

I. INTRODUÇÃO	1
II. SOBRE PROCESSOS ESTOCÁSTICOS ESTACIONÁRIOS.	11
1. Introdução	11
2. Funções média e covariância de um processo estocástico	14
3. Processos com incrementos ortogonais	16
4. Processos estacionários de segunda ordem	20
4-1. Representação e decomposição espectral de um processo estacionário	21
4-2. Operações lineares sobre processos estacionários.	28
4-3. Sobre a estimação da função densidade espectral	30
III. SOBRE PROCESSOS NÃO ESTACIONÁRIOS	41
1. Considerações gerais	41
2. O caso discreto.	44
3. O caso contínuo.	52
4. A classe dos processos harmonizáveis	57
IV. TEORIA ESPECTRAL PARA UMA CLASSE DE PROCESSOS NÃO ESTACIONÁRIOS	68
1. Introdução	68
2. A classe dos processos oscilatórios.	69
3. Efeito de filtros.	78
4. Processos semi-estacionários	82
5. Determinação de espectro evolutivo	87
6. Estimação de espectros evolutivos.	94
7. Considerações sobre o planejamento	104
8. O caso discreto.	111
APENDICE.	115
BIBLIOGRAFIA.	135

I - INTRODUÇÃO

Tratamos nesta monografia de alguns tipos de processos estocásticos com características não estacionárias analisando, em particular, a classe dos processos oscilatórios introduzida por Priestley.

Inicialmente são estabelecidos os principais resultados sobre processos estocásticos estacionários, de interesse para o desenvolvimento da teoria posterior, devendo-se atribuir particular atenção aos teoremas concernentes à representação, através de certas integrais de Fourier (estocásticas), tanto de um processo estacionário quanto de sua função covariância.

Nesse sentido, convém aqui observar que em muitas aplicações envolvendo transformadas de Fourier onde existe interesse em se estabelecer as relações entre duas funções digamos, $f(t)$ e $g(\lambda)$, com a variável t representando "tempo" e λ a "frequência angular" de algum tipo de oscilação, a relação entre $f(t)$ e $g(\lambda)$ pode ser expressa na forma de uma integral de Fourier-Stieltjes:

$$f(t) = \int_{-\infty}^{\infty} e^{it\lambda} dg(\lambda) \quad (1.1)$$

que mostra como a função $f(t)$ é aditivamente construída pelas oscilações harmônicas elementares $e^{it\lambda} dg(\lambda)$, tendo cada uma delas frequência angular λ , com fase e amplitude determinadas por $dg(\lambda)$.

Os teoremas de representação apresentam resultados envolvendo integrais do mesmo tipo de (1.1) e estabelecem que, se (X_t) for um processo estocástico estacionário ao qual está associado um processo (Y_λ) com incrementos ortogonais, para cada t fixado é possível a representação:

$$X_t = \int e^{it\lambda} dY_\lambda \quad \text{com } E(|dY_\lambda|^2) = dF(\lambda) \quad (1.2)$$

onde o processo (Y_λ) é definido a menos de uma variável aleatória aditiva e F é uma função não decrescente e limitada denominada *função distribuição espectral* do processo (X_t) , em termos da qual a função covariância deste processo admite uma representação da forma:

$$R(u) = \int e^{iu\lambda} dF(\lambda) \quad (1.3)$$

Pode-se observar, em analogia com o que foi mencionado de (1.1), que em (1.2) o processo (X_t) é aditivamente construído por oscilações harmônicas mutuamente ortogonais $e^{it\lambda} dY_\lambda$, cada uma das quais tendo frequência angular λ . Dessa forma pode-se dizer que (1.2) fornece uma representação espectral do processo estacionário (X_t) .

Por outro lado, como toda função covariância de um processo estacionário pode ser representada na forma (1.3), consi-

dera-se esta representação também como uma representação espectral de $R(u)$.

Sendo de se notar que através da função distribuição espectral é possível se obter informações sobre um processo estacionário em estudo e a fim de se estabelecer comparações com resultados apresentados no caso de um processo com características não estacionárias, são feitas no final do capítulo II rápidas considerações sobre o problema da estimação do espectro (denominação usual para $dF(\lambda)$).

Muito embora toda uma bem desenvolvida teoria sobre processos estocásticos estacionários já tenha sido construída, na prática no entanto, muitas vezes a suposição de estacionariedade se apresenta como duvidosa.

Em Economia, por exemplo, é comum se supor que (X_t) possa ser representado na forma

$$X_t = m_t + c_t + s_t + \varepsilon_t \quad (1.4)$$

onde m_t é o termo de tendência, c_t representa os ciclos econômicos, s_t a componente sazonal e ε_t é o resíduo, considerado estacionário, de modo que (X_t) se torna um processo não estacionário. Problemas de importância que se apresentam são os de remover os componentes m_t , c_t e s_t . Para detalhes sobre o assun-

to, ver Granger e Hatanaka [11] e Gait [10].

Outra situação é a considerada por Brillinger e Hatanaka [2] onde são analisados processos multivariados não estacionários, em Economia, considerando-se modelos lineares da forma

$$Y_t = \sum_{j=1}^n \int a_j(t-u) X_j(u) du + \epsilon_t ,$$

sendo $\{Y_t, X_1(t), \dots, X_n(t)\}$ um processo estocástico $(n+1)$ variado estacionário e ϵ_t o distúrbio. Tratamento análogo é feito pelos autores para sistemas de equações simultâneas.

Nesta mesma linha podem ser considerados sistemas não estacionários de equações de diferenças, estocásticas, lineares, da forma

$$\tilde{Y}_t = \tilde{A}\tilde{Y}_{t-1} + \tilde{\epsilon}_t$$

onde algumas raízes características de \tilde{A} são maiores do que 1, em valor absoluto. O comportamento de tais sistemas é difícil de se analisar pois as funções de covariância não convergem e as densidades espectrais não estão definidas. Ver, por exemplo, Chow e Levitan [4] e Rao [23].

Para um tratamento de modelos lineares não estacionários

rios utilizados para fins de previsão e controle, basicamente modelos do tipo auto regressivo-médias móveis, a melhor referência é o livro de Box e Jenkins [3].

Pode-se ver ainda o recente artigo de Priestley e Tong [19], que estende a análise a ser aqui apresentada para o caso de processos não estacionários bivariados.

Em tais casos, portanto, em que a suposição de estacionariedade não é assumida, a análise espectral clássica baseada em um modelo estacionário, dificilmente pode ser admitida com convicção.

O capítulo III vem então considerar processos estocásticos gerais apresentando resultados sobre a decomposição e representação de tais processos. Generalizam-se certas propriedades dos processos estacionários para certas classes de processos não estacionários, em particular, propriedades relacionadas ao problema da predição linear de mínimos quadrados. Nas secções 1, 2 e 3 deste capítulo o caso geral de um processo estocástico vetorial com momentos de segunda ordem finitos, é estudado. Para um tal processo existe uma decomposição, determinada de maneira única, em um componente determinístico e um puramente não determinístico que são mutuamente ortogonais. No caso de um processo vetorial a parâmetro discreto as propriedades dessa decomposição são uma generalização direta da conhecida decomposição de Wold para processos univariados estacioná-

rios a parâmetro discreto; este caso é tratado com detalhes na secção 2. Para um processo a parâmetro contínuo são estabelecidos resultados análogos ao caso discreto, de uma forma mais sus
cinta.

A decomposição em um componente determinístico e um puramente não determinístico formam a base de uma análise de um processo estocástico no domínio do parâmetro t (ou "tempo"), generalizando as bem conhecidas propriedades dos processos estacionários. Para a classe destes últimos, a grande vantagem de se passar da análise no domínio do tempo para a no domínio da frequência (ou análise espectral) é a possibilidade de se usar os poderosos métodos da análise harmônica. Isto pode ser feito para processos estacionários já que para estes existem as repre
sentações especiais (1.2) e (1.3) tanto para os processos como para as funções covariância associadas. No entanto, um problema que surge ao se tratar com processos não estacionários é o da possibilidade de se formular uma teoria espectral envolvendo conceitos clássicos tais como "energia" e "frequência" de modo que uma função espectral definida, ainda possua uma interpretação significativa e útil. Deve-se observar (secções 2 e 3 do ca
pítulo) que nas expressões de representação de um processo não estacionário não ocorre o aparecimento da função exponencial sur
gindo daí a dificuldade em se estabelecer uma definição razoável do espectro para este caso.

Muitas tentativas já foram feitas para se estabelecer uma tal definição, cada uma delas tendo como objetivo se determinar uma única função cujas propriedades dependessem do comportamento do processo sobre todo o espaço dos parâmetros. Cramér considerou uma classe de processos (Processos Harmonizáveis, primeiramente introduzida por Loève) que admitem uma representação da forma

$$X_t = \int e^{it\lambda} dZ_\lambda$$

sem a restrição de (Z_λ) ser um processo com incrementos ortogonais, e definiu para este caso o *espectro integrado*, agora uma função de duas variáveis, como sendo

$$dF(\lambda, \nu) = E(|dZ_\lambda \cdot d\bar{Z}_\nu|)$$

Esta classe de processos é apresentada encerrando o capítulo III.

Como uma referência dentro do problema de se definir o espectro para o caso não estacionário, é de valor o artigo de R.M. Loynes [15], que apresenta considerações sobre as condições necessárias a uma função para que seja admitida como o espectro e sugere algumas propriedades desejáveis ao se tentar estabelecer essa definição, a saber:

A_1 : o espectro deverá ser uma função real do "tempo"

e da "frequência", completamente determinado pela função covariância.

A_2 : deverá descrever a distribuição de energia sobre frequência.

A_3 : a correspondência entre o espectro e a função covariância deverá ser biunívoca.

A_4 : o espectro deverá se transformar razoavelmente e preferivelmente de maneira mais simples, se o processo (X_t) for transformado linearmente. Em particular, um conhecimento do espectro de (X_t) determinará o espectro do processo transformado.

A_5 : deverá se reduzir ao espectro usual ou a alguma transformação simples deste, se (X_t) for de fato estacionário.

A_6 : se o processo for constituído de uma sucessão de partes estacionárias, digamos $(X_t^{(1)})$ para $t \leq 0$ e $(X_t^{(2)})$ se $t > 0$, o espectro será composto de uma sucessão dos espectros estacionários correspondentes.

A_7 : deverá ser uma transformada de Fourier, ou alguma transformada relacionada, de alguma quantidade aparentemente significativa.

São discutidas, ainda no mesmo artigo, as definições dos espec-

tros sugeridos por Fano (1950), Page (1952), Levin (1964), Dubman (1965) e por Priestley (1965), sendo a teoria relacionada com este último de especial interesse neste trabalho e apresentada aqui a partir do capítulo IV.

Em Priestley [20], este apresenta seu espectro introduzindo uma classe de processos não estacionários, os Processos Oscilatórios, e definindo para estes o que ele chama de *espectro de evolução* (ou *espectro evolutivo*) que é uma quantidade espectral com interpretação física semelhante à do espectro de um processo estacionário e que possui uma certa relação com o resultado apresentado por Page [16] quando introduziu a noção de *espectro instantâneo de potência*.

A fim de estabelecer uma definição de modo a tornar, também para o caso não estacionário, a noção de frequência fisicamente significativa, Priestley deliberadamente considera uma representação espectral para o processo, da forma

$$X_t = \int A_t(\lambda) e^{it\lambda} dY_\lambda$$

impondo condições especiais à família de funções $\{A_t(\lambda)\}$.

A sua definição de espectro evolutivo é estabelecida de modo a possibilitar uma interpretação física muito semelhante à do espectro de um processo estacionário e descreve uma distribuição local de energia sobre frequência, com "frequência" se referin-

do $\tilde{\lambda}$ constante λ da função exponencial complexa $e^{i\lambda t}$.

É mostrado que a noção de espectro evolutivo generaliza a definição usual de espectro para processos estacionários e que, em cada instante de tempo, sob certas condições, pode-se estimar o espectro evolutivo a partir de uma única realização do processo.

A fim de examinar a validade dos métodos de estimação sugeridos por Priestley, são planejadas no final deste trabalho, realizações de um processo não estacionário artificial, gerado a partir de modelos conhecidos.

II - SOBRE PROCESSOS ESTOCÁSTICOS ESTACIONÁRIOS

1. Introdução

Dado um espaço probabilístico (Ω, B, P) , seja $L_2(\Omega, B, P)$ o conjunto de todas as variáveis aleatórias X , a valores complexos, definidas em Ω , tais que

$$E(|X|^2) = \int_{\Omega} |X(\omega)|^2 dP(\omega) < \infty \quad (2.1.1)$$

Definição 1.1 :

Duas variáveis aleatórias X e Y , pertencentes a $L_2(\Omega, B, P)$ serão chamadas *variáveis aleatórias equivalentes* se

$$P(X=Y) = 1 \quad \text{ou} \quad E(|X-Y|^2) = 0$$

Variáveis aleatórias equivalentes serão observadas aqui como idênticas, de modo que ao nos referirmos a uma particular variável aleatória, estaremos nos referindo à classe de todas as que lhe forem equivalentes.

Com essa convenção, o conjunto $L_2(\Omega, B, P)$ será um espaço de Hilbert (com as operações de adição e multiplicação por escalares definidas respectivamente como a adição usual de

variáveis aleatórias e a multiplicação usual por números complexos), com produto interno entre duas variáveis aleatórias X e Y definido por

$$\langle X, Y \rangle = E(X \cdot \bar{Y}) \quad (2.1.2)$$

onde \bar{Y} representa a conjugada de Y .

Consequentemente, a norma da variável aleatória X será dada por :

$$\|X\| = \langle X, X \rangle^{1/2} = [E(|X|^2)]^{1/2} \quad (2.1.3)$$

Com esta definição, a convergência em norma de uma sequência de pontos em $L_2(\Omega, B, P)$ coincide com a convergência em média quadrática da sequência de variáveis aleatórias correspondente.

Seja agora $(X_t : t \in T)$ um processo estocástico complexo, isto é, existe um espaço probabilístico (Ω, B, P) tal que para cada t , $X_t : \Omega \longrightarrow \mathbb{C}$ é uma variável aleatória. O conjunto T , em geral, será considerado \mathbb{Z} (e o processo será dito processo a parâmetro discreto) ou um intervalo real (processo a parâmetro contínuo).

Definição 1.2 :

Um processo estocástico complexo será de segunda or-

dem se para todo $t \in T$, $X_t \in L_2(\Omega, B, P)$.

Portanto para cada $t \in T$, X_t pode ser observada como um ponto no espaço de Hilbert $L_2(\Omega, B, P)$ sendo então possível se definir o espaço de Hilbert gerado por um processo estocástico.

Definição 1.3 :

Dado um processo estocástico complexo $(X_t : t \in T)$, o espaço vetorial gerado por tal processo, representado por $S(X_t : t \in T)$, é por definição o conjunto de todas as variáveis aleatórias U que puderem ser escritas na forma $U = \sum_{i=1}^n z_i X_{t_i}$ para algum inteiro n , constantes complexas z_1, z_2, \dots, z_n e pontos t_1, t_2, \dots, t_n em T .

Definição 1.4 :

Dado um processo estocástico complexo de segunda ordem define-se o espaço de Hilbert gerado por tal processo, representado por $L(X_t : t \in T)$, como o constituído de todas as variáveis aleatórias do sub-espaço $S(X_t : t \in T)$ juntamente com todas as variáveis aleatórias U tais que existe uma sequência de variáveis aleatórias U_n em $L(X_t : t \in T)$ convergindo para U no sentido que, quando $n \rightarrow \infty$,

$$\|U_n - U\|^2 = E(|U_n - U|^2) \rightarrow 0$$

Prova-se que $L(X_t : t \in T)$ é um espaço de Hilbert. De fato, pode-se definir $L(X_t : t \in T)$ como sendo o menor subconjunto de $L_2(\Omega, B, P)$ contendo a família de variáveis aleatórias, que possui as propriedades de um espaço de Hilbert.

Como somente serão de interesse em nosso estudo, processos estocásticos complexos de segunda ordem, sempre que nos referirmos a um processo $(X_t : t \in T)$, este deverá ser considerado, salvo menção em contrário, como um processo estocástico nas condições da definição 1.2.

2. Funções Média e Covariância de um Processo Estocástico

Definição 2.1 :

Para todo $t \in T$, a função valor médio do processo $(X_t : t \in T)$, representada por $m(t)$, é por definição

$$m(t) = E(X_t) \quad (2.2.1)$$

Observemos que para todo $t \in T$, $(X_t - m(t))$ possui média zero e sendo usualmente mais simples se estudar esse processo em lugar do processo (X_t) , suporemos realizada essa substituição de modo que, se nada for mencionado em contrário, estaremos tratando com processos (X_t) tais que $E(X_t) = 0$, para todo $t \in T$.

Definição 2.2 :

A função covariância do processo $(X_t : t \in T)$, representada por $R(s, t)$, é definida para todo (s, t) em $T \times T$ por

$$R(s, t) = \langle X_s, X_t \rangle = E(X_s \cdot \bar{X}_t) \quad (2.2.2)$$

Pode-se verificar (Loève [14]) que:

- i) $R(s, t)$ existe e é finita para todo $(s, t) \in T \times T$
- ii) $R(s, t) = \overline{R(t, s)}$
- iii) $R(t, t) = E(|X_t|^2)$ é real e não negativa
- iv) $R(s, t)$ é uma função definida não negativa no sentido que, dado qualquer conjunto de pontos t_1, t_2, \dots, t_n em T e quaisquer números complexos z_1, z_2, \dots, z_n

$$\begin{aligned} \sum_{j, k=1}^n R(t_j, t_k) z_j \bar{z}_k &= E \left(\sum_{j, k=1}^n X_{t_j} \bar{X}_{t_k} z_j \bar{z}_k \right) = \\ &= E \left(\left| \sum_{j=1}^n X_{t_j} z_j \right|^2 \right) \end{aligned}$$

é sempre real e não negativa. Esta é uma propriedade característica da classe de todas as funções covariância pois, dada qualquer função $R(s, t)$ com essa propriedade, existe sempre um processo estocástico $(X_t : t \in T)$ cuja função covariância é a $R(s, t)$ dada. (Cramér e Leadbetter [8]).

Admitindo-se como convergência natural, a convergência em média quadrática já citada em II.1, pode-se provar (Loève

[14]) que um processo $(X_t : t \in T)$ será contínuo em t_0 se, e somente se, sua função covariância $R(s, t)$ for contínua em (t_0, t_0) ; e ainda, se $R(s, t)$ for contínua na diagonal, será contínua em toda parte.

3. Processos com Incrementos Ortogonais

Definição 3.1 :

Um processo $(Y_t : t \in T)$ será considerado com incrementos ortogonais se

$$E(|Y_s - Y_t|^2) < \infty, \text{ para todo } s, t \text{ em } T$$

e se para quaisquer $s_1 < t_1 \leq s_2 < t_2$ em T ,

$$E[(Y_{t_2} - Y_{s_2})(\overline{Y_{t_1} - Y_{s_1}})] = 0 \quad (2.3.1)$$

A cada processo com incrementos ortogonais corresponde uma função F monótona, não decrescente, determinada de maneira única, a menos de uma constante aditiva, por

$$E(|Y_t - Y_s|^2) = F(t) - F(s), \quad s < t \quad (2.3.2)$$

(ou em uma forma simbólica, $E(|dY_t|^2) = dF(t)$).

As propriedades relativas à continuidade de F determinam as do processo $(Y_t : t \in T)$ da seguinte forma (Doob [9]; Cramér e Leabetter [8]) :

Teorema 3.1 :

Seja $(Y_t : t \in T)$ um processo com incrementos ortogonais onde T é um intervalo real. Seja F a função definida por (2.3.2). Então,

i) para cada $t \in T$, $h > 0$, existem os limites laterais:

$$\lim_{h \rightarrow 0} Y_{t+h} = Y_{t+}$$

$$\lim_{h \rightarrow 0} Y_{t-h} = Y_{t-}$$

ii) $Y_{t-} = Y_t = Y_{t+}$ é válida com probabilidade 1 exceto possivelmente para reais t pertencentes a um subconjunto de T no máximo enumerável.

Faremos uso de certas integrais, mais propriamente de denominadas Integrais Estocásticas, apresentadas como segue:

Seja $(Y_t : t \in T)$ um processo com incrementos ortogonais onde T é um intervalo real. Seja ϕ uma função fixada, com domínio T . Daremos significado ao símbolo:

$$\int_T \phi(t) dY_t \quad (2.3.3)$$

Tomemos ϕ a função escada tal que

$$\phi(t) = \begin{cases} 0 & , \quad \forall t \notin [a_1, a_n) \subset T \\ c_j & , \quad t \in [a_{j-1}, a_j) , \quad j \leq n \end{cases}$$

Então definimos

$$\psi = \int_T \phi(t) dY_t = \sum_2^n c_j [Y_{a_j} - Y_{a_{j-1}}] \quad (2.3.4)$$

Observemos que a classe de equivalência de ψ é univocamente determinada por ϕ e pelo processo, evidentemente; ainda, a cada combinação linear de tais ϕ 's corresponde a mesma combinação linear das correspondentes ψ 's e

$$E \left[\left(\int_T \phi(t) dY_t \right) \overline{\left(\int_T \psi(t) dY_t \right)} \right] = \int_T \phi(t) \overline{\psi(t)} dF(t) \quad (2.3.5)$$

Estabelecemos, portanto, uma correspondência entre certas funções de t (as funções escada ϕ) e variáveis aleatórias ψ . Se forem definidas

$$\| \phi_1 - \phi_2 \| = \left[\int_T |\phi_1(t) - \phi_2(t)|^2 dF(t) \right]^{1/2} \quad (2.3.6)$$

$$\| \psi_1 - \psi_2 \| = [E(|\psi_1 - \psi_2|^2)]^{1/2} \quad (2.3.7)$$

como a distância entre as ϕ 's e a distância entre as ψ 's, respectivamente, por (2.3.5) teremos que distância \bar{e} é preservada pela correspondência.

Seja agora ϕ um limite (no sentido da distância (2.3.6)) de uma sequência $\{\phi_n\}$ de funções escada do tipo definido por (2.3.4). Então,

$$\|\phi - \phi_n\|^2 = \int_T |\phi(t) - \phi_n(t)|^2 dF(t) \rightarrow 0 \quad \text{quando } n \rightarrow \infty$$

$$\text{isto } \bar{e}, \quad \text{l.i.m}_{n \rightarrow \infty} \phi_n = \phi$$

Logo, como \bar{e} é preservada a distância,

$$\text{l.i.m}_{n \rightarrow \infty} \psi_n = \psi$$

Novamente, a classe desta ψ é determinada pelo limite de $\{\phi_n\}$ e não pelas ϕ_n 's. Definimos portanto

$$\int_T \phi(t) dY_t = \psi \quad (2.3.8)$$

A classe das funções ϕ para as quais a integral (2.3.8) é definida, \bar{e} a classe das funções em T que são mensuráveis com respeito à medida dF de Lebesgue-Stieltjes e para as quais

$$\int_T |\phi(t)|^2 dF(t) < \infty$$

4. Processos Estacionários de segunda ordem

Definição 4.1 :

Um processo $(X_t : t \in T)$, genérico, será um processo *estritamente estacionário* se a família de suas distribuições finito-dimensionais for invariante sob uma translação no parâmetro t , isto é, para qualquer sequência finita t_1, t_2, \dots, t_n de T , a distribuição conjunta das variáveis aleatórias $X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h}$ independe de $h \in T$.

Obviamente as características das distribuições conjuntas de um processo estritamente estacionário independem de h e, em particular, os momentos das distribuições, se existirem. Neste caso,

$$E(X_t) = E(X_{t+h}) = m, \quad \forall h, \quad t \in T$$

$$R(s, t) = E(X_s \cdot \bar{X}_t) = E(X_{s+h} \cdot \overline{X_{t+h}}) = E(X_{s-t} \cdot \bar{X}_0) = R(s-t)$$

Definição 4.2 :

Um processo $(X_t : t \in T)$, genérico, será um processo *estacionário no sentido amplo* (estacionário de segunda ordem ou *fracamente estacionário*) se

$$E(|X_t|^2) < \infty \quad \text{para } t \in T \tag{2.4.1}$$

$$E(X_t) = m \quad \text{para todo } t \in T$$

$$R(s,t) = R(s-t) \quad \text{para qualquer } (s,t) \text{ em } T \times T$$

(considerando a observação em II.2, poderemos sempre assumir, sem perda de generalidade, que $m=0$).

Observemos que um processo estritamente estacionário será estacionário no sentido amplo se, para todo $t \in T$, $X_t \in L_2(\Omega, B, P)$.

Serão de nosso interesse apenas os processos estacionários correspondentes à definição 4.2; sendo assim, a estacionariedade à qual nos referiremos será sempre a no sentido amplo.

4-1. Representação e Decomposição Espectral de um Processo Estacionário.

Seja $(X_t : t \in T)$ um processo estacionário contínuo, isto é, a função que a cada t associa X_t é uma função contínua. O conjunto T , se igual a Z , será considerado dotado da métrica discreta, e no caso de T ser um intervalo real, a métrica considerada será a habitual.

Seja ainda $R(u) = E(X_{t+u} \cdot \bar{X}_t)$ a função covariância do processo considerado.

O teorema abaixo melhora o resultado dado em II.3 através da propriedade iv).

Teorema 4.1 (Doob [9]):

Qualquer função $R : T \rightarrow C$ definida positiva, isto é, contínua e satisfazendo

$$R(-u) = \overline{R(u)} \tag{2.4.2}$$

$$\sum_{i,j} R(t_i - t_j) z_i \bar{z}_j \geq 0$$

para cada sequência finita t_1, t_2, \dots, t_n de elementos de T e toda sequência z_1, z_2, \dots, z_n de números complexos, é a função covariância de um processo estacionário. Ainda, se a função covariância for real, o processo poderá ser tomado a valores reais.

O teorema seguinte, devido a Bochner, descreve as funções definidas positivas como sendo transformadas de Fourier.

Teorema 4.2 :

Uma função $R : T \rightarrow C$ será definida positiva quando e somente quando seus valores forem dados por :

$$R(u) = \int_I e^{iu\lambda} dF(\lambda) \quad (2.4.3)$$

onde F é uma função não decrescente e limitada. Ainda, se R for real, então

$$R(u) = \int_J \cos u\lambda dG(\lambda) \quad (2.4.4)$$

onde G é uma função não decrescente e limitada.

Se o processo for a parâmetro discreto os intervalos de integração em (2.4.3) e (2.4.4) serão respectivamente $[-\pi, \pi]$ e $[0, \pi]$; no caso do processo ser a parâmetro contínuo, serão respectivamente $(-\infty, \infty)$ e $[0, \infty)$.

Condições de unicidade para as funções F e G e a demonstração do teorema acima podem ser vistas em Doob [9].

Se a função F for uma função absolutamente contínua, diremos que o processo $(X_t : t \in T)$ tem uma densidade espectral. A derivada de F , representada por $F' = f$, é chamada a função densidade espectral do processo (na forma complexa).

Se $(X_t : t \in T)$ for a valores reais, a função F será absolutamente contínua se e somente se a função G o for. Neste caso, $G' = 2F'$ será a função densidade espectral do processo (forma real).

Em qualquer dos casos, a função F é chamada a fun-

ção distribuição espectral do processo.

Existirá uma função densidade espectral, contínua, dada por:

$$F'(\lambda) = \sum_{u=-\infty}^{\infty} R(u) e^{-i\lambda u} \quad (\text{forma complexa})$$

$$G'(\lambda) = 2R(0) + 4 \sum_{u=1}^{\infty} R(u) \cos \lambda u \quad (\text{forma real})$$

quando o processo for discreto, se a soma $\sum_{u=-\infty}^{\infty} |R(u)|$ for convergente.

De modo análogo, existirá uma função densidade espectral, contínua, dada por :

$$F'(\lambda) = \int_{-\infty}^{\infty} R(u) e^{-i\lambda u} du \quad (\text{caso complexo})$$

$$G'(\lambda) = 4 \int_0^{\infty} R(u) \cos \lambda u du \quad (\text{caso real})$$

quando o processo for a parâmetro contínuo, se $\int_{-\infty}^{\infty} |R(u)| du < \infty$.

O conjunto de todos os valores λ tais que, para qualquer $\epsilon > 0$,

$$F(\lambda + \epsilon) - F(\lambda - \epsilon) > 0$$

é definido como sendo o *espectro* do processo considerado.

O seguinte importante teorema, sob um ponto de vista abstrato, é equivalente a um teorema devido a M.H.Stone [24]

sobre grupos de transformações unitárias em espaços de Hilbert. Sob linguagem probabilística ele é devido a Cramér e independentemente a Loève.

Teorema 4.3 :

Todo processo estacionário contínuo $(X_t : t \in T)$ tem a representação espectral

$$X_t = \int_I e^{it\lambda} dY_\lambda \quad (2.4.5)$$

onde $(Y_\lambda : \lambda \in I)$ é um processo com incrementos ortogonais com $E(|dY_\lambda|^2) = dF(\lambda)$. Se o processo $(X_t : t \in T)$ for a valores reais, a representação espectral será

$$X_t = \int_J [\cos \lambda t dU_\lambda + \text{sen} \lambda t dV_\lambda] \quad (2.4.6)$$

onde $(U_\lambda : \lambda \in J)$ e $(V_\lambda : \lambda \in J)$ são processos reais, com incrementos ortogonais e tais que

$$E[(dU_\lambda)^2] = E[(dV_\lambda)^2] = dG(\lambda), \text{ para } \lambda > 0$$

$$E[(dU_\lambda)^2] = dG(\lambda), \text{ para } \lambda \geq 0$$

$$E(dU_\lambda \cdot dV_\mu) = 0, \text{ para } \lambda, \mu \in J$$

Observação: Sendo de interesse no capítulo IV alguns aspectos práticos que podem ser associados à representação es-

pectral de um processo estacionário, serão feitas aqui algumas considerações a respeito: Se $(X_t : t \in \mathbb{R})$ for observado como representando algum processo físico concreto, por exemplo, flutuações em voltagens elétricas, a representação espectral dada através do teorema 4.3 fornecerá a decomposição da flutuação total em seus componentes harmônicos elementares. A relação $dF(\lambda) = E(|dY_\lambda|^2)$ mostra que $dF(\lambda)$ é a potência média dissipada, através de uma unidade de resistência, pelo componente com frequência no elemento $(\lambda, \lambda+d\lambda)$. A função distribuição espectral F então determina a distribuição da potência média total da flutuação sob a variação da frequência angular λ . A potência média associada ao intervalo de frequência $\lambda_1 < \lambda \leq \lambda_2$ será $F(\lambda_2) - F(\lambda_1)$ que, para o conjunto todo de variação de λ se torna

$$E(|X_t|^2) = R(0) = F(\infty) - F(-\infty)$$

Então a função F determina o *espectro de potência* do processo $(X_t : t \in T)$. Pode-se pensar então disto como uma distribuição de uma massa espectral, cujo valor total é $R(0)$, sobre o conjunto dos λ . E será possível se dizer que λ pertence ao espectro de potência sempre que o intervalo $(\lambda-h, \lambda+h)$ levar uma massa positiva, para qualquer $h > 0$.

Suponhamos agora que A_1, A_2, \dots, A_q formem uma partição mensurável com respeito à medida $\int_{(e)} dF(\lambda)$, do intervalo $(-\infty, \infty)$. Então é possível escrever o processo $(X_t : t \in T)$ como uma soma de processos mutuamente ortogonais $(E(X_{t_i}^{(n)} \overline{X_{t_j}^{(m)}}) = 0$

para quaisquer t_i, t_j em T e $1 \leq n, m \leq q$) cujas distribuições espectrais estão confinadas aos respectivos conjuntos A_1, A_2, \dots, A_q .

Uma decomposição do processo, de interesse, é a correspondente à decomposição da sua função distribuição espectral F , na forma

$$F = F_1 + F_2 + F_3 \quad (2.4.7)$$

onde F_1 é uma função puramente descontínua, isto é, constante exceto em um conjunto no máximo enumerável de pontos de descontinuidade, F_2 é a componente absolutamente contínua de F e F_3 a componente contínua singular de F . Então F_1 é confinada aos pontos onde a F é descontínua, a distribuição F_2 aos pontos de continuidade de F nos quais F' é finita e a distribuição F_3 é restrita ao conjunto de medida nula onde F é contínua e F' ou não existe ou é infinita. Essa decomposição de F gera uma correspondente decomposição do processo em três processos mutuamente ortogonais, isto é,

$$X_t = X_t^{(1)} + X_t^{(2)} + X_t^{(3)} \quad (2.4.8)$$

Se $\lambda_1, \lambda_2, \dots$ forem os pontos de descontinuidade de F (suposta contínua à direita) e se $E\{[dY_\lambda]^2\} = dF(\lambda)$, então

$$X_t^{(1)} = \sum_j e^{it\lambda_j} [Y_{\lambda_j} - Y_{\lambda_j^-}] \quad (2.4.9)$$

onde as variáveis aleatórias $Y_{\lambda_j} - Y_{\lambda_j^-}$ formam um conjunto orto

gonal e para cada $t \in T$ fixo, a série em (2.4.9) converge em média.

4-2. Operações lineares sobre processos estacionários

Se $(X_t : t \in T)$ for um processo estacionário com representação espectral :

$$X_t = \int_I e^{it\lambda} dY_\lambda \quad , \quad E(|dY_\lambda|^2) = dF(\lambda)$$

entenderemos por uma operação linear sobre tal processo como sendo uma transformação que leva $(X_t : t \in T)$ em $(\hat{X}_t : t \in T)$ que é uma soma finita da forma

$$\hat{X}_t = \sum_j \gamma_j X_{t+t_j} = \int_I e^{it\lambda} \left[\sum_j \gamma_j e^{t_j \lambda} \right] dY_\lambda \quad (2.4.10)$$

ou um limite em média de tais somas finitas. A relação acima sugere a seguinte generalização:

$$\hat{X}_t = \int_I e^{it\lambda} \Gamma(\lambda) dY_\lambda$$

onde $\Gamma(\lambda)$ é qualquer função que seja mensurável com respeito a F e para a qual

$$\int_I |\Gamma(\lambda)|^2 dF(\lambda) < \infty$$

A função Γ é chamada *função transferência* da operação. Logo, toda operação linear tem uma função transferência e

toda função transferência determina uma operação linear que define um novo processo estacionário.

O novo processo é também contínuo pois

$$E(|\hat{X}_s - \hat{X}_t|^2) = \int_I |e^{is\lambda} - e^{it\lambda}|^2 |\Gamma(\lambda)|^2 dF(\lambda) \rightarrow 0$$

quando $(s-t) \rightarrow 0$

A função covariância de $(\hat{X}_t : t \in T)$ será dada por

$$\hat{R}(u) = \int_I e^{iu\lambda} |\Gamma(\lambda)|^2 dF(\lambda), \quad (2.4.11)$$

de modo que a função distribuição espectral de $(\hat{X}_t : t \in T)$, será obtida de

$$d\hat{F}(\lambda) = |\Gamma(\lambda)|^2 dF(\lambda) \quad (2.4.12)$$

Se a função transferência da operação for dada por

$$\Gamma(\lambda) = \int_I e^{-i\lambda s} g(s) ds \quad \text{com} \quad \int_I |g(s)| ds < \infty \quad (2.4.13)$$

(e aqui estamos supondo o processo a parâmetro contínuo),

$$\begin{aligned} \hat{X}_t &= \int_I e^{it\lambda} \Gamma(\lambda) dY_\lambda = \int_I e^{it\lambda} \left[\int_I e^{-i\lambda s} g(s) ds \right] dY_t = \\ &= \int_I \left[\int_I e^{i(t-s)\lambda} dY_\lambda \right] g(s) ds = \int_I X_{t-s} g(s) ds \end{aligned} \quad (2.4.14)$$

A operação linear cuja função transferência é da forma (2.4.13)

é denominada *filtragem linear*, nome que provem do fato desta operação fornecer uma descrição matemática para um filtro físico linear cuja entrada e saída estão relacionadas por uma equação da forma (2.4.14). Esta equação exhibe explicitamente a relação linear entre \hat{X}_t e X_t .

Para um processo a parâmetro discreto são válidos os mesmos resultados apresentados através de (2.4.13) e (2.4.14) com as habituais modificações.

4-3. Sobre a estimação da função densidade espectral

Consideraremos agora $(X_t : t \in T)$ um processo estacionário a valores reais e a parâmetro contínuo, tal que:

$$m(t) = E(X_t) = 0 \quad \text{para todo } t$$

(o caso de processos estacionários a parâmetro discreto pode ser discutido de modo semelhante e todas as considerações a serem feitas nesta secção se estendem para o caso em que a função valor médio não é nula).

Assumiremos ainda que a função covariância

$$R(u) = E(X_t \cdot X_{t+u})$$

satisfaz
$$\int_{-\infty}^{\infty} |R(u)| \, du < \infty \quad (2.4.15)$$

Segue então que o processo (X_t) possui uma função

densidade espectral $f(\lambda)$ satisfazendo

$$f(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\lambda u} R(u) du \quad (2.4.16)$$

Em vista da equação (2.4.16) parece natural se pensar em um estimador para $f(\lambda)$ como sendo a transformada de Fourier do estimador da função covariância.

Definição 4.3 :

Dada uma amostra de (X_t) , com $0 \leq t \leq T_0$, a função covariância amostral, $R_{T_0}(u)$, é definida por

$$R_{T_0}(u) = \begin{cases} \frac{1}{T_0} \int_0^{T_0 - |u|} X_t \cdot X_{t+|u|} dt & , \quad |u| < T_0 \\ 0 & , \quad |u| \geq T_0 \end{cases} \quad (2.4.17)$$

e sua transformada de Fourier:

$$\begin{aligned} f_{T_0}(\lambda) &= \frac{1}{2\pi} \int_{-T_0}^{T_0} e^{-i\lambda u} R_{T_0}(u) du = \frac{1}{\pi} \int_0^{T_0} \cos \lambda u R_{T_0}(u) du \\ &= \frac{1}{2\pi T_0} \left| \int_0^{T_0} e^{-it\lambda} X_t dt \right|^2 \end{aligned} \quad (2.4.18)$$

é denominada *função densidade espectral amostral* ou *periodograma*.

Muito embora a função covariância amostral $R_{T_0}(u)$ seja, em cada u , um estimador consistente de $R(u)$, é fato conhecido

do que o periodograma não constitui um estimador consistente para a densidade espectral $f(\lambda)$. As sugestões apresentadas são então de se estimar $f(\lambda)$ em um ponto λ_0 , considerando-se a média dos valores de $f_{T_0}(\lambda)$ em uma vizinhança desse ponto. Entretanto isto leva a um estimador consistente não de $f(\lambda_0)$ mas de uma média espectral na vizinhança de λ_0 , isto é, uma média da forma $\int A(\lambda)f(\lambda)d\lambda$ onde $A(\lambda)$ é uma função escolhida convenientemente (por exemplo, $A(\lambda)$ pode ser escolhida como uma função cujos valores máximos se concentram em torno de λ_0).

Definição 4.4 :

Para toda função contínua e limitada $A(\lambda)$, define-se

$$J(A) = \int_{-\infty}^{\infty} A(\lambda)f(\lambda)d\lambda \quad (2.4.19)$$

como sendo a *média espectral* correspondendo à *janela espectral* $A(\lambda)$.

Será de importância no que se segue a noção de *largura de faixa de uma janela espectral*. Diversas definições para essa medida têm sido sugeridas por diversos autores, todas razoáveis interpretações da noção intuitiva de "largura". A adequada para nosso desenvolvimento é devida a Parzen.

Suporemos $A(\lambda)$ satisfazendo certas condições de regularidade de modo a ter sentido as considerações que se seguem:

Definição 4.5 :

A largura de faixa de uma janela espectral $A(\lambda)$ é definida como o comprimento da base do retângulo que possui mesma área e mesma altura máxima que o gráfico de $A(\lambda)$. Em símbolos,

$$\beta(A) = \frac{\int_{-\infty}^{\infty} A(\lambda) d\lambda}{\max_{\lambda} |A(\lambda)|} \quad (2.4.20)$$

O valor λ_0 tal que $A(\lambda_0) = \max_{\lambda} |A(\lambda)|$ é chamado *pico* da janela espectral $A(\lambda)$.

Parzen [17] mostra que médias espectrais correspondendo a janelas espectrais com largura de faixa $\frac{1}{T_0}$ são quantidades naturais de se desejar estimar e sendo assim, formula o problema da análise espectral empírica como segue:

- Dada uma amostra de um processo (X_t) com $0 \leq t \leq T_0$, formar estimadores ou do valor $f(\lambda_0)$ da função densidade espectral em um dado λ_0 , ou de médias amostrais $J(A)$ correspondentes a janelas espectrais $A(\lambda)$ cujo pico é λ_0 e cuja largura de faixa é da ordem de $\frac{1}{T_0}$.

Grenander e Rosenblatt [12] mostraram que é necessário somente se considerar estimadores da forma:

$$f_{T_0}^*(\lambda_0) = \frac{1}{2\pi} \int_{-T_0}^{T_0} e^{-i\lambda_0 u} k_{T_0}(u) R_{T_0}(u) du \quad (2.4.21)$$

onde as constantes $k_{T_0}(u)$ são escolhidas como uma função par

de u . Estimadores dessa forma podem também ser escritos como médias espectrais amostrais :

$$f_{T_0}^*(\lambda_0) = \int_{-\infty}^{\infty} K_{T_0}(\lambda - \lambda_0) f_{T_0}(\lambda) d\lambda \quad (2.4.22)$$

onde a janela espectral $K_{T_0}(\lambda)$ é definida por :

$$K_{T_0}(\lambda) = \frac{1}{2\pi} \int_{-T_0}^{T_0} e^{i\lambda u} k_{T_0}(u) du \quad (2.4.23)$$

Nota-se que $f_{T_0}^*(\lambda)$ também pode ser escrito como uma média discreta sobre os valores do periodograma nos pontos

$\lambda_m(T_0) = \frac{\pi m}{T_0}$ para $m=0, \pm 1, \pm 2, \dots$, em virtude de

$$f_{T_0}^*(\lambda) = \frac{\pi}{T_0} \sum_{m=-\infty}^{\infty} f_{T_0}(\lambda_m(T_0)) K_{T_0}(\lambda - \lambda_m(T_0)) \quad (2.4.24)$$

Assumindo-se que a janela espectral $K_{T_0}(\lambda)$ atinge seu máximo em $\lambda=0$,

$$\beta(K_{T_0}) = \frac{\int_{-\infty}^{\infty} K_{T_0}(\lambda) d\lambda}{K_{T_0}(0)} = \frac{2\pi k_{T_0}(0)}{\int_{-T_0}^{T_0} k_{T_0}(u) du} \quad (2.4.25)$$

Em muitos casos $k_{T_0}(0)=1$ e nestas circunstâncias, a classe de estimadores da forma (2.4.21) pode ser caracterizada como segue: - Como um estimador $f_{T_0}^*(\lambda_0)$ de uma média espectral $J(A)$ cuja janela espectral $A(\lambda)$ tem pico λ_0 e largura de faixa

da ordem de $\frac{1}{T_0}$, considera-se médias espectrais amostrais

$$J_{T_0}(K_{T_0}(\lambda_0 - \lambda)) = \int_{-\infty}^{\infty} K_{T_0}(\lambda_0 - \lambda) f_{T_0}(\lambda) d\lambda \quad (2.4.26)$$

cuja janela espectral $K_{T_0}(\lambda_0 - \lambda)$ tem pico em λ_0 e largura de faixa da ordem de $\left[\int_{-T_0}^{T_0} k_{T_0}(u) du \right]^{-1}$ onde

$$k_{T_0}(u) = \int_{-\infty}^{\infty} e^{-iu\lambda} K_{T_0}(\lambda) d\lambda \quad (2.4.27)$$

Procura-se determinar dentre os estimadores da forma (2.4.21) aquele que é melhor de acordo com algum critério.

A fim de se especificar um estimador da forma (2.4.21) deve-se estabelecer a função k_{T_0} ; dois métodos são considerados, originando-se duas classes de estimadores que incluem como casos especiais a maioria dos estimadores já sugeridos por diversos autores.

- Seja h uma função limitada, par, quadraticamente integrável, definida para todo real v , tal que $|1-h(v)|/|v|$ seja uma função limitada de v .

Uma classe de estimadores $f_{T_0}^*(\lambda)$ é obtida através de (2.4.21) definindo-se

$$k_{T_0}(u) = h\left(\frac{u}{M_{T_0}}\right) \quad (2.4.28)$$

onde as constantes positivas $M_{T_0} \rightarrow 0$ quando $T_0 \rightarrow \infty$ de tal forma

que $\frac{M_{T_0}}{T_0} \rightarrow 0$. Esta é a classe dos estimadores chamados de *estimadores do tipo algébrico*.

- A segunda classe dos estimadores é obtida através de (2.4.21) definindo-se

$$k_{T_0}(u) = h(A_{T_0} e^{\alpha|u|}) \quad (2.4.29)$$

onde α e A_{T_0} são constantes positivas, sendo que $A_{T_0} \rightarrow 0$ quando $T_0 \rightarrow \infty$ de tal forma que $\frac{\log A_{T_0}}{T_0} \rightarrow 0$. Estes estimadores são chamados *estimadores do tipo exponencial*.

Estas duas classes de estimadores foram introduzidas e suas propriedades assintóticas extensivamente discutidas por Parzen [18]. Os vários estimadores são obtidos de acordo com diferentes escolhas do "kernel" h . Como exemplos,

O estimador sugerido por Bartlett corresponde à escolha

$$h(v) = \begin{cases} 1 - |v| & |v| \leq 1 \\ 0 & |v| \geq 1 \end{cases} \quad (2.4.30)$$

O estimador, usualmente chamado de *periodograma truncado* corresponde ao kernel

$$h(v) = \begin{cases} 1 & |v| \leq 1 \\ 0 & |v| > 1 \end{cases} \quad (2.4.31)$$

Parzen sugeriu kernels da forma:

$$h(v) = \begin{cases} 1 - |v|^q & , \quad |v| \leq 1 \\ 0 & , \quad |v| > 1, \quad q \geq 1. \end{cases} \quad (2.4.32)$$

O estimador de Daniel corresponde ao uso do kernel

$$h(v) = \frac{\text{sen } v}{v} \quad (2.4.33)$$

Tukey sugeriu o estimador correspondente ao kernel

$$h(v) = \begin{cases} \frac{1}{2} (1 + \cos \pi v) & , \quad |v| \leq 1 \\ 0 & |v| \geq 1 \end{cases} \quad (2.4.34)$$

Outras possíveis escolhas do kernel h podem ser obtidas em Parzen [18].

A fim de escolher um estimador $f_{T_0}^*(\lambda)$, deve-se observar que sendo $f(\lambda)$ uma função e não simplesmente um parâmetro, existem conseqüentemente vários critérios possíveis que podem ser usados para julgar o mérito dos diversos estimadores. Aqueles que têm sido sugeridos até agora podem ser sumarizados como segue:

i) critério do erro quadrático médio integrado -

Lomnicki e Zaremba propuseram como uma "figura de mérito", o erro quadrático médio integrado definido por

$$\int_{-\infty}^{\infty} E [| f_{T_0}^*(\lambda) - f(\lambda) |^2] d\lambda = \int_{-\infty}^{\infty} \text{Var}(f_{T_0}^*(\lambda)) d\lambda + \int_{-\infty}^{\infty} v^2 [f_{T_0}^*(\lambda)] d\lambda \quad (2.4.35)$$

onde
$$\text{Var}(f_{T_0}^*(\lambda)) = E |f_{T_0}^*(\lambda) - E(f_{T_0}^*(\lambda))|^2$$

e
$$v(f_{T_0}^*(\lambda)) = E(f_{T_0}^*(\lambda)) - f(\lambda)$$
 (2.4.36)

são respectivamente a variância e o vício do estimador $f_{T_0}^*(\lambda)$. Este critério tem sido criticado pelo fato de dar pesos iguais a todos os λ ; no entanto é o único até agora que fornece uma forma explícita para o desejado estimador "ótimo".

ii) critério do máximo erro quadrático médio -

Parzen sugeriu a medida

$$E \left[\sup_{\lambda} |f_{T_0}^*(\lambda) - f(\lambda)|^2 \right] \quad (2.4.37)$$

mencionando que é a quantidade apropriada a se considerar ao se construir faixas de confiança para a função densidade espectral.

iii) critério do erro quadrático médio em um particular valor λ -

Parzen propôs como "figura de mérito" o erro quadrático médio

$$E \left[|f_{T_0}^*(\lambda) - f(\lambda)|^2 \right] = \text{Var}(f_{T_0}^*(\lambda)) + v^2(f_{T_0}^*(\lambda)) \quad (2.4.38)$$

em cada λ . Este critério é criticado pelo fato de não se ter usualmente interesse em estimar $f(\lambda)$ em um particular valor λ . No entanto como o desvio padrão e o vício de $f(\lambda)$ são ambos em geral assintoticamente proporcionais a $f(\lambda)$, o erro quadrático

médio relativo, definido por

$$E \left[\left| f_{T_0}^* (\lambda) - f(\lambda) \right|^2 / f^2(\lambda) \right] \quad (2.4.39)$$

é aproximadamente assintoticamente independente de λ .

Uma outra forma de se atacar o problema da estimação da função densidade espectral é utilizar uma técnica de filtragem obtendo-se um novo processo (\hat{X}_t) relacionado com (X_t) através de

$$\hat{X}_t = \int_0^t X_s g(t-s) ds, \quad 0 < t < T_0 \quad (2.4.40)$$

onde a função transferência da operação é do tipo

$$\Gamma(\lambda) = \begin{cases} 1 & , \quad |\lambda - \lambda_0| \leq \delta \\ 0 & , \quad |\lambda - \lambda_0| > \delta \end{cases} \quad (2.4.41)$$

O estimador para $f(\lambda_0)$ é então obtido calculando-se a variância (isto é, a potência total) do processo filtrado (\hat{X}_t), dada por

$$\frac{1}{T_0} \int_0^{T_0} \hat{X}_t^2 dt \quad (2.4.42)$$

É resultado conhecido (Grenander e Rosenblatt [12]) que as formas (2.4.21) e (2.4.42) para os estimadores são formas teoricamente equivalentes (assintoticamente) visto que $|\Gamma(\lambda)|^2$, desempenha a mesma função que a janela espectral considerada anteriormente.

Finalizando as considerações sobre a estimação da função densidade espectral, seguem representações para a média, variância e vício de um estimador da forma de (2.4.21); estes resultados são apresentados com detalhes em Parzen [18].

Considerando-se, novamente, o estimador

$$f_{T_0}^*(\lambda) = \frac{1}{2\pi} \int_{-T_0}^{T_0} e^{-i\lambda u} k_{T_0}(u) R_{T_0}(u) du,$$

pode-se verificar que:

i) a esperança de $f_{T_0}^*(\lambda)$ é dada por

$$E[f_{T_0}^*(\lambda)] = \frac{1}{2\pi} \int_{-T_0}^{T_0} e^{-i\lambda u} k_{T_0}(u) R(u) \left(1 - \frac{|u|}{T_0}\right) du$$

ii) a variância de $f_{T_0}^*(\lambda)$ é obtida como

$$\text{Var}[f_{T_0}^*(\lambda)] = \int [f_{T_0}^*(\lambda) - E(f_{T_0}^*(\lambda))]^2 d\lambda = \frac{1}{2\pi} \int_{-T_0}^{T_0} k_{T_0}^2(u) [R_{T_0}(u) - E(R_{T_0}(u))]^2 du$$

onde $R_{T_0}(u)$ é definida por (2.4.17)

iii) o vício do estimador $f_{T_0}^*(\lambda)$, conseqüentemente,

é dado por

$$v[f_{T_0}(\lambda)] = E[f_{T_0}^*(\lambda)] - f(\lambda)$$

$$= \frac{1}{2\pi} \int_{-T_0}^{T_0} e^{-i\lambda u} [k_{T_0}(u) - 1] R(u) du - \frac{1}{2\pi T_0} \int_{-T_0}^{T_0} e^{-i\lambda u} k_{T_0}(u) |u| R(u) du - \frac{1}{2\pi} \int_{|u| \geq T_0} e^{-i\lambda u} R(u) du$$

III - SOBRE PROCESSOS NÃO ESTACIONÁRIOS

1. Considerações gerais

Seja $L_2 = L_2(\Omega, B, P)$ o espaço de Hilbert definido em II.1 com produto interno dado por :

$$\langle X, Y \rangle = E(X \cdot \bar{Y})$$

e conseqüente norma

$$\| X \| = [E(|X|^2)]^{1/2}$$

Consideraremos processos vetoriais de segunda ordem, contínuos, a valores complexos, $X_t = (X_t^1, X_t^2, \dots, X_t^q)'$, onde cada $(X_t^j : t \in T)$ é um processo estocástico complexo, contínuo, de segunda ordem.

Assumiremos o conjunto T como igual a R ou Z , e o processo $(X_t : t \in T)$ tal que, para todo $j=1, 2, \dots, q$,

$$\| X_t^j \| = [E(|X_t^j|^2)]^{1/2} > 0 \text{ para pelo menos um índice } t$$

e

$$E(X_t^j) = 0 \text{ para todo } t \in T \quad (3.1.1)$$

Definição 1.1 :

Dois processos vetoriais (X_t) e (Y_t) serão ortogonais se para qualquer $(s, t) \in T^2$ e qualquer $(j, k) \in [1, q]^2$

$$\langle X_s^j, Y_t^k \rangle = 0$$

Em L_2 , o subespaço linear fechado gerado pelo conjunto $\{X_s^j : j=1,2,\dots,q \text{ e } s \leq t\}$ será denotado por $\mathcal{L}(X,t)$. Ainda,

$$\mathcal{L}(X,\infty) = \overline{\bigcup_{t \in T} \mathcal{L}(X,t)} = \mathcal{L}(X)$$

$$\mathcal{L}(X,-\infty) = \bigcap_{t \in T} \mathcal{L}(X,t)$$

Logo, para $t \leq t'$, são imediatas as inclusões

$$0 \subset \mathcal{L}(X,-\infty) \subset \mathcal{L}(X,t) \subset \mathcal{L}(X,t') \subset \mathcal{L}(X) \subset L_2 \quad (3.1.2)$$

Se Y for de L_2 , será representada por $\pi_t Y$ a projeção de Y sobre o subespaço $\mathcal{L}(X,t)$. Do mesmo modo, se Y^1, \dots, Y^q for uma seqüência de elementos de L_2 , $\pi_t Y$ representará o vetor coluna $(\pi_t Y^1 \dots \pi_t Y^q)'$ onde $Y = (Y^1 \dots Y^q)'$.

Observemos que dentre todos os elementos Z do subespaço $\mathcal{L}(X,t-h)$, $h \geq 0$, a projeção $\pi_{t-h} X_t^j$ é aquele que minimiza $\|X_t^j - Z\|$. Com a indicação:

$$\sigma_{th}^j = \|X_t^j - \pi_{t-h} X_t^j\|, \quad (3.1.3)$$

por (3.1.1), temos $0 \leq \sigma_{th}^j \leq \sigma_{tk}^j$ para $0 < h < k$.

(Sob o ponto de vista da teoria da predição linear de mínimos quadrados, a projeção $\pi_{t-h} X_t^j$ é a melhor predição possível de X_t^j em termos de todas as variáveis X_s^1, \dots, X_s^q com $s \leq t-h$. O correspondente erro de predição é então dado por (3.1.3)).

Analisaremos agora na relação (3.1.2) duas situações

que serão de bastante interesse, a saber:

$$\begin{aligned} \text{A)} \quad \mathcal{L}(X, -\infty) &= \mathcal{L}(X, \infty) \\ \text{B)} \quad \mathcal{L}(X, -\infty) &= \mathbf{0} \end{aligned} \tag{3.1.4}$$

Em A), $\mathcal{L}(X, -\infty) = \mathcal{L}(X, t)$ para todo t ; portanto $\mathcal{L}(X, t) = \mathcal{L}(X, t-h)$ para cada h e $\sigma_{th}^j = 0$ para $j=1, 2, \dots, q$ e todo $t \in T$.

Um processo (X_t) em que a situação A) for verificada, será chamado um processo *determinístico*; caso A) não se verifique, (X_t) será um processo *não determinístico* e para estes, se for satisfeita a relação B), será reservada a denominação *não determinístico puro*.

Para um processo vetorial de segunda ordem a valores complexos, existe uma decomposição, definida de maneira única, em termos de um componente não determinístico puro e um componente determinístico, mutuamente ortogonais. Essa decomposição forma a base de uma análise no domínio T de um dado processo estocástico, generalizando as propriedades dos processos estacionários.

A análise dessa decomposição será feita a seguir, separadamente, para processos vetoriais a parâmetro discreto e a parâmetro contínuo. Para o caso discreto, as propriedades dessa decomposição são uma generalização direta da conhecida decomposição de Wold para processos estacionários univariados a parâmetro discreto.

2. O caso discreto

Seja $(X_t : t \in T)$ um processo vetorial a parâmetro discreto, nas condições estabelecidas em III.1. Consideremos a projeção de (X_t) no subespaço $\mathcal{L}(X, t-1)$, isto é,

$$\pi_{t-1} X_t = (\pi_{t-1} X_t^1 \dots \pi_{t-1} X_t^q)'$$

Definição 2.1 :

O processo $\xi_t = (\xi_t^1 \dots \xi_t^q)'$ onde, para $j=1, \dots, q$ e $t \in T$,

$$\xi_t^j = X_t^j - \pi_{t-1} X_t^j \quad (3.2.1)$$

é por definição o *processo inovação* de (X_t) .

Essa denominação para (ξ_t) é motivada no fato de $\mathcal{L}(X, t-1) \neq \mathcal{L}(X, t)$ se e só se $\xi_t = (\xi_t^1 \dots \xi_t^q)' \neq (0 \dots 0)'$.

O conjunto $\Theta = \{t \in T : \xi_t \neq (0 \dots 0)'\}$ é chamado o *espectro inovação* de (X_t) .

Evidentemente, um processo será determinístico se e somente se $\Theta = \emptyset$. Se (X_t) for um processo estacionário não determinístico então $\Theta = Z$. Vamos mostrar que esta afirmação é verdadeira para um processo (X_t) univariado. A generalização para o caso multivariado, embora trabalhosa, segue linhas semelhantes.

Seja portanto (X_t) um processo a parâmetro discreto, univariado e Θ o seu espectro inovação. Suponhamos (X_t) estacio

nário. Então se $\Theta \neq \emptyset$ vale que $\Theta = Z$.

Como $\Theta \neq \emptyset$, $\Theta \subset Z$, existe $t_0 \in \Theta$ tal que $X_{t_0} \notin \mathcal{L}(X, t_0 - 1)$. Sendo $\mathcal{L}(X, t_0 - 1)$ fechado, a distância $d(X_{t_0}, \mathcal{L}(X, t_0 - 1)) = \varepsilon_0^{1/2} > 0$. Ou, de forma equivalente, (1) $\|X_{t_0} - \sum_i \alpha_i X_{t_i}\|^2 \geq \varepsilon_0$ para qualquer combinação finita $\sum_i \alpha_i X_{t_i}$, $\forall i, t_i < t_0$.

Seja agora $t' \in Z$; mostremos que $X_{t'} \notin \mathcal{L}(X, t' - 1)$. Isto é o mesmo que mostrar que $\exists \varepsilon_0 > 0$ tal que, para qualquer combinação $\sum_j \alpha_j X_{t_j}$ vale

$$\begin{aligned} & \|X_{t'} - \sum_j \alpha_j X_{t_j}\|^2 \geq \varepsilon_0 \quad \text{sendo } t_j < t' \\ & \|X_{t'} - \sum_j \alpha_j X_{t_j}\|^2 = \langle X_{t'}, X_{t'} \rangle - \sum_j \alpha_j [\langle X_{t'}, X_{t_j} \rangle + \langle X_{t_j}, X_{t'} \rangle] \\ & + \sum_l \sum_k \alpha_l \alpha_k \langle X_{t_l}, X_{t_k} \rangle \\ & = \langle X_{t_0}, X_{t_0} \rangle - \sum_j \alpha_j [\langle X_{t_0}, X_{t_0 - (t' - t_j)} \rangle + \langle X_{t_0 - (t' - t_j)}, X_{t_0} \rangle] \\ & + \sum_l \sum_k \alpha_l \alpha_k \langle X_{t_0 - (t' - t_l)}, X_{t_0 - (t' - t_k)} \rangle \\ & = \|X_{t_0} - \sum_j \alpha_j X_{t_0 - (t' - t_j)}\|^2 \geq \varepsilon_0 \quad \text{por (1)}. \end{aligned}$$

Então, $Z \subset \Theta$. Portanto, $\Theta = Z$.

Seja agora t um inteiro dado e formemos a matriz covariância $R(t) = \langle \xi_t^j, \xi_t^k \rangle$ das q variáveis ξ_t^j , $j=1, 2, \dots, q$. O rank de $R(t)$ é o número máximo, r_t , de variáveis linearmente independentes entre as ξ_t^1, \dots, ξ_t^q . Claramente $0 \leq r_t \leq q$ e $r_t > 0$ se

e somente se $t \in \mathbb{0}$.

Seja $\xi(X,t)$ o subespaço r_t -dimensional (portanto fechado) gerado pelas variáveis ξ_t^j , $j=1,2,\dots,q$. Por (3.2.1) podemos afirmar que $\xi(X,t)$ é um subespaço de $\mathcal{L}(X,t)$ e mais ainda, ξ_t^j é ortogonal a $\mathcal{L}(X,t-1)$; conseqüentemente $\xi(X,t)$ é ortogonal a $\mathcal{L}(X,t-1)$. Pelo mesmo fato segue também que ξ_s^j e ξ_t^k são variáveis ortogonais, se $s \neq t$, $j=1,2,\dots,q$, $k=1,\dots,q$. Logo, vale que $\xi(X,s)$ é ortogonal a $\xi(X,t)$ sempre que $s \neq t$.

Pelo processo de ortonormalização das variáveis ξ_t^1, \dots, ξ_t^q , obtemos um conjunto de r_t variáveis $\eta_t^1, \dots, \eta_t^{r_t}$, linearmente independentes (sistema ortonormal completo em $\xi(X,t)$) e mais $q-r_t$ variáveis nulas.

Consideremos o subespaço-soma da família ortogonal de subespaços $\xi(X,s)$ de $\mathcal{L}(X,t)$ com $s \leq t$, ou seja,

$$\bigoplus_{s \leq t} \xi(X,s) = \xi(X,t) \oplus \xi(X,t-1) \oplus \dots \oplus \xi(X,t-s) \oplus \dots \quad (3.2.2)$$

$\bigoplus_{s \leq t} \xi(X,s)$ é portanto o subespaço gerado por todas as variáveis ξ_s^j com $j=1,\dots,q$ e $s \leq t$. Sob a notação precedente podemos escrever $\bigoplus_{s \leq t} \xi(X,s) = \mathcal{L}(\xi,t)$. Ainda, o conjunto de todas as variáveis η_s^j onde $j=1,2,\dots,r_s$ e $s \leq t$ constitui um sistema ortonormal completo em $\mathcal{L}(\xi,t)$.

Lema

Para cada $t \in \mathbb{Z}$, $\mathcal{L}(X,t)$ é a soma ortogonal de $\mathcal{L}(\xi,t)$ e

$\mathcal{L}(X, -\infty)$, ou seja,

$$\mathcal{L}(X, t) = \mathcal{L}(\xi, t) \oplus \mathcal{L}(X, -\infty) \quad (3.2.3)$$

Prova

Como $\xi(X, s)$ é ortogonal a $\mathcal{L}(X, s-1)$ e este último contém $\mathcal{L}(X, -\infty)$, segue que $\xi(X, s)$ é ortogonal a $\mathcal{L}(X, -\infty)$ para todo s . Logo, $\mathcal{L}(\xi, t) = \bigoplus_{s \leq t} \xi(X, s)$ será ortogonal a $\mathcal{L}(X, -\infty)$. Disto obtemos

$$\mathcal{L}(\xi, t) \oplus \mathcal{L}(X, -\infty) \subset \mathcal{L}(X, t) ,$$

visto que $\mathcal{L}(\xi, t)$ e $\mathcal{L}(X, -\infty)$ são subespaços de $\mathcal{L}(X, t)$.

Por outro lado, cada elemento V de $\mathcal{L}(X, t)$ é o limite de uma seqüência de variáveis, cada uma das quais sendo a soma de uma combinação linear das ξ_t^j e um elemento de $\mathcal{L}(X, t-1)$. (Note que $X_t^j = \xi_t^j + \pi_{t-1} X_t^j$). Assim, segue que $V = V_1 + V_2$ onde $V_1 \in \mathcal{L}(\xi, t)$ e $V_2 \in \mathcal{L}(X, t-1)$. Logo,

$$\mathcal{L}(X, t) \subset \mathcal{L}(\xi, t) \oplus \mathcal{L}(X, t-1) \subset \mathcal{L}(\xi, t) \oplus \mathcal{L}(X, t-1)$$

De repetidas aplicações dessa relação obtemos

$$\mathcal{L}(X, t) \subset \mathcal{L}(\xi, t) \oplus \mathcal{L}(X, t-p) \quad \text{para todo } p > 0.$$

Disto obtemos finalmente que $\mathcal{L}(X, t) \subset \mathcal{L}(\xi, t) \oplus \mathcal{L}(X, -\infty)$, o que conclui a prova.

Pode-se provar agora o análogo da decomposição de Wold para o processo $(X_t : t \in \mathbb{Z})$ generalizando o teorema correspondente de Wiener e Masani.

Teorema 2.1 :

Para cada processo vetorial discreto (X_t) existe um par de processos, (U_t) e (V_t) , satisfazendo :

(a) $X_t = U_t + V_t$ sendo que U_t^j e V_t^j , $j=1,2,\dots,q$ pertencem a $\mathcal{L}(X,t)$.

(b) (U_t) e (V_t) são processos ortogonais sendo (U_t) não determinístico puro e (V_t) determinístico. Desta forma, são únicos.

(c) $U_t = \sum_{s \leq t} A(t,s) \xi_s$ onde $A(t,s) = (a(t,s))^{(jk)}$ é uma matriz $q \times q$ tal que U_t^j é obtido como sendo a soma

$$\sum_{s \leq t} \sum_{1 \leq k \leq q} a(t,s)^{(jk)} \xi_s^k \text{ em } L_2 .$$

(d) Escrevendo $c(t,s)^{(j)} = \left\| \sum_{k=1}^q a(t,s)^{(jk)} \xi_s^k \right\| =$

$$\left[E \left(\left| \sum_{k=1}^q a(t,s)^{(jk)} \xi_s^k \right|^2 \right) \right]^{1/2} \text{ temos que } \sum_{s \leq t} (c(t,s)^{(j)})^2 < \infty$$

para todo j e todo t . Os coeficientes $a(t,s)^{(j,k)}$ serão determinados de maneira única se e somente se $r_s = q$, enquanto que $c(t,s)^{(j)}$ são únicos para todo t,s e j .

Prova

Para todo t e para $j=1,2,\dots,q$ sejam U_t^j e V_t^j as projeções de X_t^j nos subespaços $\mathcal{L}(\xi,t)$ e $\mathcal{L}(X,-\infty)$, respectivamente. Pelo lema anterior,

$$\mathcal{L}(X,t) = \mathcal{L}(\xi,t) \oplus \mathcal{L}(X,-\infty)$$

Logo, $X_t = U_t + V_t$, ou seja, para $j=1,2,\dots,q$ e $t \in T$,

$X_t^j = U_t^j + V_t^j$. Como para todo (j,k) e t , $\langle U_t^j, V_t^k \rangle = 0$, os processos (U_t) e (V_t) são ortogonais.

Consideremos, como em notações precedentes, os subespaços $\mathcal{L}(U,t)$ e $\mathcal{L}(V,t)$, observando que $\mathcal{L}(U,t)$ é ortogonal a $\mathcal{L}(V,s)$ para todo par de inteiros t e s . Como $X_t = U_t + V_t$, temos pelo lema anterior que

$$\mathcal{L}(X,t) = \mathcal{L}(\xi,t) \oplus \mathcal{L}(X,-\infty) = \mathcal{L}(U,t) \oplus \mathcal{L}(V,t)$$

Da definição de (U_t) e (V_t) obtemos

$$\mathcal{L}(U,t) = \mathcal{L}(\xi,t) \tag{3.2.4}$$

$$\mathcal{L}(V,t) = \mathcal{L}(X,-\infty) \tag{3.2.5}$$

Pelo mesmo lema e por (3.2.4) temos que para $t \rightarrow -\infty$,

$$\mathcal{L}(U,-\infty) = \mathcal{L}(\xi,-\infty) = 0$$

Portanto, (U_t) é um processo não determinístico puro.

Por outro lado,

$\mathcal{L}(V,t) = \mathcal{L}(X,-\infty) = \mathcal{L}(V,-\infty)$ para cada t e consequentemente (V_t) é um processo determinístico.

Tratemos agora da prova relativa à unicidade da decomposição.

Suponhamos (W_t) e (Z_t) quaisquer processos satisfazendo

$$X_t = W_t + Z_t \quad \text{onde} \quad X_t^j = W_t^j + Z_t^j, \quad \text{cada } t \text{ e } j \text{ e tendo}$$

as propriedades (a) e (b) afirmadas no teorema. Então são válidas ainda as relações :

$$\mathcal{L}(X, t) = \mathcal{L}(W, t) \oplus \mathcal{L}(Z, t) \quad \text{para todo } t \quad (3.2.6)$$

$$\mathcal{L}(X, -\infty) = \mathcal{L}(W, -\infty) \oplus \mathcal{L}(Z, -\infty) \quad (3.2.7)$$

Considerando a propriedade (b), obtemos $\mathcal{L}(W, -\infty) = \mathbb{0}$ e portanto

$$\mathcal{L}(X, -\infty) = \mathcal{L}(W, -\infty) = \mathcal{L}(W, t) \quad \text{para todo } t.$$

Pelo lema e por (3.2.6) vemos que ainda são válidas as relações:

$$\mathcal{L}(W, t) = \mathcal{L}(\xi, t)$$

e

$$\mathcal{L}(Z, t) = \mathcal{L}(X, -\infty)$$

Na decomposição $X_t^j = W_t^j + Z_t^j$, a componente W_t^j é então a projeção de X_t^j sobre $\mathcal{L}(\xi, t)$ e a componente Z_t^j , a projeção de X_t^j sobre $\mathcal{L}(X, -\infty)$.

Finalmente para estabelecermos a validade da propriedade (c) usaremos o fato daquelas variáveis η_s^j constituírem um sistema ortonormal completo em $\mathcal{L}(\xi, t)$. Cada elemento U_t^j de $\mathcal{L}(\xi, t)$ tem então seu correspondente desenvolvimento de Fourier, a saber

$$U_t^j = \sum_{s \leq t} \sum_{k=1}^{r_s} b^{(jk)}(t, s) \eta_s^k \quad (3.2.8)$$

sendo $\sum_{s \leq t} \sum_{k=1}^{r_s} |b^{(jk)}(t, s)|^2 < \infty$

Para cada s , as variáveis η_s^j , $1 \leq j \leq r_s$, são certas combinações lineares das componentes ξ_s^k , $1 \leq k \leq q$. Logo, os coeficientes dessas combinações serão univocamente determinados se e somente se $r_s = q$.

Substituindo em (3.2.8) as variáveis η_s^k por suas expressões em termos das ξ_s^k obtemos o desenvolvimento dado sob (c) e ainda

$$(C^j(t, s))^2 = \sum_{k=1}^{r_s} |b^{(jk)}(t, s)|^2$$

o que completa a prova.

Como aplicação deste teorema, vejamos o resultado abaixo que tem uma interpretação natural na teoria da predição :

Teorema 2.2 :

Seja h qualquer inteiro positivo e (X_t) um processo nas condições do teorema 2.1. A melhor predição da componente X_t^j em termos de todas as variáveis X_s^j , $j=1, 2, \dots, q$, $s \leq t-h$, será dada por

$$\pi_{t-h} X_t^j = \sum_{s=-\infty}^{t-h} \sum_{k=1}^q a^{(jk)}(t, s) \xi_s^k + v_t^j$$

com o correspondente erro de predição

$$\sigma_{th}^j = \|X_t^j - \pi_{t-h} X_t^j\| = \left[\sum_{s=t-h+1}^t (C^j(t, s))^2 \right]^{1/2}$$

Se observarmos que v_s^j , assim como todas as ξ_s^k , com

$s \leq t-h$, pertencem a $\mathcal{L}(X, t-h)$, enquanto que todas as ξ_s^k , com $s > t-h$, são ortogonais a esse espaço, a prova deste teorema segue diretamente do teorema 3.1 e da definição de σ_{th}^j .

Deve ser notado que os coeficientes $a_{(t,s)}^{(jk)}$ na expressão para a melhor predição de X_t^j dada no teorema acima dependem de certas covariâncias do processo (X_t) até o ponto t . Em concordância, qualquer estimativa estatística desta predição por meio deste teorema deve ser baseada em informação concernente à estrutura da covariância do processo até o ponto t , ou a partir de um conhecimento a priori (como no caso em que o processo é suposto estacionário), ou a partir de prévias experiências estatísticas.

3. O caso contínuo

Consideremos agora um processo vetorial

$$X_t = (X_t^1 \dots X_t^q)'$$

nas condições estabelecidas em III.1, onde o conjunto dos parâmetros, T , é o conjunto dos números reais.

Como em III.2, $\mathcal{L}(X, t)$ será o subespaço linear fechado de L_2 , gerado pelo conjunto $\{X_s^j : j=1, 2, \dots, q \text{ e } s \leq t\}$ e

$$\mathcal{L}(X, \infty) = \overline{\bigcup_{t \in T} \mathcal{L}(X, t)} = \mathcal{L}(X)$$

$$\mathcal{L}(X, -\infty) = \bigcap_{t \in T} \mathcal{L}(X, t)$$

Assim como para processos vetoriais a parâmetro discreto, é possível se estabelecer um teorema (teorema 3.1 a seguir) que garanta a existência de uma decomposição e representação para processos vetoriais a parâmetro contínuo, satisfazendo certas condições. A primeira parte desse teorema é diretamente análoga à do teorema 2.1 e pode ser provada usando argumentos semelhantes. No entanto, a segunda parte, relativa à representação do componente não determinístico do processo (X_t) , envolve circunstâncias mais complicadas do que no caso discreto, de modo que nos restringiremos apenas em afirmar os resultados principais e fazer algumas discussões preliminares.

Para nossos propósitos admitiremos que o processo (X_t) seja não determinístico puro, isto é, $\mathcal{L}(X, -\infty) = \emptyset$.

Se $(t, t+h)$ e $(s, s+k)$ forem intervalos disjuntos, os complementos ortogonais de $\mathcal{L}(X, t)$ em $\mathcal{L}(X, t+h)$ e de $\mathcal{L}(X, s)$ em $\mathcal{L}(X, s+k)$ serão denotados respectivamente por

$$\mathcal{L}(X, t+h) \ominus \mathcal{L}(X, t) \quad \text{e} \quad \mathcal{L}(X, s+k) \ominus \mathcal{L}(X, s)$$

Observemos que estes últimos são mutuamente ortogonais. (3.3.0)

Definição 3.1 :

Seja $\Theta = \{t \in T : \forall h > 0 \mathcal{L}(X, t+h) \ominus \mathcal{L}(X, t-h) \neq \emptyset\}$

O conjunto Θ é chamado o *espectro inovação* do processo (X_t) .

Como em III.2 denotaremos por π_t a projeção de $\mathcal{L}(X, +\infty)$ em $\mathcal{L}(X, t)$. Obtemos então uma família (π_t) de projeções com domí

nio $\mathcal{L}(X, +\infty)$ sendo que $\pi_{-\infty} = 0$ e $\pi_{+\infty} = I$. Notemos que $\pi_{t+h}^{-1}\pi_t$ é a projeção de $\mathcal{L}(X, +\infty)$ sobre $\mathcal{L}(X, t+h) \ominus \mathcal{L}(X, t)$.

Cada variável aleatória Z de $\mathcal{L}(X, +\infty)$ dá origem a um novo processo unidimensional (Z_t) definido, para todo t , por

$$Z_t = \pi_t Z \tag{3.3.1}$$

Segue que $Z_t \in \mathcal{L}(X, t)$ e ainda, (Z_t) é um processo com incrementos ortogonais (veja-se observação 3.3.0) a valores complexos com as propriedades:

$$Z_{-\infty} = \pi_{-\infty} Z = 0$$

$$Z_{\infty} = \pi_{\infty} Z = Z$$

$$E(Z_t) = 0$$

$$E(|Z_t|^2) = F_Z(t)$$

onde para cada Z , $F_Z(t)$ é uma função limitada, não decrescente e contínua (pois (X_t) é contínuo), tal que

$$F_Z(-\infty) = \lim_{t \rightarrow -\infty} F_Z(t) = 0$$

$$F_Z(+\infty) = \lim_{t \rightarrow \infty} F_Z(t) = E(|Z|^2)$$

Os pontos t tais que, para qualquer $h > 0$,

$$E(|Z_{t+h} - Z_{t-h}|^2) = F_Z(t+h) - F_Z(t-h) > 0$$

formam um subconjunto do espectro inovação de (X_t) . Qualquer incremento $Z_{t+h} - Z_t$ pertence ao subconjunto $\mathcal{L}(X, t+h) \ominus \mathcal{L}(X, t)$ e pode então ser observado como uma parte da inovação recebida pelo processo (X_t) durante o intervalo $(t, t+h)$.

Por (3.3.1), observemos que $\mathcal{L}(Z, t) \subset \mathcal{L}(X, t)$.

Seja $\mathcal{L}(Z, \infty) = \overline{\cup \mathcal{L}(Z, t)}$. Segue portanto que $\mathcal{L}(Z, t)$ é a projeção de $\mathcal{L}(Z, \infty)$ sobre $\mathcal{L}(X, t)$.

Consideremos agora $\mathcal{L}^*(Z, t)$ o conjunto de todas as variáveis W representáveis na forma

$$W = \int_{-\infty}^t g(s) dZ_s \text{ onde } g \text{ é uma função } F_Z\text{-mensurável e}$$

$$E(|W|^2) = \int_{-\infty}^t |g(\lambda)|^2 dF_Z(\lambda) < \infty$$

Então, de acordo com os resultados apresentados no capítulo II, item 3, sobre integrais estocásticas, segue que

$$\mathcal{L}^*(Z, t) = \mathcal{L}(Z, t) \tag{3.3.2}$$

Com base na teoria da multiplicidade espectral em espaços de Hilbert (ver Halmos [13]) temos que :

- Existe uma sequência finita ou infinita de elementos não nulos Z^1, Z^2, \dots de $\mathcal{L}(X, \infty)$ tal que para todo $t, t \leq \infty$ valem:

$$(a) \mathcal{L}(Z^j, t) \text{ é ortogonal a } \mathcal{L}(Z^i, t) \text{ para } i \neq j \tag{3.3.3}$$

$$(b) \mathcal{L}(X, t) = \mathcal{L}(Z^1, t) \oplus \mathcal{L}(Z^2, t) \oplus \dots$$

Por (3.3.2) e (3.3.3) temos então que, para cada $j=1, 2, \dots, q$ e $t \in \mathbb{R}$

$$X_t^j = \sum_{n=1}^N \int_{-\infty}^t g_{jn}(t, \lambda) dZ_\lambda^n \tag{3.3.4}$$

Para $N=\infty$, a s\u00e9rie no segundo membro converge em m\u00e9dia;

portanto

$$\sum_{n=1}^{\infty} \int_{-\infty}^t |g_{jn}(t, \lambda)|^2 dF_{Z^n}(\lambda) < \infty$$

Se definirmos o processo vetorial N -dimensional,

$$Z_{\lambda} = (Z_{\lambda}^1 \dots Z_{\lambda}^N)' \quad (3.3.5)$$

e uma matriz $q \times N$, $G(t, \lambda) = (g_{jn}(t, \lambda))$, $j=1, 2, \dots, q$, $n=1, 2, \dots, N$, obteremos a rela\u00e7\u00e3o (3.3.4) na forma

$$X_t = \int_{-\infty}^t G(t, \lambda) dZ_{\lambda} \quad , \quad (3.3.6)$$

chamada a representa\u00e7\u00e3o can\u00f4nica do processo (X_t) .

O processo ortogonal (Z_{λ}) \u00e9, \u00e0s vezes, chamado de *processo inova\u00e7\u00e3o total* associado a (X_t) e o cardinal N \u00e9 chamado de *multiplicidade espectral* de (X_t) .

Em resumo, estabeleceu-se o seguinte teorema:

Teorema 3.1 :

Qualquer processo vetorial n\u00e3o nulo, cont\u00ednuo, de segunda ordem, com fun\u00e7\u00e3o m\u00e9dia identicamente nula, pode ser decomposto de maneira \u00fanica como soma de outros dois processos, mutuamente ortogonais, um deles determin\u00edstico e o outro n\u00e3o determin\u00edstico puro.

Ainda, para o componente (X_t) n\u00e3o determin\u00edstico puro, existe um processo vetorial N -dimensional $Z_t = (Z_t^1 \dots Z_t^N)'$ deter

minado por variáveis não nulas $Z^1 \dots Z^N$, com incrementos ortogonais, e uma matriz $G(t, \lambda)$ definida como em (3.3.5), tais que

$$X_t = \int_{-\infty}^t G(t, \lambda) dZ_\lambda .$$

Em analogia com o teorema 2.2 segue o seguinte:

Teorema 3.2 :

Seja $h > 0$ dado. Para qualquer processo (X_t) não determinístico puro, contínuo, sob as condições do teorema 3.1, a melhor predição da componente X_t^j em termos de todas as variáveis X_s^j , $j=1, \dots, q$, $s \leq t-h$, será

$$\pi_{t-h} X_t^j = \sum_{k=1}^N \int_{-\infty}^{t-h} g_{jk}(t, s) dZ_s^k$$

com o correspondente erro de predição

$$\sigma_{th}^j = \|X_t^j - \pi_{t-h} X_t^j\| = \left[\sum_{k=1}^N \int_{t-h}^t |g_{jk}(t, s)|^2 dF_{Z^k}(s) \right]^{1/2}$$

Observação: Uma prova detalhada do teorema 3.1, com uma hipótese mais fraca com relação à continuidade do processo, pode ser encontrada em H. Cramér [6].

4. Uma classe de processos não estacionários :

- Processos Harmonizáveis -

Consideraremos agora um processo discreto (X_t) , dado

por:

$$X_t = \int_0^{2\pi} e^{it\lambda} dZ_\lambda \quad (3.4.1)$$

onde $(Z_\lambda : \lambda \in [0, 2\pi])$ é um processo satisfazendo as condições:

$$E(Z_\lambda) = 0 \quad , \quad \text{para todo } \lambda \in [0, 2\pi] \quad (3.4.2)$$

e

$$E(Z_\lambda \overline{Z_\nu}) = F(\lambda, \nu) \in \mathbb{C}, \text{ para todo } (\lambda, \nu) \in [0, 2\pi]^2$$

onde F é de variação limitada em seu domínio, isto é, existe $k > 0$ tal que para cada par de partições $(t_i), (s_j)$ de $[0, 2\pi]$,

$$\sum_{i,j} |\Delta_{2F_{ij}}| = \sum_{i,j} |F(t_{i+1}, s_{j+1}) - F(t_i, s_{j+1}) - F(t_{i+1}, s_j) + F(t_i, s_j)| < k \quad (3.4.3)$$

Portanto, sendo $F(\lambda, \nu)$ de variação limitada, a integral (3.4.1) existe e o processo está definido.

Processos desse tipo foram introduzidos por Loève e foram chamados por ele de *processos harmonizáveis*. A função covariância correspondente ao processo definido por (3.4.1) é

$$R(s, t) = E(X_s \overline{X_t}) = \int_0^{2\pi} \int_0^{2\pi} e^{-i(s\lambda - t\nu)} d_{\lambda, \nu} F(\lambda, \nu) \quad (3.4.4)$$

Reciprocamente, se a função covariância for dada por (3.4.4) onde $F(\lambda, \nu)$ é uma função covariância satisfazendo (3.4.3), então existirá um processo (Z_λ) satisfazendo (3.4.2) tal que o processo (X_t) será dado por (3.4.1). Para mostrarmos a validade desta afirmação, estabeleçamos alguns resultados:

Observemos que sendo $F(\lambda, \nu)$ de variação limitada, o processo (X_t) é contínuo em média e as covariâncias são contínuas e limitadas. Além disso, $F(\lambda \pm 0, \nu \pm 0)$, $Z_{\lambda \pm 0}$ e $Z_{\mp \infty}$ existem.

Sejam (\tilde{Z}_λ) e $\Delta_o Z_\lambda$, o processo (Z_λ) normalizado e o salto de (Z_λ) em λ , definidos por:

$$\tilde{Z}_\lambda = \frac{1}{2} [Z_{\lambda+0} + Z_{\lambda-0}], \quad \Delta_o Z_\lambda = Z_{\lambda+0} - Z_{\lambda-0}$$

De modo análogo, sejam $\hat{F}(\lambda, \nu)$ e $\Delta_o \Delta_o' F(\lambda, \nu)$ a função $F(\lambda, \nu)$ normalizada e o salto de $F(\lambda, \nu)$ em (λ, ν) definidos por:

$$\hat{F}(\lambda, \nu) = \frac{1}{4} [F(\lambda+0, \nu+0) + F(\lambda+0, \nu-0) + F(\lambda-0, \nu+0) + F(\lambda, \nu)]$$

$$\Delta_o \Delta_o' F(\lambda, \nu) = F(\lambda+0, \nu+0) - F(\lambda+0, \nu-0) - F(\lambda-0, \nu+0) + F(\lambda-0, \nu-0).$$

Segue portanto que

$$E(\tilde{Z}_\lambda \tilde{Z}_\nu) = \hat{F}(\lambda, \nu), \quad E(\Delta_o \tilde{Z}_\lambda \Delta_o' \tilde{Z}_\nu) = \Delta_o \Delta_o' \hat{F}(\lambda, \nu).$$

Consideremos o operador $\Delta_h \Delta_h'$, tal que:

$$\Delta_h \Delta_h' F(\lambda, \nu) = F(\lambda+h, \nu+h') - F(\lambda+h, \nu-h') - F(\lambda-h, \nu+h) + F(\lambda-h, \nu-h')$$

$$\text{Seja } a_r(\eta) = \frac{\text{sen } r\eta}{r\eta} \text{ e seja } b_r(\nu, h) = \frac{1}{\pi} \int_{r(\nu-h)}^{r\nu} \frac{\text{sen } \eta}{\eta} d\eta$$

Usaremos repetidamente os fatos:

$$b_r(\nu, h) \rightarrow \begin{cases} 0 & \text{se } \nu(\nu-h) > 0 \\ 1 & \text{se } \nu(\nu-h) < 0 \\ \frac{1}{2} & \text{se } \nu(\nu-h) = 0 \end{cases}$$

Lema 4.1 :

Se a covariância $R(s, t)$ satisfizer (3.4.4), então quando $r, r' \rightarrow \infty$

$$\frac{1}{4rr'} \int_{-r}^r \int_{-r'}^{r'} e^{-i(\lambda s - \nu t)} R(s, t) ds dt \rightarrow \Delta_o \Delta_o' F(\lambda, \nu) \quad (1)$$

e

$$\frac{1}{4\pi^2} \int_{-r}^r \int_{-r'}^{r'} \frac{1}{st} \Delta_h \Delta_h' e^{-i(\lambda s - \nu t)} R(s, t) ds dt \rightarrow \Delta_h \Delta_h' F(\lambda, \nu) \quad (2)$$

Prova

De (3.4.4) segue, por transformações elementares, que as integrais (1) e (2) são respectivamente

$$\int \int a_r(\eta - \lambda) a_{r'}(\eta' - \nu) d_{\eta, \eta'} F(\eta, \eta')$$

$$\int \int b_r(\eta - \lambda, h) b_{r'}(\eta' - \nu, h') d_{\eta, \eta'} F(\eta, \eta')$$

e o resultado é obtido fazendo-se $r, r' \rightarrow \infty$

Lema 4.2 :

Se (X_t) for um processo com função covariância dada por (3.4.4), existem processos $(\Delta_o Z_\lambda)$ com variância $\Delta_o \Delta_o' F(\lambda, \nu)$, e (\hat{Z}_λ) com covariância $\hat{F}(\lambda, \nu)$, tais que, quando $r \rightarrow \infty$,

$$\frac{1}{2r} \int_{-r}^r e^{-i\lambda s} X_s ds \xrightarrow{m.q.} \Delta_o Z_\lambda$$

$$\frac{1}{2\pi} \int_{-r}^r -\frac{1}{is} \Delta_h e^{-i\lambda s} X_s ds \xrightarrow{m.q.} \Delta_h \hat{Z}_\lambda$$

Se o processo (X_t) for dado por (3.4.1) então $(\Delta_o Z_\lambda)$ será o salto de (Z_λ) e (\tilde{Z}_λ) será o (Z_λ) normalizado.

Prova

A primeira afirmação segue do lema anterior e a segunda segue diretamente de (3.4.1) com a observação que, por transformações elementares, as integrais anteriores se tornam respectivamente

$$\int a_r(\eta-\lambda) dZ_\eta \quad \int b_r(\eta-\lambda) dZ_\eta .$$

Estamos agora em condições de mostrar que se a função covariância for dada por

$$R(s, t) = \int_0^{2\pi} \int_0^{2\pi} e^{-i(s\lambda - t\nu)} d_{\lambda, \nu} F(\lambda, \nu) \text{ com } \iint |d_{\lambda, \nu} F(\lambda, \nu)| < \infty$$

onde $F(\lambda, \nu)$ é uma função de variação limitada, existirá um processo (Z_λ) satisfazendo

$$E(Z_\lambda) = 0, \quad E(Z_\lambda \bar{Z}_\nu) = F(\lambda, \nu) \text{ para todo } (\lambda, \nu) \in [0, 2\pi],$$

tal que

$$X_t = \int_0^{2\pi} e^{it\lambda} dZ_\lambda$$

De fato, sendo (X_t) um processo com covariância dada por (3.4.4) com

$$\iint |d_{\lambda, \nu} F(\lambda, \nu)| < \infty, \text{ sendo o integrando em (3.4.4)}$$

uma função contínua e limitada e ainda, $F(\lambda, \nu)$ de variação limi-

tada, podemos afirmar, sem restrição à generalidade, que $F(\lambda, \nu)$ é normalizada.

De acordo com o lema 4.2 existe um processo (Z_λ) cuja covariância é $F(\lambda, \nu)$ tal que

$$\frac{1}{2\pi} \int_{-r}^r -\frac{1}{is} \Delta_h e^{-is\lambda} X_s ds \xrightarrow{m.q} \Delta_h Z_\lambda, \text{ quando } r \rightarrow \infty$$

Também pelos lemas, segue, por cálculos elementares, que

$$Y_t = \int e^{it\lambda} dZ_\lambda \text{ existe e } E(|X_t|^2) = E(X_t \overline{Y_t}) = E(|Y_t|^2),$$

de modo que $E(|X_t - Y_t|^2) = 0$, e a prova está completa.

Admitamos que a função F defina uma distribuição de "massa complexa" sobre $Q = [0, 2\pi]^2$ (estamos supondo que $\lambda \mapsto Z_\lambda$ seja contínua); assim sendo, a massa em um retângulo $[t_i, t_{i+1}] \times [s_j, s_{j+1}]$ será $\Delta_2 F_{ij}$. Segue então, das propriedades de uma função covariância, que as massas admitidas por conjuntos simétricos em relação à diagonal $\lambda = \nu$ de Q são complexas conjugadas e, em particular, que cada subconjunto da diagonal admite massa real.

A função F é chamada *função espectral* do processo (X_t) enquanto que a distribuição definida por F é chamada a *distribuição espectral* de (X_t) .

Nestes termos, quando toda a massa espectral estiver situada na diagonal, como $E(dZ_\lambda \overline{dZ_\nu}) = d_{\lambda, \nu} F(\lambda, \nu)$, teremos o processo (Z_λ) , de incrementos ortogonais, e (X_t) será então um proces

so estacionário.

No caso geral temos que

$$F = F_1 + F_2 + F_3 \quad (3.4.5)$$

sendo cada uma das covariâncias F_i , de variação limitada sobre Q . Aqui F_2 é absolutamente contínua com densidade espectral f_2 , de modo que

$$F_2(\lambda, \nu) = \int_0^\lambda \int_0^\nu f_2(s, t) \, ds dt$$

Por outro lado, as distribuições F_1 e F_3 têm suas massas totais concentradas em conjuntos de medida nula em Q . Para F_1 , esse conjunto é no máximo enumerável e cada ponto leva uma massa não nula, enquanto que F_3 é não enumerável e cada ponto leva massa nula. Desta forma, sendo no caso estacionário $F_2=0$, teremos F_1 e F_3 com massas totais concentradas sobre a diagonal.

Vejamos agora uma condição para que o processo harmonizável (X_t) dado por (3.4.1) seja determinístico. Para isto, lembremos que (X_t) será determinístico se e somente se para cada t e h dados, existirem constantes c_0, c_1, \dots, c_r de modo que

$$W = \left\| X_t - c_0 X_{t-h} - c_1 X_{t-h-1} - \dots - c_r X_{t-h-r} \right\|^2 \quad (3.4.6)$$

seja arbitrariamente pequeno. (notar que isto significa que

$$X_t \in \mathcal{L}(X, t-h)$$

Seja agora

$$g(\lambda) = e^{it\lambda} - c_0 e^{i(t-h)\lambda} - \dots - c_r e^{i(t-h-r)\lambda} \quad (3.4.7)$$

Com apoio em (3.4.4) concluimos que

$$W = \int_0^{2\pi} \int_0^{2\pi} g(\lambda) \overline{g(v)} dF(\lambda, v)$$

e pela desigualdade de Schwarz temos então que

$$W^2 \leq \int_0^{2\pi} \int_0^{2\pi} |g(\lambda)|^2 |dF(\lambda, v)| \int_0^{2\pi} \int_0^{2\pi} |g(v)|^2 |dF(\lambda, v)|.$$

Pela simetria da distribuição espectral os fatores no segundo membro são iguais e então

$$W^2 \leq \int_0^{2\pi} |g(\lambda)|^2 dG(\lambda) \text{ onde } G(\lambda) = \int_0^\lambda \int_0^{2\pi} |dF(s, t)|$$

Como G é não decrescente e limitada então G' é não negativa quase sempre em [0, 2π].

Consideremos por fim a integral

$$\omega = \int_0^{2\pi} \log G'(\lambda) d\lambda .$$

Vale então o teorema (H.Cramér [6]):

Teorema 4.1

Se $\omega = -\infty$, o processo (X_t) será determinístico.

Em particular, se as componentes F_1 e F_3 em (3.4.5)

forem nulas, a massa total da distribuição estará concentrada em pontos isolados e então $G'(\lambda)=0$ quase sempre, de modo que ainda $\omega=-\infty$ e o processo será determinístico.

Consideremos agora, para finalizar, um processo (X_t) com uma função espectral F tendo uma componente absolutamente contínua F_2 não identicamente nula e ainda, que a densidade espectral f_2 seja de $L_2(Q)$. Disto temos

$$f_2(\lambda, \nu) = \sum_{p \geq 1} \mu_p \psi_p(\lambda) \overline{\psi_p(\nu)} \quad (3.4.8)$$

onde μ_p e ψ_p são respectivamente os autovalores e as auto-funções de f_2 , sendo os μ_p reais positivos ($\neq 0$) e (ψ_p) , família ortogonal em $L_2(0, 2\pi)$.

Temos então o

Teorema 4.2

Suponhamos que na expansão em (3.4.8) exista p de modo que

$$\psi_p(\lambda) = \sum_{q=m_0}^{\infty} b_{pq} e^{-iq\lambda}, \quad m_0 > -\infty \quad \text{e} \quad b_{p m_0} \neq 0.$$

Então (X_t) é não determinístico e m_0 pertence ao seu espectro de inovação.

Prova

Fazendo-se $n=m_0$ e $h=1$ nas igualdades (3.4.6) e

e (3.4.7), obtem-se

$$\begin{aligned} W &= \int_0^{2\pi} \int_0^{2\pi} g(\lambda) \overline{g(\nu)} dF(\lambda, \nu) \geq \int_0^{2\pi} \int_0^{2\pi} g(\lambda) \overline{g(\nu)} f_2(\lambda, \nu) d\lambda d\nu \\ &= \sum_{p=1}^{\infty} \mu_p \left| \int_0^{2\pi} g(\lambda) \psi_p(\lambda) d\lambda \right|^2 = 4\pi^2 \mu_p |b_{pm_0}|, \end{aligned}$$

independentemente da escolha dos coeficientes c_j . Logo,

$$\|X_{m_0} - \pi_{m_0-1} X_{m_0}\| \neq 0 \text{ e conclue-se a tese.}$$

IV - TEORIA ESPECTRAL PARA UMA CLASSE DE PROCESSOS NÃO-ESTACIONÁRIOS

1. Introdução

Como já foi mencionado inicialmente, nem sempre a su posição de estacionariedade pode ser admitida na prática ao se estudar um processo estocástico; sendo assim, uma das dificulda des encontradas é formular uma teoria espectral, envolvendo con ceitos clássicos tais como frequência e energia, que permita u- ma interpretação significativa e útil para uma função espectral então definida.

Uma idéia que se apresenta é generalizar a seguinte si tuação: Se $(X_t : t \in T)$ for um processo não estacionário da for ma

$$X_t = \begin{cases} X_t^{(1)} & \text{se } t \leq t_0 \\ X_t^{(2)} & \text{se } t > t_0 \end{cases}$$

onde $(X_t^{(1)} : t \in T)$ e $(X_t^{(2)} : t \in T)$ são processos estacionários com funções covariância diferentes, será certamente possível se obter alguma informação sobre o espectro (novamente, a denominação usual para $dF(\lambda)$) de $(X_t : t \in T)$ através de uma amostra de $t = t_0 - N$ a $t = t_0 + N$. Se ainda t_0 for conhecido, presumivelmente seja possí vel se estimar duas funções espectrais - uma para cada processo estacionário. A generalização que se considera é então supor o

espectro para um processo não estacionário como sendo dependente do parâmetro t . É claro que neste caso não se deve esperar se estimar o espectro em cada particular t considerado, mas se for assumido que este espectro varia "lentamente" em T , pode-se tentar estimá-lo em alguma vizinhança do ponto considerado, usando-se estimadores que envolvam apenas funções "locais" do processo. O que será apresentado, segundo Priestley [20], é uma definição de uma quantidade espectral possuindo uma certa relação com o resultado apresentado por Page [16] quando introduziu a noção de "espectro instantâneo de potência" ao definir a função densidade espectral como

$$f^*(\lambda) = \lim_{N \rightarrow \infty} f_N^*(\lambda) \quad \text{onde} \quad f_N^*(\lambda) = E\left(\left|\int_0^N X_t e^{-i\lambda t} dt\right|^2\right)$$

e escrevendo, para cada λ ,

$$f_N^*(\lambda) = \int_0^N \rho_t(\lambda) dt$$

O espectro instantâneo de potência, representando a diferença entre o conteúdo espectral do processo no intervalo $(0, t+\delta t)$ e o intervalo $(0, t)$, foi definido por Page como

$$\rho_t(\lambda) = \frac{d}{dt} (f_t^*(\lambda)) \quad \text{sendo} \quad f^*(\lambda) = \int_0^\infty \rho_t(\lambda) dt$$

A diferença entre o que Page obteve e o que será desenvolvido é que neste caso será estudado o conteúdo espectral do processo dentro do intervalo $(t, t+\delta t)$.

2. Uma classe de processos não estacionários:

- Processos Oscilatórios -

Seja $(X_t : t \in T)$ um processo estocástico de segunda ordem, complexo, a parâmetro contínuo, com

$$\begin{aligned} E(X_t) &= 0 && \text{para todo } t \in R \\ R(s,t) &= E(X_s \bar{X}_t) && \text{para todo } (s,t) \in R^2 \end{aligned} \quad (4.2.1)$$

(os resultados a serem apresentados se aplicam igualmente ao caso discreto com as usuais modificações).

Vamos nos restringir à classe de processos para os quais existe uma família $(\phi_t(\lambda))$ de funções definidas em R e indexadas em $T=R$ e uma medida $\mu(\lambda)$ em R , tal que para cada $(s,t) \in R^2$, a função covariância $R(s,t)$ definida em (4.2.1) admite uma representação da forma

$$R(s,t) = \int_{-\infty}^{\infty} \phi_s(\lambda) \overline{\phi_t(\lambda)} d\mu(\lambda) \quad (4.2.2)$$

Observemos que se $\phi_t(\lambda)$ for, para cada t , quadraticamente integrável com respeito à medida μ , então $R(t,t) < \infty$ para todo $t \in R$.

Vamos mostrar que se a função covariância admitir uma representação da forma (4.2.2), o processo admitirá uma representação da forma

$$X_t = \int_{-\infty}^{\infty} \phi_t(\lambda) dY_\lambda \quad (4.2.3)$$

onde (Y_λ) é um processo com incrementos ortogonais com

$$E(|dY_\lambda|^2) = d\mu(\lambda) \quad (4.2.4)$$

De fato, seja $L_2(\mu)$ o conjunto de todas as funções quadraticamente integráveis com respeito a μ e suponhamos que a família de funções de λ , $(\phi_t(\lambda))$, quando t percorrer R , seja uma base para $L_2(\mu)$.

Se $A \subset R$ for de medida μ finita, sua função característica $I_A(\lambda) \in L_2(\mu)$ e poderá ser aproximada (em média quadrática) por somas

$$I_A^{(n)}(\lambda) = \sum_{v=1}^n a_v^{(n)} \phi_{t_v}^{(n)}(\lambda)$$

Como

$$\int_R \left| I_A^{(n)}(\lambda) - I_A^{(m)}(\lambda) \right|^2 d\mu(\lambda) = E\left(\left| Y_n - Y_m \right|^2 \right)$$

onde

$$Y_n = \sum_{v=1}^n a_v^{(n)} X_{t_v}^{(n)},$$

segue que a sequência $\{Y_n\}$ é uma sequência de Cauchy e converge para uma variável aleatória que será representada por $Y(A)$.

Se A_1 e A_2 forem disjuntos,

$$E\left(\left| Y(A_1 \cup A_2) - Y(A_1) - Y(A_2) \right|^2 \right) = \int \left| I_{A_1 \cup A_2}(\lambda) - I_{A_1}(\lambda) - I_{A_2}(\lambda) \right|^2 d\mu(\lambda) = 0$$

Pode-se verificar facilmente que

$$E[Y(A_1) \cdot \overline{Y(A_2)}] = \mu(A_1 \cap A_2).$$

$$\text{Façamos } Z_t = \int_R \phi(\lambda) dY_\lambda$$

$$\begin{aligned} E(X_t \cdot \overline{Y(A)}) &= \lim_{n \rightarrow \infty} E(X_t \sum_{v=1}^n \overline{a_v^{(n)}} \overline{X}_{t_v^{(n)}}) \\ &= \lim_{n \rightarrow \infty} \sum_{v=1}^n \overline{a_v^{(n)}} \int_R \phi(\lambda) \overline{\phi_{t_v^{(n)}}(\lambda)} d\mu(\lambda) = \int_A \phi_t(\lambda) d\mu(\lambda) \end{aligned}$$

Usando novamente um procedimento de aproximação, obtemos

$$E(X_t \overline{Z_t}) = \int_R |\phi_t(\lambda)|^2 d\mu(\lambda)$$

Então,

$$E(|X_t - Z_t|^2) = 0$$

e portanto (X_t) admite uma representação da forma (4.2.3)

Se a família $(\phi_t(\lambda))$ não constituir uma base para $L_2(\mu)$, acrescentaremos todas as funções ortogonais a cada $\phi_t(\lambda)$, $t \in T$, representadas por $g_t(\lambda)$, $t \in T'$ onde $T' \cap T = \emptyset$.

$$\text{Definamos } h_t(\lambda) = \begin{cases} \phi_t(\lambda) & \text{se } t \in T \\ g_t(\lambda) & \text{se } t \in T' \end{cases}$$

Consideremos $Z_t = X_t$ quando $t \in T$ e ainda, (Z_t) o pro -

cesso normal com covariância

$$\int_R g_s(\lambda) \overline{g_t(\lambda)} d\mu(\lambda) = [\overline{g_t(\lambda)} \cdot g_s(\lambda)]$$

independente de (X_t) , para $t \in T'$.

Então, (Z_t) possui covariância dada por

$$\int_R h_s(\lambda) \cdot \overline{h_t(\lambda)} d\mu(\lambda) \quad \text{para } s, t \in T \cup T',$$

onde $h_t(\lambda)$ é uma base para $L_2(\mu)$ quando t percorrer $T \cup T'$.

Logo,

$$Z_t = \int_R h_t(\lambda) dY_\lambda \quad \text{para } t \in T \cup T'$$

onde (Y_λ) é um processo com incrementos ortogonais.

Mas

$$X_t = Z_t = \int_R h_t(\lambda) dY_\lambda = \int_R \phi_t(\lambda) dY_\lambda \quad \text{para } t \in T$$

e obtemos para (X_t) a representação dada em (4.2.3).

Este resultado dá uma representação do processo como o limite de certas combinações lineares de funções de $t, \phi_t(\lambda)$, com pesos dY_λ .

A medida $\mu(\lambda)$ na representação dada através de (4.2.2) e (4.2.4) atua da mesma forma que a $F(\lambda)$ no caso do processo estacionário de modo que a situação análoga ao caso de $F(\lambda)$ ser absolutamente contínua é obtida assumindo-se $u(\lambda)$ como

absolutamente contínua com respeito à medida de Lebesgue.

Para a classe dos processos estacionários foi visto que uma escolha válida da família $(\phi_t(\lambda))$ é a dada por $\phi_t(\lambda) = e^{i\lambda t}$ que fornece a decomposição espectral

$$X_t = \int_I e^{i\lambda t} dY_\lambda \quad \text{com } E(|dY_\lambda|^2) = dF(\lambda), \quad (4.2.5)$$

em termos de "ondas seno e cosseno", constituindo a base da interpretação física da análise espectral como uma distribuição de energia sobre frequência. Já para a classe dos processos não estacionários no entanto, a escolha da família $(\phi_t(\lambda))$ como sendo a família das exponenciais complexas não é tão válida (mesmo porque a representação (4.2.5) implica que o processo é estacionário) e se desejarmos introduzir a noção de frequência também para este caso, seremos levados a considerar novos elementos básicos que, embora sejam não estacionários, possuam uma forma oscilatória na qual a noção de frequência ainda seja dominante.

A fim de obtermos uma classe de tais elementos (na verdade família de funções) suponhamos que para cada λ fixado, $\phi_t(\lambda)$, considerada como uma função de t , possua uma transformada de Fourier generalizada (via distribuições) cujo módulo tenha máximo absoluto em algum $\theta(\lambda)$. Poderemos então observar $\phi_t(\lambda)$ como uma onda seno de amplitude modulada com frequência $\theta(\lambda)$ e escrever

$$\phi_t(\lambda) = A_t(\lambda) \cdot e^{i\theta(\lambda)t} \quad (4.2.6)$$

onde a função modulação $A_t(\lambda)$ é tal que o módulo de sua transformada de Fourier generalizada tem um máximo absoluto na origem.

Definição 2.1

$\phi_t(\lambda)$, considerada como uma função de t , será chamada uma *função oscilatória* se para algum $\theta(\lambda)$, necessariamente único, puder ser escrita na forma

$$\phi_t(\lambda) = A_t(\lambda) e^{i\theta(\lambda)t}$$

onde $A_t(\lambda) = \int_{-\infty}^{\infty} e^{i\theta t} dH_{\lambda}(\theta)$ com $|dH_{\lambda}(\theta)|$ tendo um máximo

absoluto em $\theta=0$.

Se ainda $(\phi_t(\lambda))$ for tal que elementos distintos da família não possuam transformadas de Fourier cujos máximos ocorram em um mesmo ponto, mudando a variável de λ para $\theta(\lambda)$ em (4.2.2) e redefinindo convenientemente $A_t(\lambda)$ e $\mu(\lambda)$, poderemos escrever

$$R(s, t) = \int_{-\infty}^{\infty} A_s(\lambda) \overline{A_t(\lambda)} e^{i\lambda(s-t)} d\mu(\lambda) \quad (4.2.7)$$

e correspondentemente,

$$X_t = \int_{-\infty}^{\infty} A_t(\lambda) e^{i\lambda t} dY_{\lambda} \quad \text{onde } E(|dY_{\lambda}|^2) = d\mu(\lambda) \quad (4.2.8)$$

Definição 2.2 :

Se existir uma família de funções oscilatórias em termos das quais o processo (X_t) possui uma representação da forma (4.2.3), então (X_t) será chamado um *processo oscilatório*.

Portanto, qualquer processo oscilatório pode ser representado na forma (4.2.8) onde $A_t(\lambda)$ satisfaz a condição da definição 2.1 podendo-se escrever, sem perda de generalidade,

$$\phi_t(\lambda) = A_t(\lambda) e^{it\lambda}$$

Notemos que, sendo $\phi_t(\lambda) = e^{i\lambda t}$ um caso particular de (4.2.6) com $A_t(\lambda) \equiv 1$ para todo t , e $\theta(\lambda) \equiv \lambda$, a classe dos processos oscilatórios inclui todos os processos estacionários de segunda ordem.

Para qualquer particular processo oscilatório, em geral, existem diferentes famílias de funções oscilatórias cada uma das quais induzindo uma medida $\mu(\lambda)$ diferente e em termos das quais o processo possui uma representação da forma (4.2.8). Para cada família, procura-se definir o espectro do processo em relação a essa família como sendo simplesmente a medida $\mu(\lambda)$. No entanto, uma tal definição poderá não representar uma distribuição de energia sobre frequência.

Se considerarmos que a variância de um processo pode ser interpretada como a medida da sua energia total no tempo t , então

$$\text{Var}(X_t) = R(t, t) = \int_{-\infty}^{\infty} |A_t(\lambda)|^2 d\mu(\lambda) \quad (4.2.9)$$

fornecerá uma decomposição da energia total na qual a contribuição de λ é $|A_t(\lambda)|^2 d\mu(\lambda)$, resultado este que será consistente com o de interpretarmos (4.2.8) como sendo a forma limite de uma soma de ondas seno de diferentes frequências tendo amplitudes aleatórias $(A_t(\lambda)dY_\lambda)$ variando em T . Sendo assim, somos levados a definir o espectro para um processo oscilatório, como segue:

Definição 2.3 :

Se \mathcal{F} for uma família de funções oscilatórias, $(\phi_t(\lambda)) = (A_t(\lambda)e^{i\lambda t})$, em termos da qual um processo oscilatório (X_t) possui uma representação da forma (4.2.8), define-se o espectro evolutivo de potência no ponto t com respeito à família \mathcal{F} , como sendo

$$dF_t(\lambda) = |A_t(\lambda)|^2 d\mu(\lambda) \quad (4.2.10)$$

Observemos que no caso de (X_t) ser estacionário e \mathcal{F} a família das exponenciais complexas,

$$dF_t(\lambda) = d\mu(\lambda) = E(|dY_\lambda|^2) = dF(\lambda)$$

Embora, de acordo com a definição 2.3, o espectro evolutivo dependa da família \mathcal{F} considerada, por (4.2.9) e (4.2.10) temos

$$\text{Var}(X_t) = R(t, t) = \int_{-\infty}^{\infty} dF_t(\lambda) \quad (4.2.11)$$

de modo que o valor desta integral independe da particular família \mathcal{F} e representa a energia total do processo no ponto t , qualquer que seja \mathcal{F} .

Será conveniente uniformizarmos as funções $A_t(\lambda)$ de modo que, para todo λ , $A_0(\lambda)=1$, isto é, incorporarmos $|A_0(\lambda)|$ na medida $\mu(\lambda)$. Com esta convenção, $d\mu(\lambda)$ representará o espectro evolutivo em $t=0$ e $|A_t(\lambda)|^2$ a variação no espectro relativa a $t=0$. Portanto, para cada λ ,

$$\int_{-\infty}^{\infty} dH_{\lambda}(\theta) = 1$$

e as transformadas de Fourier de $(A_t(\lambda))$ serão normalizadas a fim de ter integrais unitárias.

Como um exemplo de processo não estacionário admitindo uma representação da forma (4.2.8), podemos apresentar o processo (X_t) dado por

$$X_t = c(t)X_t^{(0)} \quad (4.2.12)$$

onde $c(t)$ (com $c(0)=1$) é uma função com transformada de Fourier generalizada cujo módulo tem um máximo absoluto na origem (por exemplo, $c(t)$ pode ser qualquer função real não negativa cuja transformada de Fourier existe) e $X_t^{(0)}$ é um processo estacionário com função média identicamente nula e função distribuição espectral $F(\lambda)$.

Portanto,

$$X_t = \int_{-\infty}^{\infty} c(t) e^{i\lambda t} dY_\lambda$$

onde (Y_λ) é um processo com incrementos ortogonais e $E(|dY_\lambda|^2) = dF(\lambda)$.

Como a família $\mathcal{F}_0 = (c(t)e^{i\lambda t})$ é tal que qualquer de seus elementos está nas condições da definição 2.1, o processo definido por (4.2.12) é um processo oscilatório e com respeito à família \mathcal{F}_0 tem espectro evolutivo dado por

$$dF_t(\lambda) = |c(t)|^2 dF(\lambda) .$$

Observemos que o processo definido em (4.2.12) é um caso especial do modelo (4.2.8) no sentido que, para qualquer par (λ_1, λ_2) e (t_1, t_2) ,

$$\frac{dF_{t_1}(\lambda_1)}{dF_{t_2}(\lambda_1)} = \frac{dF_{t_1}(\lambda_2)}{dF_{t_2}(\lambda_2)} \quad (4.2.13)$$

Um processo para o qual existe uma família \mathcal{F}_0 tal que o espectro evolutivo com relação a \mathcal{F}_0 satisfaz (4.2.13) é chamado um processo *uniformemente modulado*.

3. Efeito de filtros

Se para um processo estacionário (X_t) com representa

ção espectral

$$X_t = \int_{-\infty}^{\infty} e^{i\lambda t} dY_\lambda, \quad E(|dY_\lambda|^2) = dF(\lambda)$$

for considerada uma transformação linear do tipo

$$\hat{X}_t = \int_{-\infty}^{\infty} X_{t-s} g(s) ds, \quad (4.3.1)$$

foi visto (Cap. II, 4-2) que as distribuições espectrais de (X_t) e (\hat{X}_t) estão relacionadas por:

$$d\hat{F}(\lambda) = |\Gamma(\lambda)|^2 dF(\lambda) \quad (4.3.2)$$

onde $\Gamma(\lambda) = \int_{-\infty}^{\infty} e^{-is\lambda} g(s) ds$ é a função transferência do filtro g .

Podemos observar que uma das consequências mais úteis da representação espectral de um processo estacionário é a de ser possível se descrever o efeito de operações lineares (ou filtros) somente em termos do efeito sobre componentes espectrais individuais. Esta propriedade também será válida (em um sentido aproximado) para espectros evolutivos quando forem consideradas transformações lineares de processos não estacionários.

Seja então (X_t) um processo satisfazendo um modelo da forma (4.2.8) e consideremos a transformação

$$\hat{X}_t = \int_{-\infty}^{\infty} X_{t-s} e^{-i\lambda_0(t-s)} g(s) ds \quad \text{onde } \lambda_0 = \text{constante qq}$$

Então podemos escrever:

$$\begin{aligned} \hat{X}_t &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} A_{t-s}(\lambda + \lambda_0) e^{i(\lambda + \lambda_0)(t-s)} dY_{\lambda + \lambda_0} \right] e^{-i\lambda_0(t-s)} g(s) ds \\ &= \int_{-\infty}^{\infty} g(s) \int_{-\infty}^{\infty} (A_{t-s}(\lambda + \lambda_0) / A_t(\lambda + \lambda_0)) A_t(\lambda + \lambda_0) e^{i\lambda t} \cdot e^{-i\lambda s} dY_{\lambda + \lambda_0} ds \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} g(s) (A_{t-s}(\lambda + \lambda_0) / A_t(\lambda + \lambda_0)) e^{-i\lambda s} ds \right] A_t(\lambda + \lambda_0) e^{i\lambda t} dY_{\lambda + \lambda_0} \end{aligned}$$

Ou então,

$$\hat{X}_t = \int_{-\infty}^{\infty} \Gamma_{t, \lambda + \lambda_0}(\lambda) A_t(\lambda + \lambda_0) e^{i\lambda t} dY_{\lambda + \lambda_0} \quad (4.3.3)$$

onde, para qualquer t, ω, Θ ,

$$\Gamma_{t, \omega}(\Theta) = \int_{-\infty}^{\infty} g(s) (A_{t-s}(\omega) / A_t(\omega)) e^{-is\Theta} ds \quad (4.3.4)$$

A função $\Gamma_{t, \omega}(\lambda)$ será chamada a *função transferência generalizada* do filtro g com respeito à família \mathcal{F} .

A representação de (\hat{X}_t) dada em (4.3.3) não é necessariamente da forma (4.2.8) pois a transformada de Fourier generalizada de $(\Gamma_{t, \lambda + \lambda_0}(\lambda) \cdot A_t(\lambda + \lambda_0))$ pode não ter um máximo absoluto na origem.

Neste caso podemos considerar a família das funções

$$\hat{\phi}_t(\lambda) = \Gamma_{t, \lambda + \lambda_0}(\lambda) A_t(\lambda + \lambda_0) e^{i\lambda t} \quad (4.3.5)$$

que em geral são oscilatórias, embora sua frequência dominante seja um pouco diferente de λ .

Consideraremos agora o caso em que $A_{t-s}(\lambda)$, para cada t , λ varia lentamente comparada a $g(s)$, isto é, $g(s)$ tende rapidamente a zero quando $|s| \rightarrow \infty$ enquanto $A_{t-s}(\lambda)$ é aproximadamente constante sob a variação de s para a qual $g(s)$ é não desprezível.

Heuristicamente, então, para cada t, λ e todo θ ,

$$\Gamma_{t,\lambda}(\theta) \sim \Gamma(\theta) \quad (4.3.6)$$

e por (4.3.3),

$$\hat{X}_t \sim \int_{-\infty}^{\infty} A_t(\lambda + \lambda_0) e^{i\lambda t} d\tilde{Y}_\lambda$$

onde

$$(4.3.7)$$

$$E(|d\tilde{Y}_\lambda|^2) = |\Gamma(\lambda)|^2 d\mu(\lambda + \lambda_0)$$

Portanto, os espectros evolutivos de (X_t) e (\hat{X}_t) definidos com relação à mesma família $\mathcal{F} = (A_t(\lambda) e^{i\lambda t})$ estão heurísticamente relacionados por:

$$d_t \hat{F}(\lambda) \sim |\Gamma(\lambda)|^2 dF_t(\lambda + \lambda_0) \quad (4.3.8)$$

A fim de estabelecer mais precisamente a noção de uma função "variando lentamente" e examinar com mais detalhes a aproximação (4.3.6) introduziremos a noção de processos semi-estacionários.

4. Processos Semi-Estacionários

Suponhamos (X_t) um processo oscilatório cujas características não estacionárias variem lentamente em T , tendo uma representação da forma (4.2.8) em termos de uma família $\mathcal{F} = (A_t(\lambda)e^{i\lambda t})$ de funções oscilatórias tais que, para cada λ , $A_t(\lambda)$ é uma função de t variando lentamente, no sentido de possuir uma transformada de Fourier fortemente concentrada na região de $\theta = 0$.

Para cada família \mathcal{F} vamos definir uma medida para a "largura" de $|dH_\lambda(\theta)|$ como sendo a função $B_{\mathcal{F}}(\lambda)$ dada por

$$B_{\mathcal{F}}(\lambda) = \int_{-\infty}^{\infty} |\theta| |dH_\lambda(\theta)| \quad (4.4.1)$$

Definição 4.1 :

Uma família \mathcal{F} de funções oscilatórias será chamada *semi estacionária* se a função $B_{\mathcal{F}}(\lambda)$ for limitada para todo λ . A constante $B_{\mathcal{F}}$ definida por

$$B_{\mathcal{F}} = \frac{1}{\sup_{\lambda} (B_{\mathcal{F}}(\lambda))} \quad (4.4.2)$$

será a *largura característica* da família \mathcal{F} .

Definição 4.2 :

(X_t) será um processo semi-estacionário se existir uma

família semi-estacionária \mathcal{F} em termos da qual o processo admite uma representação da forma (4.2.8).

Observemos que um exemplo de processo semi-estacionário é o processo uniformemente modulado dado por (4.2.12) visto que a família $\mathcal{F}_0 = (c(t)e^{i\lambda t})$ é semi-estacionária.

Consideremos a classe \mathcal{E} das famílias \mathcal{F} semi-estacionárias em termos de cada uma das quais um particular processo semi-estacionário (X_t) admite uma representação espectral.

Denotaremos por B_x a largura característica do processo, dada por

$$B_x = \sup_{\mathcal{F} \in \mathcal{E}} (B_{\mathcal{F}}) \quad (4.4.3)$$

Deve-se notar que para processos estacionários a classe \mathcal{E} contém a família das exponenciais complexas que tem largura característica infinita; conseqüentemente, todo processo estacionário tem largura característica infinita.

Seja agora $\mathcal{E}^* \subset \mathcal{E}$ a sub-classe de famílias cujas larguras características são cada uma igual a B_x e seja \mathcal{F}^* um elemento qualquer de \mathcal{E}^* . Se $\mathcal{E}^* = \emptyset$, consideraremos \mathcal{F}^* qualquer família cuja largura característica seja suficientemente próxima de B_x .

Admitamos que seja

$$X_t = \int_{-\infty}^{\infty} A_t^*(\lambda) e^{i\lambda t} dY_{\lambda}^* \quad (4.4.4)$$

onde $E(|dY_\lambda^*|^2) = d\mu^*(\lambda)$ e $\phi_t(\lambda) = A_t^*(\lambda)e^{i\lambda t} \in \mathcal{F}^*$, a representação espectral do processo (X_t) em termos da família \mathcal{F}^* . Torna-se claro agora que, se for definido o espectro evolutivo de (X_t) com respeito a \mathcal{F}^* , então a aproximação (4.3.8) será válida se para cada λ , $dH_\lambda^*(\theta)$ (transformada de Fourier de $A_t^*(\lambda)$) se comportar como uma função- δ com respeito a $\Gamma(\lambda)$. Mais precisamente, consideremos a

Definição 4.3 :

Uma função $u(x)$ será considerada uma *pseudo função- δ de ordem ϵ* com respeito a uma função $v(x)$ se, para todo k , existir ϵ ($\ll 1$) independente de k tal que

$$\left| \int_{-\infty}^{\infty} u(x)v(x+k)dx - v(k) \int_{-\infty}^{\infty} u(x)dx \right| < \epsilon \quad (4.4.5)$$

Ainda, admitamos que

(a) o filtro g seja quadraticamente integrável e normalizado de modo a se obter

$$2\pi \int_{-\infty}^{\infty} |g(s)|^2 ds = \int_{-\infty}^{\infty} |\Gamma(\lambda)|^2 d\lambda = 1 \quad (4.4.6)$$

(b) uma medida para a largura do filtro g , representada por B_g , seja

$$B_g = \int_{-\infty}^{\infty} |s| |g(s)| ds \quad (4.4.7)$$

Lema 4.1 :

Seja \mathcal{F} uma família semi-estacionária com largura característica $B_{\mathcal{F}} = \frac{1}{\sup_{\lambda} (B_{\mathcal{F}}(\lambda))}$ onde $B_{\mathcal{F}}(\lambda)$ é dada por (4.4.1). Então para cada t, λ , $(e^{i\theta t} dH_{\lambda}(\theta))$ é uma pseudo função- δ de ordem $B_g/B_{\mathcal{F}}$ com respeito a $\Gamma(\theta)$.

Prova

Para todo k ,

$$\int_{-\infty}^{\infty} e^{it\theta} \Gamma(\theta+k) dH_{\lambda}(\theta) - \Gamma(k) \int_{-\infty}^{\infty} e^{it\theta} dH_{\lambda}(\theta) =$$

$$\int_{-\infty}^{\infty} e^{it\theta} [\Gamma(\theta+k) - \Gamma(k)] dH_{\lambda}(\theta) = \int_{-\infty}^{\infty} e^{it\theta} \Gamma'(k+\eta(\theta) \cdot \theta) dH_{\lambda}(\theta)$$

onde para cada θ , $0 \leq \eta(\theta) \leq 1$.

Logo,

$$\int_{-\infty}^{\infty} e^{it\theta} \Gamma(\theta+k) dH_{\lambda}(\theta) = \Gamma(k) \int_{-\infty}^{\infty} e^{it\theta} dH_{\lambda}(\theta) + R(k)$$

sendo

$$R(k) = \int_{-\infty}^{\infty} \theta e^{it\theta} \Gamma'(k+\eta(\theta) \cdot \theta) dH_{\lambda}(\theta)$$

Mas

$$|R(k)| < \sup_{\theta} |\Gamma'(\theta)| \int_{-\infty}^{\infty} |\theta| |dH_{\lambda}(\theta)| \leq \frac{B_g}{B_{\mathcal{F}}} \quad \text{por (4.4.7)}$$

O resultado é portanto verificado.

Estamos agora em condições de obter uma forma mais exata para a aproximação (4.3.6).

Teorema 4.1 :

Seja g um filtro satisfazendo as condições (4.4.6) e (4.4.7) e $\Gamma_{t,\omega}(\theta)$ sua função transferência (generalizada) com respeito à família semi-estacionária \mathcal{F} de largura característica $B_{\mathcal{F}}$.

Se para qualquer $\epsilon > 0$, g for tal que $B_g \leq \epsilon B_{\mathcal{F}}$, então para todo t, ω, θ ,

$$|A_t(\omega) - \Gamma_{t,\omega}(\theta)| < \epsilon$$

Prova :

Por (4.3.4) temos:

$$A_t(\omega) \Gamma_{t,\omega}(\theta) = \int_{-\infty}^{\infty} g(s) A_{t-s}(\omega) e^{-is\theta} ds$$

Substituindo $A_{t-s}(\omega)$ em termos de $dH_{\omega}(\alpha)$, obtemos

$$\begin{aligned} A_t(\omega) \Gamma_{t,\omega}(\theta) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(s) e^{-is\theta} e^{i(t-s)\alpha} dH_{\omega}(\alpha) ds \\ &= \int_{-\infty}^{\infty} e^{it\alpha} \Gamma(\theta+\alpha) dH_{\omega}(\alpha) \end{aligned}$$

De acordo com o lema 4.1, $e^{it\alpha} dH_{\omega}(\alpha)$ é uma pseudo-função- δ de ordem $\frac{B_g}{B_{\mathcal{F}}}$ com respeito a $\Gamma(\alpha)$. Logo, se o filtro g for escolhido de modo que, para $\epsilon > 0$ dado, $B_g \leq \epsilon B_{\mathcal{F}}$, então

$$\left| \int_{-\infty}^{\infty} e^{it\alpha} \Gamma(\theta+\alpha) dH_{\omega}(\alpha) - \Gamma(\theta) \int_{-\infty}^{\infty} e^{it\alpha} dH_{\omega}(\alpha) \right| < \epsilon$$

O resultado proposto é obtido observando-se que

$$\int_{-\infty}^{\infty} e^{it\alpha} dH_{\omega}(\alpha) = A_t(\omega).$$

5. Determinação de Espectros Evolutivos

Consideremos um processo semi-estacionário (X_t) com largura característica B_x tendo, em termos da família \mathcal{F}^* , a representação

$$X_t = \int_{-\infty}^{\infty} A_t^*(\lambda) e^{i\lambda t} dY_{\lambda}^* \quad \text{com} \quad E(|dY_{\lambda}^*|^2) = d\mu^*(\lambda) \quad (4.5.1)$$

Seja ainda g um filtro nas condições (4.4.6) e (4.4.7) com largura B_g . Para qualquer λ_0 , seja (\hat{X}_t) o processo definido por

$$\hat{X}_t = \int_{-\infty}^{\infty} g(s) X_{t-s} e^{-i\lambda_0(t-s)} ds \quad (4.5.2)$$

Usando (4.5.1) e considerando (4.3.3) podemos escrever

$$\hat{X}_t = \int_{-\infty}^{\infty} \Gamma_{t, \lambda + \lambda_0}^*(\lambda) A_t^*(\lambda + \lambda_0) e^{i\lambda t} dY_{\lambda + \lambda_0}^* \quad (4.5.3)$$

onde $\Gamma_{t, \omega}^*(\theta)$ é a função transferência generalizada de g com respeito à família \mathcal{F}^* .

Devido à ortogonalidade de (Y_λ^*) segue que

$$E(|\hat{X}_t|^2) = \int_{-\infty}^{\infty} |\Gamma_{t, \lambda + \lambda_0}^*(\lambda)|^2 |A_t^*(\lambda + \lambda_0)|^2 d\mu^*(\lambda + \lambda_0) \quad (4.5.4)$$

Em termos de

$$dF_t^*(\lambda) = |A_t^*(\lambda)|^2 d\mu^*(\lambda) \quad (4.5.4)$$

o espectro evolutivo de (X_t) com respeito à família \mathcal{F}^* , temos

Teorema 5.1

$$E(|\hat{X}_t|^2) = \int_{-\infty}^{\infty} |\Gamma(\lambda)|^2 dF_t^*(\lambda + \lambda_0) + O(\epsilon) \quad (4.5.5)$$

onde $O(\epsilon)$ designa um termo que pode ser feito arbitrariamente pequeno escolhendo-se B_g suficientemente pequena em relação a B_x .

Prova

Suponhamos o filtro g escolhido de forma a $B_g \subset \epsilon B_x$; como \mathcal{F}^* é uma família com largura característica igual ou arbitrariamente próxima de B_x , podemos escrever com base no teorema 4.1 anterior,

$$\Gamma_{t, \lambda + \lambda_0}^*(\lambda) = \Gamma(\lambda) + r(t, \lambda_0, \lambda)$$

$$\text{onde } |r(t, \lambda_0, \lambda)| < \frac{\epsilon}{|A_t^*(\lambda + \lambda_0)|}$$

Então de (4.5.4),

$$E(|\hat{X}_t|^2) = \int_{-\infty}^{\infty} |\Gamma(\lambda) + r(t, \lambda_0, \lambda)|^2 |A_t^*(\lambda + \lambda_0)|^2 d\mu^*(\lambda + \lambda_0)$$

$$= \int_{-\infty}^{\infty} |\Gamma(\lambda)|^2 |A_t^*(\lambda + \lambda_0)|^2 d\mu^*(\lambda + \lambda_0) + I_1 + I_2 + I_3$$

onde

$$I_1 = \int_{-\infty}^{\infty} \Gamma(\lambda) \overline{r(t, \lambda_0, \lambda)} |A_t^*(\lambda + \lambda_0)|^2 d\mu^*(\lambda + \lambda_0)$$

$$I_2 = \int_{-\infty}^{\infty} \overline{\Gamma(\lambda)} \cdot r(t, \lambda_0, \lambda) |A_t^*(\lambda + \lambda_0)|^2 d\mu^*(\lambda + \lambda_0)$$

$$I_3 = \int_{-\infty}^{\infty} |r(t, \lambda_0, \lambda)|^2 |A_t^*(\lambda + \lambda_0)|^2 d\mu^*(\lambda + \lambda_0)$$

Agora

$$|I_3| \leq \varepsilon^2 \int_{-\infty}^{\infty} d\mu^*(\lambda) = O(\varepsilon^2)$$

e

$$|I_2| \leq \varepsilon \int_{-\infty}^{\infty} |\Gamma(\lambda)| |A_t^*(\lambda + \lambda_0)| d\mu^*(\lambda + \lambda_0)$$

Para mostrarmos que $|I_2| = O(\varepsilon)$ basta provarmos que a integral à direita da desigualdade acima permanece finita quando $B_g \rightarrow 0$. Seja então $\Lambda = \{\lambda : |\Gamma(\lambda)| |A_t^*(\lambda + \lambda_0)| \leq 1\}$

Logo,

$$\int_{-\infty}^{\infty} |\Gamma(\lambda)| |A_t^*(\lambda + \lambda_0)| d\mu^*(\lambda + \lambda_0)$$

$$\leq \int_{\Lambda} d\mu^*(\lambda + \lambda_0) + \int_{\Lambda^c} |\Gamma(\lambda)|^2 |A_t^*(\lambda + \lambda_0)|^2 d\mu^*(\lambda + \lambda_0)$$

A integral sobre Λ é finita pois $d\mu^*(\lambda)$ é o espectro evolutivo em $t=0$, assim como a integral sobre Γ , pois $\Gamma(\lambda)$ é normalizada de forma a se obter

$$\int_{-\infty}^{\infty} |\Gamma(\lambda)|^2 d\lambda = 1$$

Procedendo analogamente para o termo I_1 , obtemos o resultado.

Se a medida $\mu^*(\lambda)$ for absolutamente contínua com respeito à medida de Lebesgue, poderemos escrever

$$dF_t^*(\lambda) = f_t^*(\lambda) d\lambda,$$

existindo para todo λ a função densidade espectral evolutiva $f_t^*(\lambda)$.

O teorema 5.1, em termos da $f_t^*(\lambda)$ resulta então

$$E(|\hat{X}_t|^2) \approx \int_{-\infty}^{\infty} |\Gamma(\lambda)|^2 f_t^*(\lambda + \lambda_0) d\lambda. \quad (4.5.6)$$

Notemos que a validade da aproximação acima depende somente da condição $B_g \ll B_{\mathcal{F}^*}$. É claro que para B_g fixada, (4.5.6) valerá ainda aproximadamente se a representação do processo (X_t) for em termos de uma família semi-estacionária \mathcal{F} cuja largura característica $B_{\mathcal{F}}$ satisfaça $B_{\mathcal{F}} \gg B_g$. Se $dF_t(\lambda) = f_t(\lambda) d\lambda$ for então o espectro evolutivo do processo com respeito à essa família, (4.5.6) também valerá aproximadamente se substituirmos f_t^* por f_t .

No entanto, deve ser lembrado que se tratarmos com uma

família \mathcal{F} genérica, o valor exato de $E(|\hat{X}_t|^2)$ será dado por

$$E(|\hat{X}_t|^2) = \int_{-\infty}^{\infty} |\Gamma_{t, \lambda + \lambda_0}^*|^2 dF_t(\lambda + \lambda_0) \quad (4.5.7)$$

Logo, o valor exato de $E(|\hat{X}_t|^2)$ é uma média em λ e t de $dF_t(\lambda)$ e como $E(|X_t|^2)$ independe da escolha da família \mathcal{F} , o valor desta média de $dF_t(\lambda)$ (sobre λ e t) deve também ser independente de \mathcal{F} .

Escrevendo (4.5.6), consideramos que o efeito do cálculo da média em relação a t é desprezível pois a condição $B_g \ll B_{\mathcal{F}}$ implica que a variação de $dF_t(\lambda)$ é muito lenta sob a variação do filtro g . De qualquer modo, o grau de precisão de (4.5.6) depende da razão $B_g/B_{\mathcal{F}}$.

Se, por exemplo, $B_g = 0$, isto é, $g(s) = \delta(s)$, então (4.5.6) será exata para qualquer \mathcal{F} e

$$E(|\hat{X}_t|^2) = E\left(\left|\int_{-\infty}^{\infty} \delta(s) \left[X_{t-s} e^{-i\lambda_0(t-s)}\right] ds\right|^2\right) =$$

$$E(|X_t e^{-i\lambda_0 t}|^2) = E(|X_t|^2)$$

Portanto,

$$E(|\hat{X}_t|^2) = \int_{-\infty}^{\infty} dF_t(\lambda) \quad (4.5.8)$$

Por outro lado, se

$$g(s) = \lim_{T_0 \rightarrow \infty} [g_{T_0}(s)] \quad \text{onde} \quad g_{T_0}(s) = \begin{cases} 1/\sqrt{T_0} & |s| \leq \frac{1}{2}T_0 \\ 0 & |s| > \frac{1}{2}T_0 \end{cases}$$

de modo que $B_g = \infty$, então

$$\begin{aligned} E(|\hat{X}_t|^2) &= \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} g(s) \frac{A_{t-s}(\lambda + \lambda_0)}{A_t(\lambda + \lambda_0)} e^{-is\lambda} ds \right|^2 |A_t(\lambda + \lambda_0)|^2 d\mu(\lambda + \lambda_0) \\ &= \lim_{T_0 \rightarrow \infty} \int_{-\infty}^{\infty} |G_{T_0, \lambda_0}(\lambda)|^2 d\mu(\lambda + \lambda_0) \end{aligned} \quad (4.5.9)$$

onde

$$G_{T_0, \lambda_0}(\lambda) = \frac{1}{\sqrt{T_0}} \int_{t - \frac{1}{2}T_0}^{t + \frac{1}{2}T_0} A_u(\lambda + \lambda_0) e^{iu\lambda} du$$

Notemos que $E(|\hat{X}_t|^2)$ é independente de t e se (X_t) for estacionário, então se reduzirá à definição clássica de espectro para processos estacionários. No entanto, (4.5.6) não é verificado para qualquer família.

Uma comparação entre (4.5.8) e (4.5.9) é interessante:

Em (4.5.8), $\int_{-\infty}^{\infty} dF_t(\lambda)$ é uma função apenas do espectro evolutivo no ponto t não envolvendo valores em outros pontos; porém não fornece informação alguma sobre a distribuição de $dF_t(\lambda)$ em termos de λ .

No entanto, se em (4.5.9) supusermos que para cada T_0 , $|G_{T_0, \lambda_0}(\lambda)|^2$ seja fortemente concentrada na região de $\lambda=0$ (como é geralmente o caso desde que admitamos que para todo λ , $A_t(\lambda)$ é uma função variando lentamente em t), então

$$\lim_{T_0 \rightarrow \infty} \int_{-\infty}^{\infty} |G_{T_0, \lambda_0}(\lambda)|^2 d\mu(\lambda + \lambda_0) \approx d\mu(\lambda_0) \left[\lim_{T_0 \rightarrow \infty} \int_{-\infty}^{\infty} |G_{T_0, \lambda_0}(\lambda)|^2 d\lambda \right]$$

Esta quantidade, sendo completamente independente de t , pode ser interpretada como uma forma de "média" sobre t dos valores de $dF_t(\lambda_0)$ para $-\infty < t < \infty$.

Com esta comparação, podemos observar que quanto mais precisamente tentarmos determinar $dF_t(\lambda)$ como uma função de t , menos precisamente a determinaremos como uma função de λ , e vice-versa. Esta característica sugere uma forma de

Princípio de Incerteza:

"Na determinação de espectros evolutivos não se pode obter simultaneamente um alto grau de resolução no domínio de t e no domínio de λ ."

Se for fixado um grau de resolução no domínio de λ , isto é, se for estabelecido um limite inferior para B_g , então obtem-se a máxima resolução possível no domínio de t considerando-se famílias com máxima largura característica pois, para uma particular família \mathcal{F} , a resolução neste domínio será determinada por $\frac{B_g}{B_{\mathcal{F}}}$.

6. Estimaco de Espectros Evolutivos

No que segue, ser considerada apenas a representao do processo (X_t) em termos de uma famlia $\mathcal{F}^* \in \mathcal{C}^*$, e quando se fizer referncia ao espectro evolutivo de (X_t) , sem aluso a qualquer particular famlia, este dever ser considerado como sendo $|A_t^*(\lambda)|^2 d\mu^*(\lambda)$, ou seja, o espectro evolutivo com respeito a \mathcal{F}^* .

Os astersticos em $A_t^*(\lambda)$, $d\mu^*(\lambda)$ e $dF_t^*(\lambda)$ sero omitidos portanto, sendo entendido que todas as funces so agora definidas com respeito  famlia \mathcal{F}^* .

Consideremos dada uma amostra de (X_t) para $0 \leq t \leq T_0$. Estudaremos o problema de estimar o espectro evolutivo, $dF_t(\lambda)$, para $0 \leq t \leq T_0$, a partir dessa amostra. Trataremos apenas do caso em que a medida $\mu(\lambda)$  absolutamente contnua com respeito  medida de Lebesgue, isto , para cada t, λ , $F_t'(\lambda) = f_t(\lambda)$, onde $f_t(\lambda)$, a funco densidade espectral evolutiva, existe para todo λ .

Seja g um filtro de largura B_g , nas condices (4.4.6) e (4.4.7) e para qualquer λ_0 faamos

$$U_t = \int_{t-T_0}^t g(s) X_{t-s} e^{-i\lambda_0(t-s)} ds. \quad (4.6.1)$$

Suporemos $B_g \ll B_x \ll T_0$ de modo que para $t \gg 0$, os limites na integral acima podem ser efetivamente substituidos por $(-\infty, \infty)$ tornando (U_t) idntico ao processo (\tilde{X}_t) definido em (4.5.2).

Sendo assim, pelo teorema 5.1,

$$E(|U_t|^2) = \int_{-\infty}^{\infty} |\Gamma(\lambda)|^2 f_t(\lambda + \lambda_0) d\lambda + O(B_g/B_x) \quad (4.6.2)$$

Se estabelecermos uma comparação entre a estimação de espectros evolutivos e a estimação de espectros para processos estacionários observaremos que para estes últimos pode-se também empregar a técnica descrita acima, porém, neste caso, a largura de faixa de $|\Gamma(\lambda)|^2$ é uma função de T_0 que tende a zero quando $T_0 \rightarrow \infty$. Para os espectros evolutivos, no entanto, a largura de faixa de $|\Gamma(\lambda)|^2$, que varia inversamente com B_g , é limitada pela restrição $B_g \ll B_x$.

Em outras palavras, como a escolha do filtro é tal que ele opera localmente sobre (X_t) e desta forma garante um alto grau de resolução no domínio de t , pode-se sacrificar algum grau de resolução no domínio de λ . Portanto, a fim de se estimar $f_t(\lambda)$ deve-se supor que sua largura de faixa seja substancialmente maior do que a largura de faixa de $|\Gamma(\lambda)|^2$. Neste caso $|\Gamma(\lambda)|^2$ é uma pseudo função- δ (de ordem a razão entre as larguras) com respeito a $f_t(\lambda)$ e podemos escrever, lembrando que $\int_{-\infty}^{\infty} |\Gamma(\lambda)|^2 d\lambda = 1$,

$$E(|U_t|^2) \sim f_t(\lambda_0) \quad (4.6.3)$$

Portanto, $|U_t|^2$ é um estimador não viciado aproximado de $f_t(\lambda_0)$. Contudo, se (X_t) for um processo estocástico normal

verifica-se, por um cálculo direto, que

$$\text{Var}(|U_t|^2) \sim \left[\int_{-\infty}^{\infty} |\Gamma(\lambda)|^2 f_t(\lambda + \lambda_0) dt \right]^2 (1 + \delta_{0, \lambda_0})$$

e sendo essa quantidade independente de T_0 , conclue-se que $|U_t|^2$ não deverá ser na prática um estimador muito útil de $f_t(\lambda_0)$.

No entanto podemos reduzir as flutuações amostrais "suavizando" os valores de $|U_t|^2$ em uma vizinhança de t ; assim procedendo, aumenta-se a precisão dos estimadores sacrificando-se algum grau de resolução no domínio de t .

Consideremos portanto uma função-peso $W_{T'_0}(t)$, dependendo do parâmetro T'_0 , que satisfaça:

(a) $W_{T'_0}(t) \geq 0$, todo t, T'_0

(b) $W_{T'_0}(t) \rightarrow 0$ quando $|t| \rightarrow \infty$ para todo T'_0

(c) $\int_{-\infty}^{\infty} W_{T'_0}(t) dt = 1$, todo T'_0

(d) $\int_{-\infty}^{\infty} [W_{T'_0}(t)]^2 dt < \infty$, todo T'_0

(e) Com $W_{T'_0}^*(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} W_{T'_0}(t) dt$, (4.6.4)

Admitiremos que exista uma constante C tal que

$$\lim_{T'_0 \rightarrow \infty} \left[T'_0 \int_{-\infty}^{\infty} |W_{T'_0}^*(\omega)|^2 d\omega \right] = C \quad (4.6.5)$$

Seja agora

$$V_t = \int_{-\infty}^{\infty} W_{T_0}(s) |U_{t-s}|^2 ds \quad (4.6.6)$$

Novamente, admitiremos que os valores de $W_{T_0}(s)$ decresçam suficientemente rápido de modo que a integral acima possa ser avaliada a partir de um comprimento finito de $|U_t|^2$. De (4.6.2) segue então que:

$$\begin{aligned} E(V_t) &\sim \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_{T_0}(s) f_{t-s}(\lambda + \lambda_0) |\Gamma(\lambda)|^2 ds d\lambda \\ &= \int_{-\infty}^{\infty} \hat{f}_t(\lambda + \lambda_0) |\Gamma(\lambda)|^2 d\lambda \end{aligned} \quad (4.6.7)$$

onde

$$\hat{f}_t(\lambda + \lambda_0) = \int_{-\infty}^{\infty} W_{T_0}(s) f_{t-s}(\lambda + \lambda_0) ds$$

Por (4.6.7) vemos que $E(V_t)$ é uma forma suavizada de $f_t(\lambda_0)$, suavizada em t e λ . Se, como anteriormente em (4.6.3), supuzermos que para cada t , $f_t(\lambda)$ seja suave comparada com $|\Gamma(\lambda)|^2$, então poderemos escrever:

$$E(V_t) \sim \hat{f}_t(\lambda_0)$$

de modo que V_t é um estimador não viciado aproximado do valor médio ponderado de $f_t(\lambda_0)$ em uma vizinhança de t .

Uma investigação quanto às propriedades amostrais de V_t é direta, porém longa. Os resultados que serão apresentados podem ser vistos em Priestley [21] e foram obtidos sob a conside-

ração de ser o processo (X_t) um processo estocástico normal.

Consideremos, para λ_0 qualquer,

$$U_t = \int_{-\infty}^{\infty} g(s) X_{t-s} e^{-i\lambda_0(t-s)} ds$$

Portanto, para ω e ν quaisquer, obtemos

$$E(|U_t|^2 | U_s|^2) = \iiint_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(u)g(v)g(m)g(n) e^{-iu\omega} e^{iv\omega} e^{-im\nu} e^{in\nu} E(X_{t-u} X_{t-v} X_{s-m} X_{s-n}) dudvdmn \quad (4.6.8)$$

Sendo (X_t) um processo normal,

$$E(X_{t-u} X_{t-v} X_{s-m} X_{s-n}) = R(t-u, t-v)R(s-m, s-n) + R(t-u, s-m)R(t-v, s-n) + R(t-u, s-n)R(t-v, s-m) \quad (4.6.9)$$

onde, como em (4.2.7),

$$R(p, q) = \int_{-\infty}^{\infty} e^{i\lambda(p-q)} A_p(\lambda) \overline{A_q(\lambda)} d\mu(\lambda) \quad (4.6.10)$$

Será assumido por simplicidade e sem perda de generalidade, já que $d\mu$ é absolutamente contínua com respeito a $d\lambda$, que $d\mu(\lambda) \equiv d\lambda$.

Substituindo-se (4.6.9) em (4.6.8) e considerando-se (4.6.10), obtem-se:

$$E(|U_t|^2 | U_s|^2) = E(|U_t|^2) \cdot E(|U_s|^2) + S_1 + S_2 \quad (4.6.11)$$

onde

$$S_1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(u)g(v)g(m)g(n)e^{-iu\omega}e^{iv\omega}e^{-imv}e^{inv} \\ \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i\theta(s-n-t+u)} A_{t-u}(\theta) \overline{A_s(\theta)} \cdot e^{-i\phi(s-m-t+v)} A_{t-v}(\phi) \overline{A_s(\phi)} d\theta d\phi \right) dudvdmdn$$

e S_2 é definido de modo semelhante permutando-se m com n .

Pode-se escrever S_1 como:

$$S_1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Gamma_{t,\theta}(\theta+\omega) \cdot \Gamma_{t,\phi}(\phi-\omega) \overline{\Gamma_{s,\phi}(\phi+v)} \overline{\Gamma_{s,\theta}(\theta-v)} \\ \times \left[A_t(\theta) A_t(\phi) \overline{A_s(\phi)} \overline{A_s(\theta)} e^{-i(s-t)(\theta+\phi)} \right] d\theta d\phi$$

onde, para quaisquer t, ω, θ ,

$$\Gamma_{t,\omega}(\theta) = \int_{-\infty}^{\infty} g(s) (A_{t-s}(\omega) / A_t(\omega)) e^{-is\theta} ds$$

Mas sob a condição $B_g \ll B_x$, pelo teorema 4.1, para todo t, ω, θ ,

$$\Gamma_{t,\omega}(\theta) \sim \Gamma(\theta)$$

Portanto

$$S_1 \sim \left[\int_{-\infty}^{\infty} e^{-i(s-t)\theta} A_t(\theta) \overline{A_s(\theta)} \Gamma(\theta+\omega) \overline{\Gamma(\theta-v)} d\theta \right] \\ \times \left[e^{-i(s-t)\phi} A_t(\phi) \overline{A_s(\phi)} \Gamma(\phi-\omega) \overline{\Gamma(\phi+v)} d\phi \right]$$

Considerando-se que $\Gamma(\lambda)$ é uma pseudo função- δ com respeito a $A_t(\lambda)$ (o que parece razoável se supor, visto que es tamos sob a afirmação de ser $|\Gamma(\lambda)|^2$ uma pseudo-função- δ com respeito a $|A_t(\lambda)|^2$), obtemos com essa aproximação,

$$S_1 \sim \frac{1}{4} \left[\overline{A_t(\omega)} A_s(\omega) + \overline{A_t(\nu)} A_s(\nu) \right] \left[A_t(\omega) \overline{A_s(\omega)} + A_t(\nu) \overline{A_s(\nu)} \right] \\ \times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i(s-t)(\theta+\phi)} \Gamma(\theta+\omega) \overline{\Gamma(\theta-\nu)} \Gamma(\phi-\omega) \overline{\Gamma(\nu+\phi)} d\theta d\phi \quad (4.6.12)$$

Da mesma forma pode-se encontrar a expressão análoga para S_2 .

Portanto, de (4.6.11), segue que

$$\text{Cov}(|U_t|^2, |U_s|^2) = S_1 + S_2 \quad (4.6.13)$$

Logo,

$$\text{Cov}(V_t, V_{t'}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_{T_0}(s) W_{T_0}(\nu) \text{Cov}(|U_{t-s}|^2, |U_{t'-\nu}|^2) ds d\nu = Q_1 + Q_2 \quad (4.6.14)$$

onde Q_1 e Q_2 correspondem às contribuições de S_1 e S_2 .

Utilizando (4.6.12) - (4.6.14) e depois de alguma redução, chega-se a

$$Q_1 \sim \frac{1}{4} \sum_{j=1}^4 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i(t'-t)(\theta+\phi)} K_t^{(j)}(\theta+\phi) \overline{K_{t'}^{(j)}(\theta+\phi)} \\ \times \Gamma(\theta+\omega) \overline{\Gamma(\theta-\nu)} \Gamma(\phi-\omega) \overline{\Gamma(\phi+\nu)} d\theta d\phi \quad (4.6.15)$$

onde

$$K_t^{(j)}(\theta) = \int_{-\infty}^{\infty} W_{T_0}(s) a_{t-s}^{(j)} e^{-is\theta} ds$$

e

$$a_t^{(1)} = |A_t(\omega)|^2, \quad a_t^{(2)} = A_t(\omega) \cdot \overline{A_t(\nu)}$$

$$a_t^{(3)} = A_t(\nu) \overline{A_t(\omega)}, \quad a_t^{(4)} = |A_t(\nu)|^2$$

Mas

$$\begin{aligned} \int_{-\infty}^{\infty} K_t^{(j)}(\theta) \overline{K_{t'}^{(j)}(\theta)} d\theta &= 2\pi \int_{-\infty}^{\infty} [W_{T_0}(s)]^2 a_{t-s}^{(j)} \overline{a_{t'-s}^{(j)}} ds \\ &= 2\pi (a_t^{(j)} \cdot \overline{a_{t'}^{(j)}}) \int_{-\infty}^{\infty} [W_{T_0}(s)]^2 ds \end{aligned}$$

Se $\tilde{K}^{(j)}(\theta)$ denotar a função $K_t^{(j)}(\theta) \cdot \overline{K_{t'}^{(j)}(\theta)}$ depois de normalizada de modo a possuir integral unitária, então

$$\begin{aligned} Q_1 &\sim \frac{\pi}{2} \left[\int_{-\infty}^{\infty} [W_{T_0}(s)]^2 ds \right] \left[\int_{-\infty}^{\infty} |\Gamma(\theta+\omega)|^2 |\Gamma(\theta+\nu)|^2 d\theta \right] \\ &\times \left[\sum_{j=1}^4 a_t^{(j)} \overline{a_{t'}^{(j)}} \int_{-\infty}^{\infty} e^{-i(t'-t)\theta} \tilde{K}^{(j)}(\theta) d\theta \right] \quad (4.6.16) \end{aligned}$$

Com cálculos semelhantes obtém-se a expressão análoga para Q_2 .

Portanto, para ω e ν não simultaneamente nulos, chega-se à

$$\text{Cov}[V_t, V_{t'}] \sim \frac{\pi}{2} \left[\int_{-\infty}^{\infty} [W_{T_0}(s)]^2 ds \right] \left[\int_{-\infty}^{\infty} |\Gamma(\theta+\omega)|^2 |\Gamma(\theta+\nu)|^2 d\theta \right] \sum_{j=1}^4 a_t^{(j)} \overline{a_{t'}^{(j)}} k^{(j)}(t'-t)$$

onde

$$k^{(j)}(s) = \int_{-\infty}^{\infty} e^{-is\theta} \tilde{K}^{(j)}(\theta) d\theta \quad (4.6.17)$$

Pode-se interpretar (4.6.17) como segue:

A covariância entre V_t e $V_{t'}$, será efetivamente nula

se

i) $|\omega \pm \nu|$ for suficientemente grande de modo que

$$\int_{-\infty}^{\infty} |\Gamma(\theta+\omega)|^2 |\Gamma(\theta+\nu)|^2 d\theta \sim 0$$

isto é, se $|\omega \pm \nu| \gg$ largura de faixa de $|\Gamma(\omega)|^2$; ou

ii) $|t'-t|$ for suficientemente grande de modo que $k^{(j)}(t'-t) \sim 0$ para todo j , isto é, se $|t'-t| \gg$ largura da função peso $W_{T'_0}$.

Em particular, para $t'=t$, $\omega=\nu=\lambda_0$

$$\text{Var}(V_t) \sim (1+\delta_{0,\lambda_0}) 2\pi \tilde{F}_t(\lambda_0) \left[\int_{-\infty}^{\infty} [W_{T'_0}(s)]^2 ds \right] \left[\int_{-\infty}^{\infty} |\Gamma(\lambda)|^4 d\lambda \right] \quad (4.6.18)$$

onde

$$\tilde{F}_t(\lambda_0) = \int_{-\infty}^{\infty} f_{t-s}^2(\lambda_0) [W_{T'_0}(s)]^2 ds / \int_{-\infty}^{\infty} [W_{T'_0}(s)]^2 ds .$$

Observando-se que

$$2\pi \int_{-\infty}^{\infty} [W_{T'_0}(t)]^2 dt = \int_{-\infty}^{\infty} |W_{T'_0}^*(\omega)|^2 d\omega$$

obtem-se, para valores grandes de T'_0 ,

$$\text{Var}(V_t) \sim (1+\delta_{0,\lambda_0}) \frac{C}{T'_0} \tilde{F}_t(\lambda_0) \int_{-\infty}^{\infty} |\Gamma(\lambda)|^4 d\lambda \quad (4.6.19)$$

Uma expressão aproximada para o vício do estimador V_t pode ser obtida de (4.6.7), onde

$$E(V_t) \sim \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{t-s}(\lambda + \lambda_0) W_{T_0}(s) |\Gamma(\lambda)|^2 d\lambda ds \quad (4.6.20)$$

Expandindo $f_{t-s}(\lambda + \lambda_0)$ em uma vizinhança de $f_t(\lambda_0)$ como uma série de Taylor nas variáveis s e λ , negligenciando potências de s e λ acima da de segunda ordem e supondo $W_{T_0}(s)$ e $|\Gamma(\lambda)|^2$ como funções simétricas de s e λ , respectivamente, ($|\Gamma(\lambda)|^2$ será sempre simétrica quando a função g for real), então de (4.6.20),

$$E(V_t) \sim f_t(\lambda_0) \left[1 + \frac{B_W^2}{2B_o(t, \lambda_0)} + \frac{B_\Gamma^2}{2B_f(t, \lambda_0)} \right] \quad (4.6.21)$$

onde

$$B_W = \left[\int_{-\infty}^{\infty} s^2 W_{T_0}(s) ds \right]^{1/2}$$

é uma medida da "largura da função W_{T_0} " ;

$$B_\Gamma = \left[\int_{-\infty}^{\infty} \lambda^2 |\Gamma(\lambda)|^2 d\lambda \right]^{1/2}$$

é uma medida da "largura de faixa da função $|\Gamma(\lambda)|^2$ " ;

$$B_o(t, \lambda) = \left| f_t(\lambda) / \frac{\partial^2 f_t(\lambda)}{\partial t^2} \right|^{1/2}$$

$$B_f(t, \lambda) = \left| f_t(\lambda) / \frac{\partial^2 f_t(\lambda)}{\partial \lambda^2} \right|^{1/2}$$

Para cada λ fixado, $B_o(t, \lambda)$ pode ser interpretada como a "largura de faixa de $f_t(\lambda)$ no ponto t ", com $f_t(\lambda)$ sendo observada como uma densidade espectral sobre o domínio de t , enquanto para cada t fixado, $B_f(t, \lambda)$ pode ser interpretada como a "largura de faixa de $f_t(\lambda)$ no ponto λ ", observada (no sentido clássico) como uma densidade espectral sobre λ .

Façamos

$$B_o(\lambda) = \inf_t \{B_o(t, \lambda)\} \quad e \quad B_f(t) = \inf_\lambda \{B_f(t, \lambda)\}$$

Considerando

$$B_o = \inf_\lambda \{B_o(\lambda)\} \quad , \quad B_f = \inf_t \{B_f(t)\}$$

podemos descrever B_o e B_f , respectivamente, como a largura de faixa de $f_t(\lambda)$ no domínio de t e no domínio de λ .

Portanto, por (4.6.21) podemos obter uma expressão aproximada para o vício de V_t como sendo:

$$v(V_t) = \left| E[V_t - f_t(\lambda_o)] \right| \sim f_t(\lambda_o) \left(\frac{B_W^2}{2B_o^2} + \frac{B_\Gamma^2}{2B_f^2} \right) \quad (4.6.22)$$

7. Considerações sobre o planejamento

Antes do estimador V_t ser calculado, é necessário se escolher a forma do filtro g e a forma da função peso W_{T_o} . Geralmente o filtro é escolhido como uma das "janelas" usuais da análise espectral de processos estacionários, lembrando que

$|\Gamma(\lambda)|^2$ e não $|\Gamma(\lambda)|$ corresponde à janela espectral. O filtro envolverá um parâmetro digamos, h , de modo que ajustando o valor de h , será possível se variar os valores de B_g e B_Γ . Do mesmo modo $W_{T'_0}$, poderá ser escolhida a partir da mesma coleção de janelas e pelo ajustamento de T'_0 torna-se possível variar B_W .

Por exemplo, se o filtro escolhido for

$$g(s) = \begin{cases} \frac{1}{2\sqrt{h\pi}} & , \quad |s| \leq h \\ 0 & |s| > h \end{cases} \quad (4.7.1)$$

então

$$|\Gamma(\lambda)|^2 = \frac{1}{\pi} \frac{\text{sen}^2 h\lambda}{h\lambda^2}$$

correspondendo à janela de Bartlett.

Por outro lado, pode-se escolher

$$W_{T'_0}(t) = \begin{cases} \frac{1}{T'_0} & , \quad -\frac{1}{2}T'_0 \leq t \leq \frac{1}{2}T'_0 \\ 0 & , \quad \text{complementar} \end{cases} \quad (4.7.2)$$

correspondendo à janela de Daniell; então

$$W_{T'_0}^*(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} W_{T'_0}(t) dt = \frac{\text{sen}(\frac{1}{2}T'_0 \omega)}{\frac{1}{2}T'_0 \omega}$$

$$e \lim_{T'_0 \rightarrow \infty} \left[T'_0 \int_{-\infty}^{\infty} \left| \frac{\text{sen} \frac{1}{2}T'_0 \omega}{\frac{1}{2}T'_0 \omega} \right|^2 d\omega \right] = 2\pi$$

Uma outra escolha poderá ser

$$W_{T'_0}(t) = \begin{cases} \frac{e^{-t/T'_0}}{T'_0} & , \quad t \geq 0 \\ 0 & , \quad t < 0 \end{cases} \quad (4.7.3)$$

de modo que

$$W_{T'_0}^*(\omega) = \frac{1}{1+i\omega T'_0}$$

e

$$\lim_{T'_0 \rightarrow \infty} \left[T'_0 \int_{-\infty}^{\infty} |W_{T'_0}^*(\omega)|^2 d\omega \right] = \pi$$

Depois da escolha das formas matemáticas das funções g e $W_{T'_0}$, surge o problema de como escolher os parâmetros h e T'_0 de modo que o estimador V_t possua certas propriedades desejadas.

Consideraremos agora tres conjuntos possíveis de condições que podemos desejar que V_t obedeça:

(1) "Se a resolução no domínio de λ for fixada"

Se for exigido que para cada t , V_t deva ter um determinado grau de resolução no domínio de λ , então o parâmetro h poderá ser escolhido de modo a se obter

$$B_{\Gamma/B_f} = k_1 \quad (4.7.4)$$

onde k_1 é alguma constante determinada. Dada B_f , (4.7.4) determinará h e conseqüentemente B_g .

O parâmetro T'_0 poderá agora ser escolhido de modo a minimizar $M(T'_0)$, o erro quadrático médio relativo de V_t , sujeito às condições $B_g \ll T'_0 \ll T_0$.

$$M(T'_0) = \frac{v^2(V_t) + \text{Var}(V_t)}{f_t^2(\lambda_0)} - \frac{B_W^4}{4B_0^4} + \frac{C}{T'_0} \int_{-\infty}^{\infty} |\Gamma(\lambda)|^4 d\lambda$$

Note-se que quando $T'_0 \uparrow$, $\text{Var}(V_t) \downarrow$, $v(V_t) \uparrow$ e quando $T'_0 \downarrow$, $\text{Var}(V_t) \uparrow$ e $v(V_t) \downarrow$.

(2) "Se a resolução no domínio de t for fixada".

Se for exigido um grau determinado de resolução no domínio de t , deve-se escolher T'_0 de modo a se ter

$$\frac{B_W}{B_0} = k_2 \tag{4.7.5}$$

onde k_2 é uma constante dada.

O parâmetro h poderá agora ser escolhido de modo a minimizar $M(h)$, o erro quadrático médio relativo, sujeito à condição $B_g \ll T'_0$,

$$M(h) \sim \frac{B_\Gamma^4}{4B_f^4} + \frac{C}{T'_0} \int_{-\infty}^{\infty} |\Gamma(\lambda)|^4 d\lambda$$

Note-se que quando $B_g \uparrow$, $v(V_t) \downarrow$, $\text{Var}(V_t) \uparrow$ e quando $B_g \downarrow$, $v(V_t) \uparrow$ e $\text{Var}(V_t) \downarrow$.

(3) "Se a resolução for fixada em ambos os domínios."

Neste caso, h e T'_0 deverão ser escolhidos de modo a se obter

$$\frac{B_\Gamma}{B_f} = k_1 \quad \text{e} \quad \frac{B_W}{B_0} = k_2$$

onde k_1 e k_2 são ambas dadas. O valor do erro quadrático médio relativo será agora determinado pelas condições acima.

(4) "Se nenhuma condição for exigida na resolução."

Neste caso, pode-se escolher h e T'_0 de modo que conjuntamente minimizem $M(h, T'_0)$, o erro quadrático médio relativo, sob as condições $B_g \ll B_x$, $T'_0 \ll T_0$; isto é, h e T'_0 podem ser escolhidos de modo a

$$M(h, T'_0) = \frac{1}{4} \left(\frac{B_W^2}{B_0^2} + \frac{B_\Gamma^2}{B_f^2} \right) + \frac{C}{T'_0} \int_{-\infty}^{\infty} |\Gamma(\lambda)|^4 d\lambda \quad (4.7.6)$$

ser um mínimo.

Ainda de Priestley [21] se obtêm as seguintes observações:

1) Não se pode escolher um valor arbitrariamente pequeno para k_1 pois o valor de h , finalmente escolhido, determinará o valor de B_g e deverá ser consistente com a condição $B_g \ll B_x$. Quando esta condição for mantida, o valor de h influenciará somente a resolução no domínio de λ de modo que as exigências para uma

dada medida de "precisão" em cada domínio podem ser tratadas separadamente. Quando a condição for transgredida, o valor de h afetará a resolução tanto no domínio de λ quanto no domínio de t (de acordo com o Princípio da Incerteza).

2) Os valores admissíveis de k_2 são do mesmo modo restritos pela condição $T'_0 \ll T_0$.

3) Em princípio, não existe dificuldade em se admitir os parâmetros h e T'_0 como sendo ambos seletivos. Então, se forem conhecidos os valores de $B_0(t, \lambda)$ e $B_f(t, \lambda)$ para todo t e λ , (4.7.4) e (4.7.5) poderão ser substituídos, respectivamente, por

$$\frac{B_\Gamma(t, \lambda)}{B_f(t, \lambda)} = k_1(t, \lambda) \quad \text{e} \quad \frac{B_W(t, \lambda)}{B_0(t, \lambda)} = k_2(t, \lambda)$$

Desta forma será possível se variar a largura de faixa do filtro g e a largura da função peso $W_{T'_0}$ quando t e λ variarem.

4) O Problema da escolha do parâmetro h no filtro g ocorre também na análise espectral de processos estacionários. Neste caso, é usual se pedir que h seja escolhido como uma função de T_0 de modo que $B_\Gamma \rightarrow 0$ quando $T_0 \rightarrow \infty$, a fim de que $\hat{f}(\lambda)$ seja um estimador consistente de $f(\lambda)$. Entretanto, no caso não estacionário, esta exigência é impossível, em primeiro lugar porque o valor de B_Γ está restrito pela condição $B_g \ll B_x$ e em segundo lugar, porque a noção de consistência será irrelevante ao tratarmos com uma amo

tra de uma única realização do processo, já que o efeito de aumentar T_0 é simplesmente de nos ser possível estimar $f_t(\lambda)$ em pontos-t futuros, não proporcionando informações adicionais que possam ser usadas para melhorar o estimador de $f_t(\lambda)$ em um valor t já considerado.

É necessário se salientar que as relações apresentadas nesta secção se prestam à determinação dos parâmetros h e T'_0 desde que as formas matemáticas das funções g e $W_{T'_0}$, já tenham sido escolhidas. O problema de escolher as formas dessas funções, no entanto, é muito mais complicado e mesmo para o caso estacionário, esse problema ainda não foi resolvido satisfatoriamente. Entretanto uma "figura de mérito" conveniente que poderia ser sugerida para determinar formas "ótimas" para g e $W_{T'_0}$, seria o erro quadrático médio relativo, $M(h, T'_0)$, dado por (4.7.6). Supondo-se que $M(h, T'_0)$ fosse minimizado com respeito a h e T'_0 e denotando-se o valor mínimo por $M(g, W_{T'_0})$, então para se determinar as formas ótimas procuradas, poder-se-ia tentar minimizar esse funcional em g e $W_{T'_0}$, lembrando que em $M(g, W_{T'_0})$ a quantidade $\int_{-\infty}^{\infty} |\Gamma(\lambda)|^4 d\lambda$ depende da forma do filtro g e a constante C depende da forma da função $W_{T'_0}$.

Um problema também a se considerar é o de testar se o processo com o qual estamos tratando pode ou não ser admitido como estacionário. Uma forma de teste é sugerida em analogia com os métodos da análise de variância. Então, se for assumido que

o processo é de fato estacionário, um estimador para sua função densidade espectral poderá ser

$$f_s^{(T_0)}(\lambda) = \frac{1}{T_0} \int_0^{T_0} |U_t|^2 dt$$

correspondendo a

$$W_{T_0'}(t) = \begin{cases} \frac{1}{T_0'} & , \quad 0 \leq t \leq T_0' \\ 0 & , \quad \text{complementar} \end{cases}$$

com $T_0' \equiv T_0$.

Se $f_t^{(T_0)}(\lambda)$ denotar um estimador de $f_t(\lambda)$ no ponto t , parece ser razoável se basear um teste de estacionariedade na função

$$\psi(\lambda) = \int_0^{T_0} (f_t^{(T_0)}(\lambda) - f_s^{(T_0)}(\lambda)) dt$$

8. O caso discreto

Exatamente do mesmo modo em que foi desenvolvida uma teoria espectral para processos não estacionários a parâmetro contínuo, pode-se fazê-lo para processos a parâmetro discreto.

Neste caso, um processo oscilatório (X_t) terá uma representação da forma

$$X_t = \int_{-\pi}^{\pi} e^{i\lambda t} A_t(\lambda) dY_\lambda \quad , \quad \text{com } t \in Z$$

onde (Y_λ) é um processo ortogonal em $(-\pi, \pi)$ com $E(|dY_\lambda|^2) = d\mu(\lambda)$ e para cada λ , a sequência $\{A_t(\lambda)\}$ possui uma transformada de Fourier generalizada cujo módulo tem um máximo absoluto na origem.

O espectro evolutivo em t com respeito à família de sequências $\mathcal{F} = \{e^{it\lambda} A_t(\lambda)\}$ é definido como

$$dF_t(\lambda) = |A_t(\lambda)|^2 d\mu(\lambda) \quad (-\pi \leq \lambda \leq \pi)$$

Considerando, da mesma forma, μ como absolutamente contínua, a função densidade espectral $f_t(\lambda)$ com respeito à família de largura característica máxima pode ser estimada pelo método descrito para o caso contínuo, agora com a função g substituída por uma sequência $\{g_s\}$ e a função peso $W_{T'_0}$ por uma sequência $\{W_{T'_0, t}\}$.

Com

$$U_t = \sum_{s=-\infty}^{\infty} g_s X_{t-s} e^{-i\lambda_0(t-s)}$$

e

$$V_t = \sum_{v=-\infty}^{\infty} W_{T'_0, v} |U_{t-v}|^2$$

obtemos:

$$E(V_t) \sim \int_{-\pi}^{\pi} \hat{f}_t(\lambda + \lambda_0) |\Gamma(\lambda)|^2 d\lambda$$

onde

$$\hat{f}_t(\lambda + \lambda_0) = \sum_{v=-\infty}^{\infty} W_{T'_0, v} f_{t-v}(\lambda + \lambda_0)$$

e

$$\Gamma(\lambda) = \sum_{s=-\infty}^{\infty} g_s e^{-i\lambda s}$$

Pode-se mostrar, então, que

$$\text{Var}(v_t) \sim [\tilde{f}_t(\lambda_0)] \left[\int_{-\pi}^{\pi} |W_{T_0}^*(\omega)|^2 d\omega \right] \left[\int_{-\pi}^{\pi} |\Gamma(\lambda)|^4 d\lambda \right] (1 + \delta_{0, \lambda_0})$$

onde

$$\tilde{f}_t(\lambda_0) = \frac{\sum_{v=-\infty}^{\infty} f_{t-v}^2(\lambda_0) (W_{T_0, v})^2}{\sum_{v=-\infty}^{\infty} (W_{T_0, v})^2}$$

e

$$W_{T_0}^*(\omega) = \sum_{v=-\infty}^{\infty} W_{T_0, v} e^{i\omega v}$$

APÊNDICE

-ANÁLISE DE UM PROCESSO ARTIFICIAL-

A fim de examinar a validade dos resultados sobre a estimação dos espectros evolutivos apresentados em IV-6 foram simuladas realizações de um processo (X_t) uniformemente modulado, artificial, gerado a partir do modelo (discreto)

$$X_t = e^{-\frac{1}{2}\left(\frac{t-270}{200}\right)^2} Y_t, \quad t=0,1,2,\dots$$

sendo (Y_t) o processo auto regressivo de segunda ordem

$$Y_t = 0,75Y_{t-1} - 0,5Y_{t-2} + Z_t$$

e (Z_t) uma família de variáveis aleatórias independentes tendo cada uma distribuição $N(0,100)$.

A função densidade espectral de (Y_t) é

$$f^Y(\lambda) = \frac{1}{2\pi} \left[\frac{0,8438 \text{ Var}(Y_t)}{1,2188 - 3,3750 \cos \lambda + 3 \cos^2 \lambda} \right], \quad 0 \leq \lambda \leq \pi,$$

onde $\text{Var}(Y_t) = (133,3333)^2$, apresentando pico em $\lambda \approx \frac{3}{10}\pi$ e com largura de faixa $B_f Y \approx \frac{\pi}{5}$.

Com respeito à família $\mathcal{F}_0 = (e^{-\frac{1}{2}\left(\frac{t-270}{200}\right)^2} e^{i\lambda t})$ o processo (X_t) tem função densidade espectral evolutiva

$$f_t(\lambda) = e^{-\left(\frac{t-270}{200}\right)^2} f^Y(\lambda) \tag{5.1}$$

de modo que, para cada t , $B_f(t) = \inf_{\lambda} \{B_f(t, \lambda)\} = B_{fY}$. Portanto,
 $B_f = \inf_t \{B_f(t)\} = B_{fY}$.

Para a família \mathcal{G}_0 , $dH(\theta)$ é independente de λ podendo-se escrever $dH_{\lambda}(\theta) = h(\theta)d\theta$ onde

$$h(\theta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{t-270}{200}\right)^2} e^{-i\theta t} dt = \frac{200}{\sqrt{2\pi}} e^{-\frac{200\theta^2}{2}} e^{-270i\theta}$$

Também $B_{\mathcal{G}_0}(\lambda)$ é aqui independente de λ obtendo-se portanto

$$B_{\mathcal{G}_0} = \frac{1}{\int_{-\infty}^{\infty} |\theta| |h(\theta)| d\theta} = 200 \sqrt{\frac{\pi}{2}}$$

de modo que B_x , a largura característica do processo (X_t) , é tal que

$$B_x \geq 200 \sqrt{\frac{\pi}{2}}$$

Consideramos

$$g(u) = \begin{cases} \frac{1}{2\sqrt{h\pi}} & |u| \leq h \\ 0 & |u| > h \end{cases} \quad (5.2)$$

obtendo portanto

$$|\Gamma(\lambda)|^2 = \frac{1}{\pi h} \frac{\text{sen}^2 h\lambda}{\lambda^2} \quad (5.3)$$

e como "largura" do filtro g ,

$$B_g = \int_{-\infty}^{\infty} |u| |g(u)| du = \sqrt{\frac{h^3}{4\pi}}$$

$$\text{Ainda, } \frac{B_g}{B_x} \leq \frac{1}{200\pi} \sqrt{\frac{h^3}{2}}$$

De acordo com a definição dada em IV-6, B_Γ será infinita quando o filtro g for o considerado em (5.2), implicando que a expressão aproximada obtida para o vício em (4.6.22) não é suficientemente precisa para um filtro dessa forma. No caso discreto, porém, $\Gamma(\lambda)$ é definida apenas em $(-\pi, \pi)$ e B_Γ é sempre finita.

A razão principal de se estabelecer os resultados em termos de B_Γ é que, em geral, B_Γ é uma medida da largura de faixa de $|\Gamma(\lambda)|^2$. Porém, para o grau de aproximação a ser usado, pode-se igualmente bem usar outra medida para essa largura de faixa, ou seja, a distância entre os pontos médios na curvatura principal de $|\Gamma(\lambda)|^2$. Assim, para $|\Gamma(\lambda)|^2$ dada em (5.3), obteve-se $B_\Gamma' = \frac{\pi}{h}$.

Fixamos a resolução no domínio de λ e escolhemos primeiramente $k_1 = \frac{1}{2}$ de modo que $B_\Gamma' = \frac{1}{2} B_f$. Ou seja, escolhemos h tal que $\frac{\pi}{h} = \frac{1}{2} \frac{\pi}{5}$; portanto, $h=10$.

Com esse valor de h ,

$$\frac{B_g}{B_x} \leq \frac{1}{200\pi} \sqrt{\frac{10^3}{2}} \approx 0,035$$

Consideramos como função peso:

$$W_{T'_0}(t) = \begin{cases} \frac{1}{T'_0} & , \quad |t| \leq \frac{T'_0}{2} \\ 0 & , \quad |t| > \frac{T'_0}{2} \end{cases} \quad (5.4)$$

Então,

$$B_W = \left[\int_{-\infty}^{\infty} t^2 W_{T'_0}(t) dt \right]^{1/2} = \frac{T'_0}{\sqrt{12}}$$

Ainda,

$$W_{T'_0}^*(\omega) = \int_{-\infty}^{\infty} W_{T'_0}(t) e^{-i\omega t} dt = \frac{\text{sen}(\omega T'_0/2)}{\omega T'_0/2}$$

e

$$\lim_{T'_0 \rightarrow \infty} \left[T'_0 \int_{-\infty}^{\infty} \frac{\text{sen}^2(\omega T'_0/2)}{(\omega T'_0/2)^2} d\omega \right] = 2\pi$$

Os estimadores V_t de $f_t(\lambda)$ foram obtidos através da equação análoga \tilde{a} (4.6.6) para o caso discreto com $W_{T'_0}$ dada por (5.4), U_t dado por (4.6.1) e $g(u)$ da forma (5.2).

Sendo assim,

$$\frac{\text{Var}(V_t)}{\tilde{f}_t(\lambda_0)} \sim \frac{2h}{3T'_0} = \frac{20}{3T'_0}$$

Para a determinação de T'_0 , consideramos o erro quadrático médio relativo de V_t , $M(T'_0)$, sujeito às condições

$B_g \ll T'_0 \ll T_0$. Então

$$M(T'_0) \sim \frac{B_W^4}{4B_0^4} + \frac{2\pi}{T'_0} \int_{-\infty}^{\infty} |\Gamma(\lambda)|^4 d\lambda = \frac{T_0'^4}{576B_0^4} + \frac{2h}{3T'_0}$$

Minimizando essa quantidade com respeito a T'_0 , obteve-se

$$T'_{0\min} = \sqrt[5]{960B_0^4}$$

Como B_0 mede a "largura" de $f_t(\lambda)$ quando observada como uma função de t , considerando a expressão (5.1), tomamos como um valor aproximado de B_0 o desvio padrão de uma distribuição normal com variância $\frac{200^2}{2}$. Isto é, tomamos $B_0 \approx \frac{200}{\sqrt{2}}$. Com este valor obteve-se $T'_0 \approx 208$.

Dessa forma, o erro relativo padrão de V_t , sendo $\frac{20}{3T'_0}$, com os valores de $h=10$ e $T'_0 \approx 208$ tornou-se aproximadamente 18%.

A fim de observar o comportamento dos estimadores V_t para outros valores de h , com a mesma análise foram escolhidos novos valores para a constante k_1 , ou seja, $k_1 = \frac{1}{3}$ e $k_1 = \frac{26}{32}$. Para o primeiro caso obteve-se $h=15$ e $T'_0 \approx 224$ com erro relativo padrão de V_t de aproximadamente 21%. Para o segundo caso os valores obtidos para h e T'_0 foram respectivamente (aproximadamente) 6 e 188 obtendo-se o erro relativo padrão de V_t como aproximadamente 14,6%

Convem observar que embora deva ser válido que $B_\Gamma < B_f$, o valor da constante $k_1 = \frac{B_\Gamma}{B_f}$ não pode ser escolhido arbitrariamente pequeno porque o valor de h então obtido determinará o valor de B_g e deverá ser consistente com a condição $B_g \ll B_x$; ainda, mantida esta condição o valor de h influenciará apenas a resolução no domínio de λ .

O programa apresentado a seguir tem como objetivo de terminar os estimadores V_t para a função densidade espectral do processo (X_t) considerado; são obtidos gráficos que apresentam as formas teórica e estimada de $f_t(\lambda)$ para $t=140, 200, 260, 320, 380$ e 270 , respectivamente, com $0 \leq \lambda \leq \pi$.

Em cada gráfico, a curva contínua corresponde à forma teórica e a curva tracejada à forma estimada. Para cada t considerado, a curva tracejada na figura indicada por (1) corresponde ao estimador V_t obtido com $h=15$ e $T'_0 \approx 224$; a indicada por (2) corresponde ao estimador obtido com $h=10$ e $T'_0 \approx 208$ e a curva tracejada na figura indicada por (3) corresponde ao estimador obtido para $h=6$ e $T'_0 \approx 188$.

Os gráficos obtidos para $t=270$ mostram o estimador da função densidade espectral do processo estacionário auto-regressivo (Y_t) cuja forma é exatamente a mesma usada para o processo não estacionário (X_t) .

```

// JOB T
LOG DRIVE' 0000 CART SPEC 0016 CART AVAIL 0016 PHY DRIVE 0000
V2 M10 ACTUAL 32K CONFIG 32K
// * C.P.D. - E.E.S.C. - U.S.P.
// FOR
*LIST SCURCE PROGRAM
SUBROUTINE AUTOC(Y,N,L,R)
REAL MED
DIMENSION Y(630),R(3)
MED=0.0
IF(N-L)50,50,100
50 R(1)=0.0
RETURN
100 DC 110 I=1,N
110 MFD=MED+Y(I)
FN=N
MED=MED/FN
DO 130 J=1,L
NJ=N-J+1
SOMA=C.0
DC 120 I=1,NJ
IJ=I+J-1
120 SOMA=SOMA+(Y(I)-MED)*(Y(IJ)-MED)
FNJ=NJ
130 R(J)=SOMA/FNJ
RETURN
END

CORE REQUIREMENTS FOR AUTOC
COMMON 0 VARIABLES 20 PROGRAM 166
RELATIVE ENTRY POINT ADDRESS IS 0018 (HEX)
END OF COMPILATION
// DUP
*STORE WS UA AUTOC

// FOR
*IOCS(CARD,1132PKINTER,TYPEWRITER,KEYBOARD,DISK,PLOTTER,1403PRINTER)
*LIST SCURCE PROGRAM
INTEGER T,T1,T2
REAL MED,LBDA
DIMENSION Y(630),X(630),FWXT(301),R(3),AUTOC(3),U2(600),V(51)
ALFA=0.75
BETA=-0.50
IX=371
MED=0.
DP=100.
YBARR=G.
AUX=0.
IT1=630
Y(1)=0.
X(1)=0.
CALL GAUSS(IX,DP,MED,Z)
Y(2)=Z+BETA*Y(1)
YBARR=YBARR+Y(2)
ET=((2.-270.)/200.)*2.)*(-1./2.)
CT=EXP(ET)
X(2)=ET*Y(2)
DC 2 T=3,IT1
CALL GAUSS(IX,DP,MED,Z)
T1=T-1
T2=T-2
Y(T)=Z+ALFA*Y(T1)+BETA*Y(T2)
YBARR=YBARR+Y(T)
ET=((FLCAT(T1)-270.)/200.)*2.)*(-1./2.)
ET=EXP(ET)
X(T)=ET*Y(T)
2 CONTINUE
WRITE(5,4)
4 FORMAT(1H1,' TEMPO',5X,'Y(I)',14X,'X(I)',/)
DC 3 I=1,IT1
IT=I-1
WRITE(5,5)IT,Y(I),X(I)
5 FORMAT(2X,I4,4X,F10.5,4X,F10.5)
3 CONTINUE
YBARR=YBARR/FLOAT(IT1)
DC 6 I=1,IT1
6 AUX=AUX+(Y(I)-YBARR)**2.
S2Y=AUX/(FLCAT(IT1)-1.)
S=SQR(S2Y)
WRITE(5,7)YBARR,S
7 FORMAT(1X,////,' MEDIA DE Y(I)=' F11.5,' DESVIO PADRAO='F11.5,/)
DO 20 INUM=1,2
READ(2,800)I13
800 FORMAT(I3)

```

```

YT=FLOAT(IT3)
W=0.
A1=(133.3333**2.)/(2.*3.14159)
C1=COS(W)
C2=COS(W)**2.
A2=(0.8438)/(1.2188-3.3750*C1+3.0*C2)
FWY=A1*A2
YE=((YT-270.)/200. )**2.)*(-1./2.)
YET=EXP(YE)
YET=(YET)**2.
FWXT(1)=YET*FWY
DO 30 I=1,300
W=W+3.14159/300.
C1=COS(W)
C2=COS(W)**2.
A2=(0.8438)/(1.2188-3.3750*C1+3.0*C2)
FWY=A1*A2
I2=I+1
FWXT(I2)=YET*FWY
30 CONTINUE
W=0.
CALL SCALF(4.,1.,0.,0.)
CALL FGRID(0,0.,0.,1./16.,18)
CALL FGRID(1,0.,0.,1./2.,6)
CALL FPLCT(-2,0.,FWXT(1)/2000.)
DO 300 I=2,301
W=W+3.14159/300.
CALL FPLCT(0,W/3.14159,FWXT(I)/2000.)
300 CALL FPLCT(1,2.,0.)
N=630
L=3
CALL AUTOC(Y,N,L,R)
DC 445 I1=1,L
I=I1-1
WRITE(5,446)I,R(I1)
446 FORMAT(1X,'AUTOCOVARIANCIA PARA L='I3,' ='E14.7,/)
DC 447 I1=1,L
I=I1-1
AUTCO(I1)=R(I1)/R(1)
447 WRITE(5,448)I,AUTCO(I1)
448 FCRMAT(1X,'AUTOCORRELACAO PARA L='I3,' ='F8.5,/)
ALFA2=(AUTCO(2))*(1.-AUTCO(3))/(1.-AUTCO(2)**2.)
BETA2=(AUTCO(3)-AUTCO(2)**2.)/(1.-AUTCO(2)**2.)
WRITE(5,449)ALFA2,BETA2
449 FORMAT(1X,'ESTIMATIVA PARA OS PARAMETROS',/, ' ALFA='F8.5,3X,'BETA='
I'F8.5,///)
DC 699 K4=1,51
K3=K4-1
LBDA=K3*(3.1415/50.)
DC 700 K5=1,225
IV1=-113+K5
INDC=IT3-IV1
SCMA1=0.
SCMA2=0.
DC 701 IU=1,31
IUA=-11+IU
IND1=INDC-IUA+1
IND2=INDC-IUA
701 SCMA1=SCMA1+X(IND1)*COS(LBDA*IND2)
SCMA2=SCMA2+X(IND1)*SIN(LBDA*IND2)
PART1=SCMA1/(2.*SQRT(15.*3.1416))
PART2=SCMA2/(2.*SQRT(15.*3.1416))
U2(INDC)=PART1**2.+PART2**2.
700 CONTINUE
SCMA=
IT2=IT3-112
IT4=IT3+112
DC 702 I=IT2,IT4
702 SCMA=SCMA+U2(I)
V(K4)=SCMA/224.
699 CONTINUE

CALL SCALF(4.,1.,0.,0.)
CALL FGRID(0,0.,0.,1./16.,18)
CALL FGRID(1,0.,0.,1./2.,6)
W=0.
CALL FPLCT(-2,0.,V(1)/2000.)
DC 703 I=2,51
W=W+3.14159/50.
703 CALL FPLCT(0,W/3.14159,V(I)/2000.)
CALL FPLCT(1,2.,0.)
20 CONTINUE
CALL EXIT
END

```

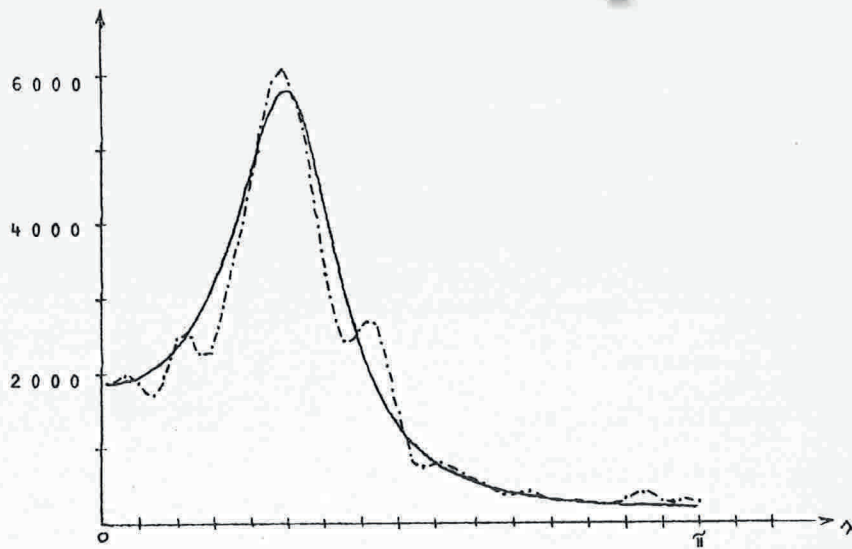
FEATURES SUPPORTED
IGCS

CORE REQUIREMENTS FOR
COMMON 0 VARIABLES 4548 PROGRAM 1326

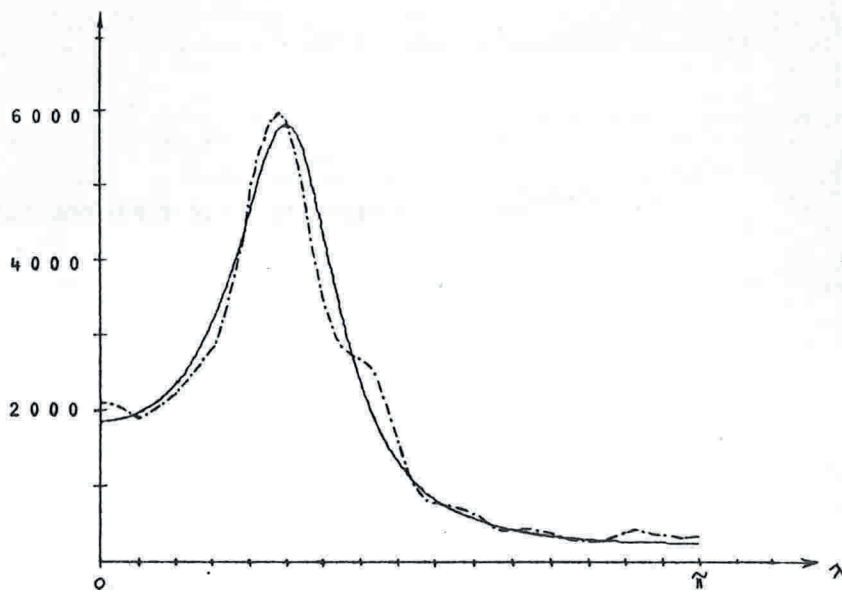
END OF COMPILATION

// XEQ

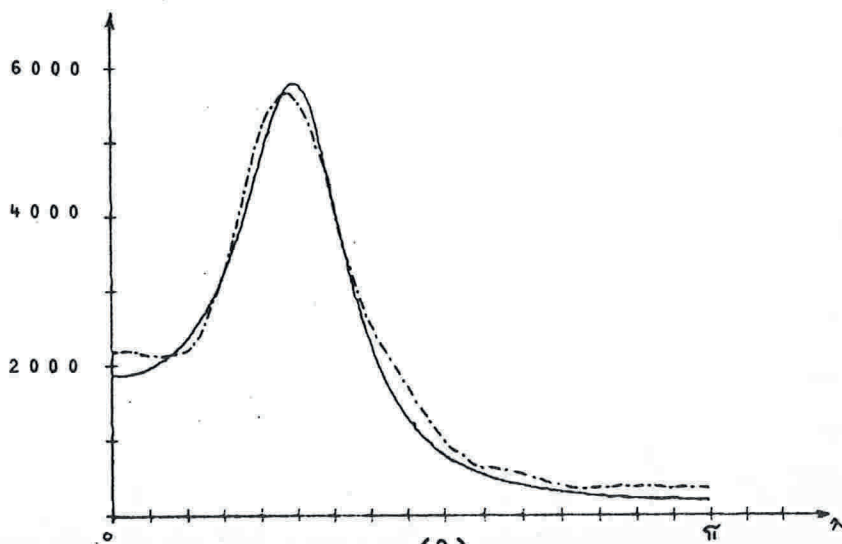
Espectros teórico e estimado para $t=140$



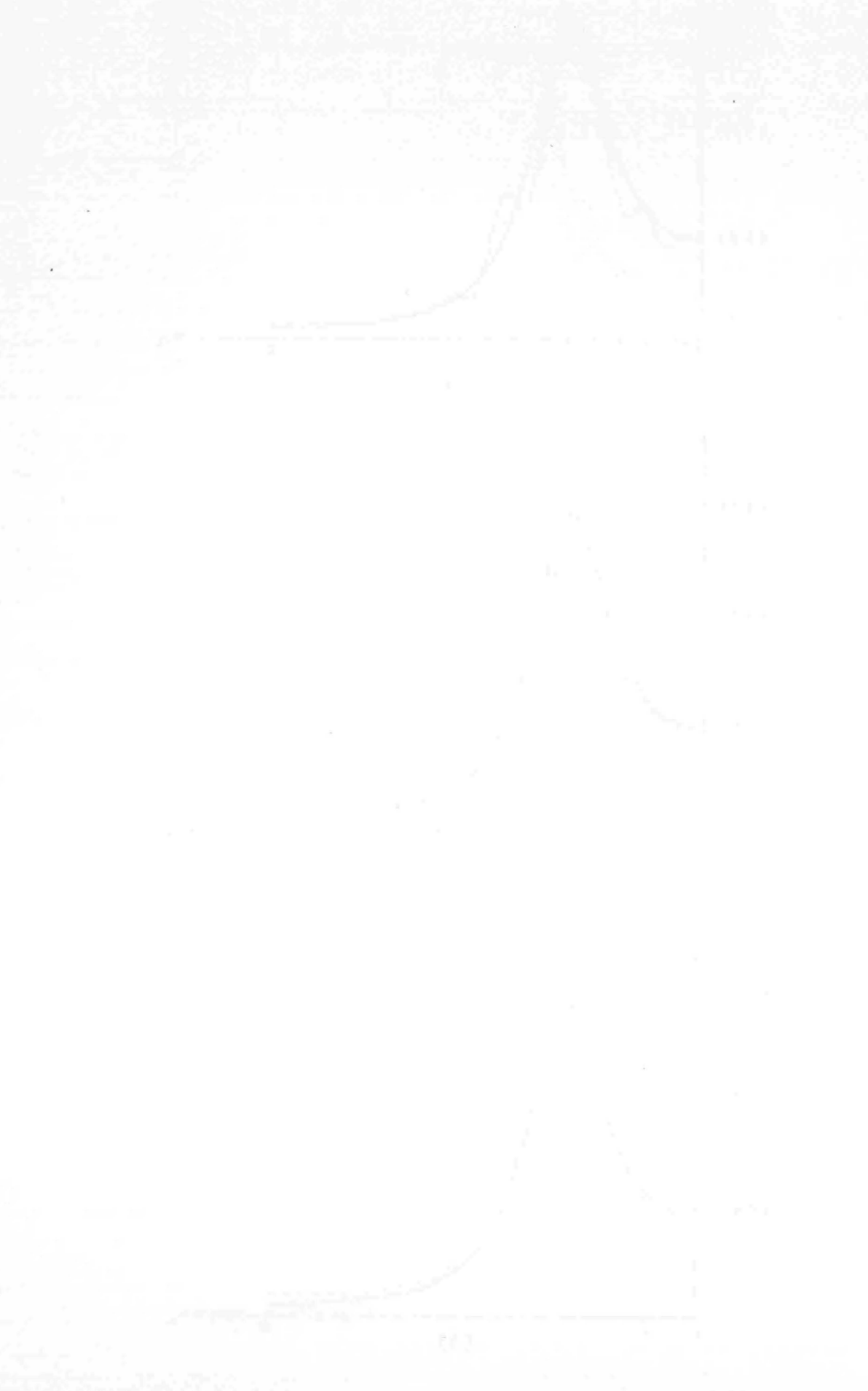
(1)



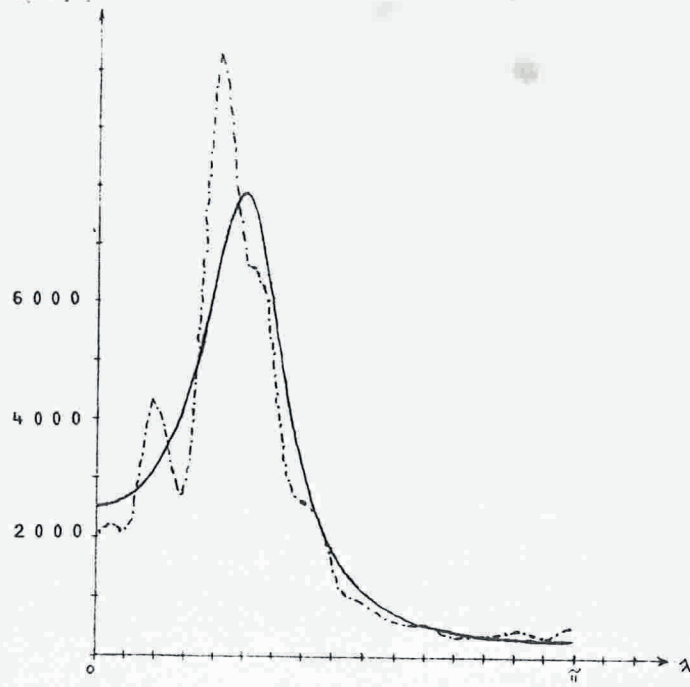
(2)



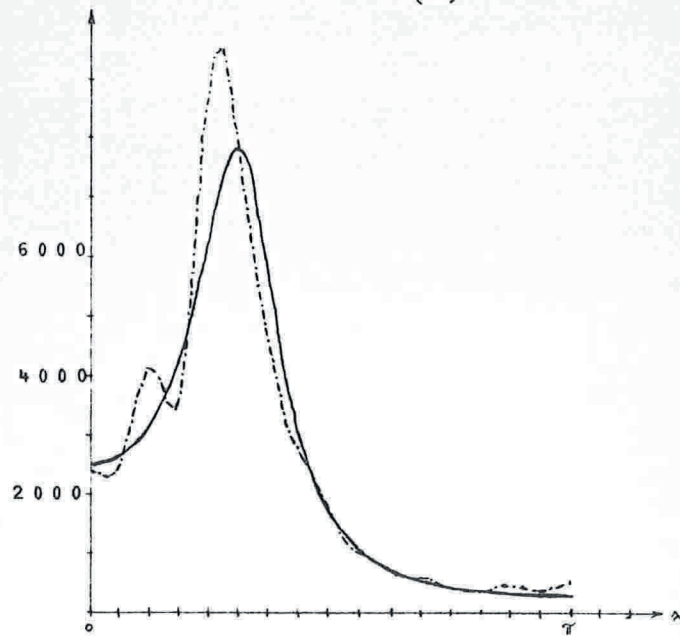
(3)



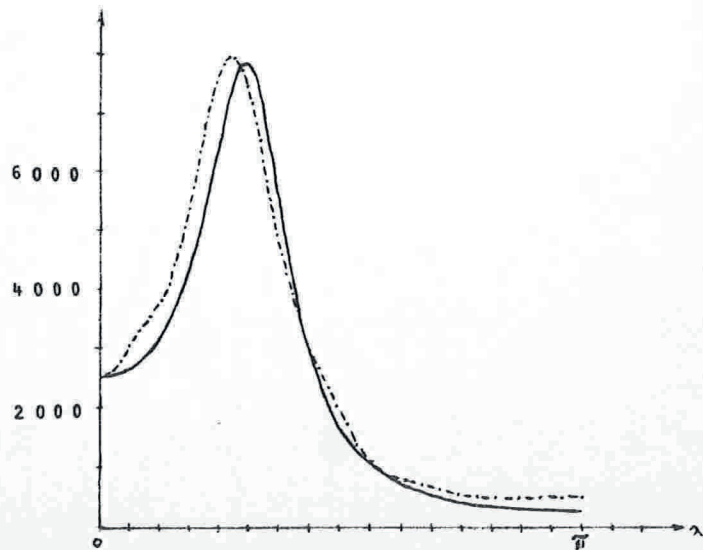
Espectros teóricos e estimado para $t=200$



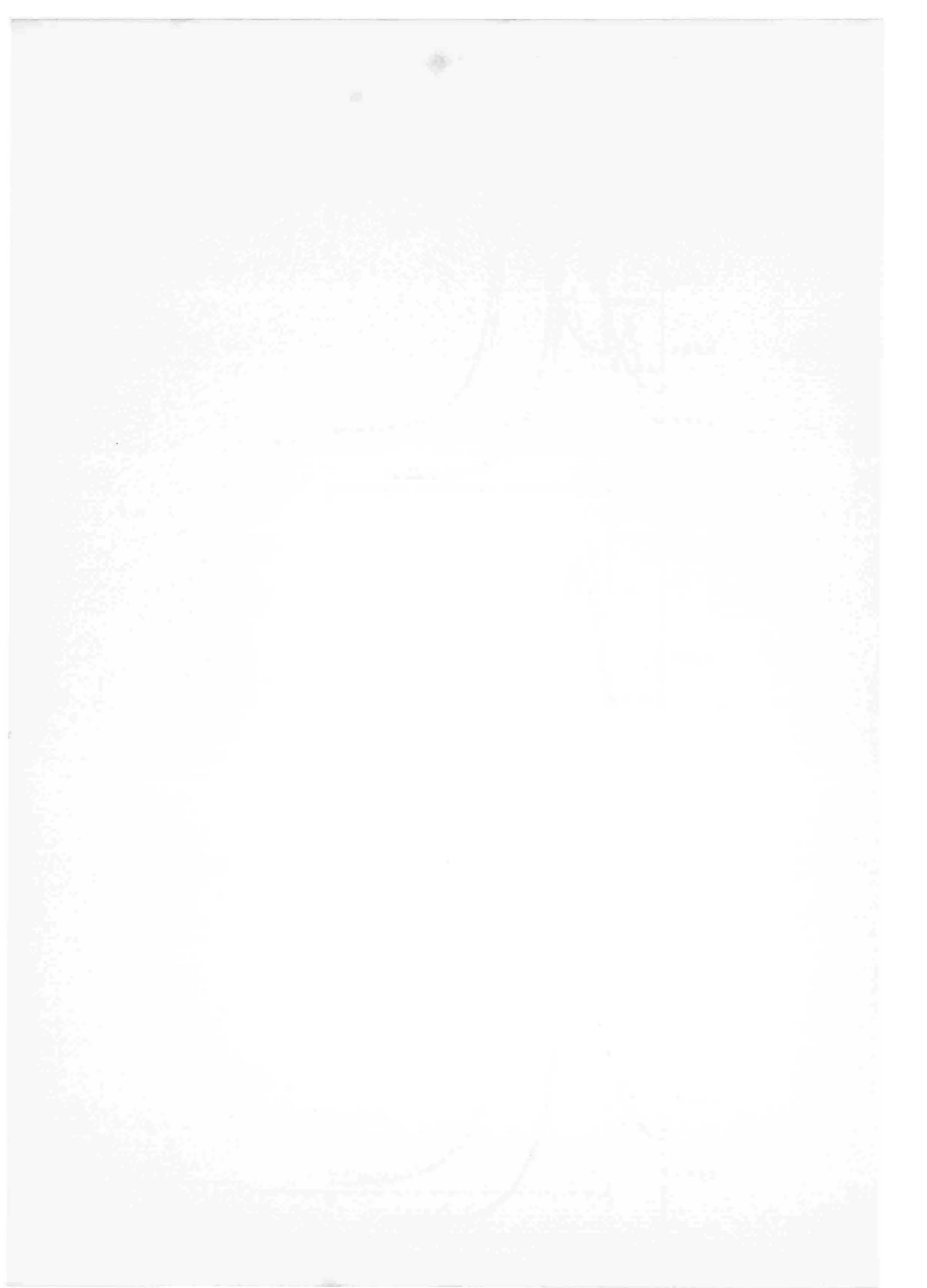
(1)



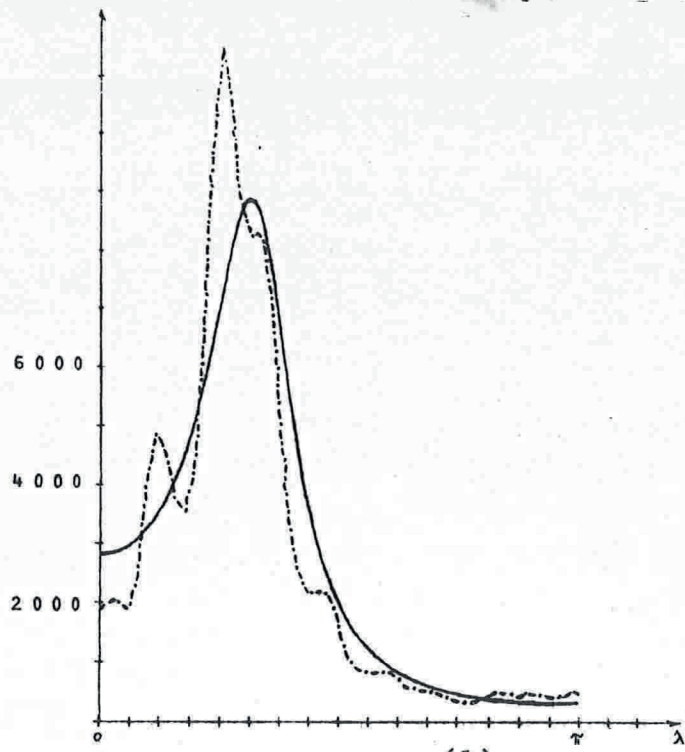
(2)



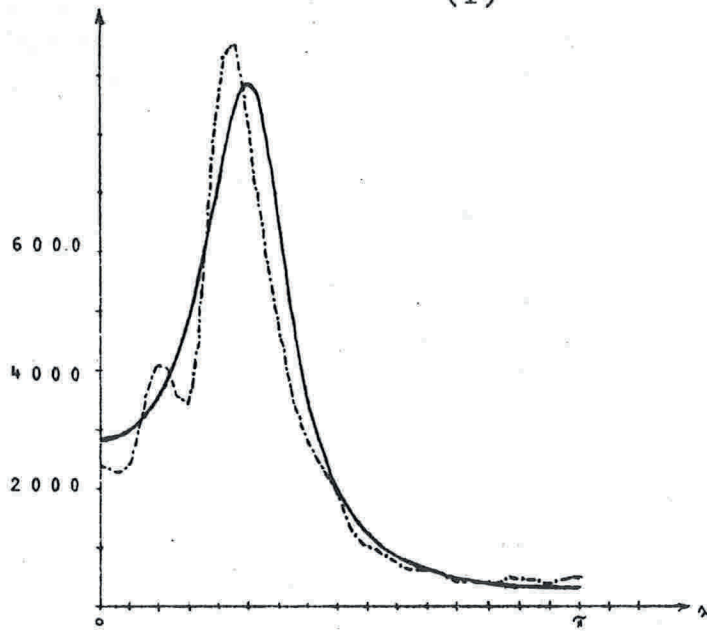
(3)



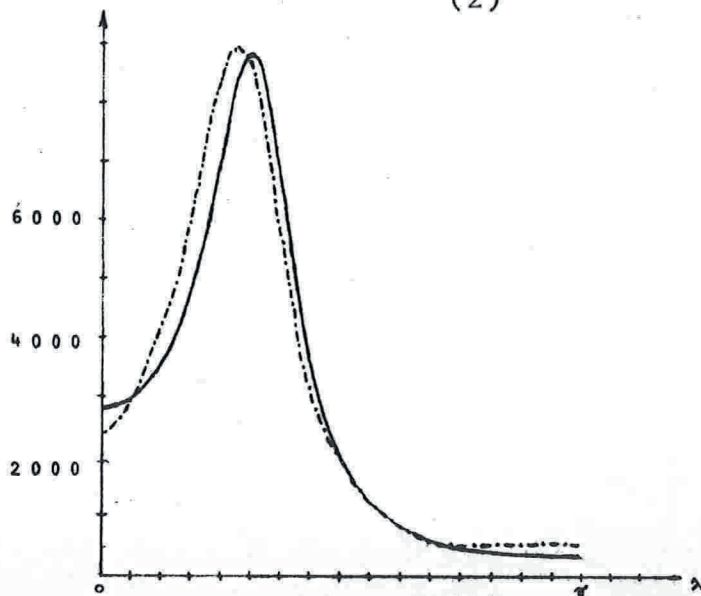
Espectros teórico e estimado para $t=260$



(1)



(2)



(3)



(a)

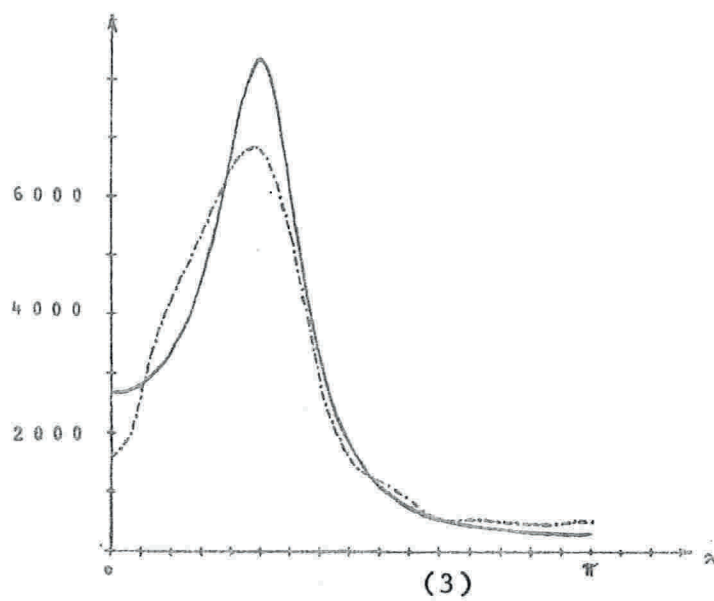
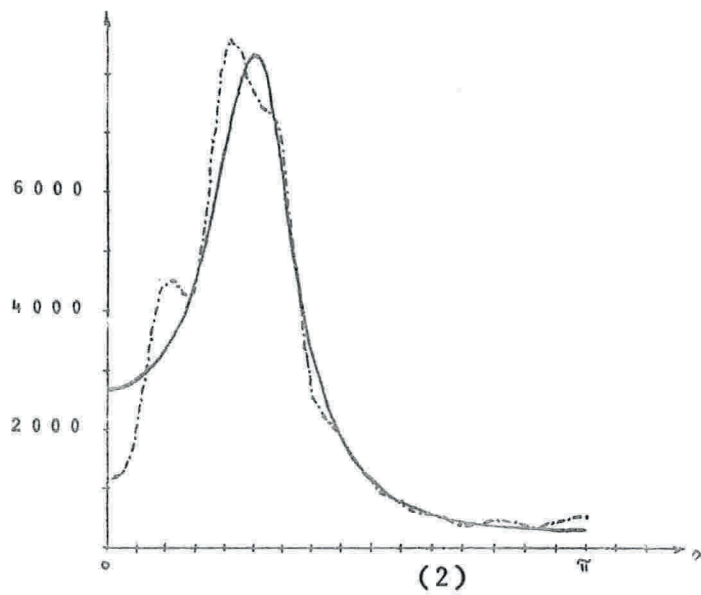
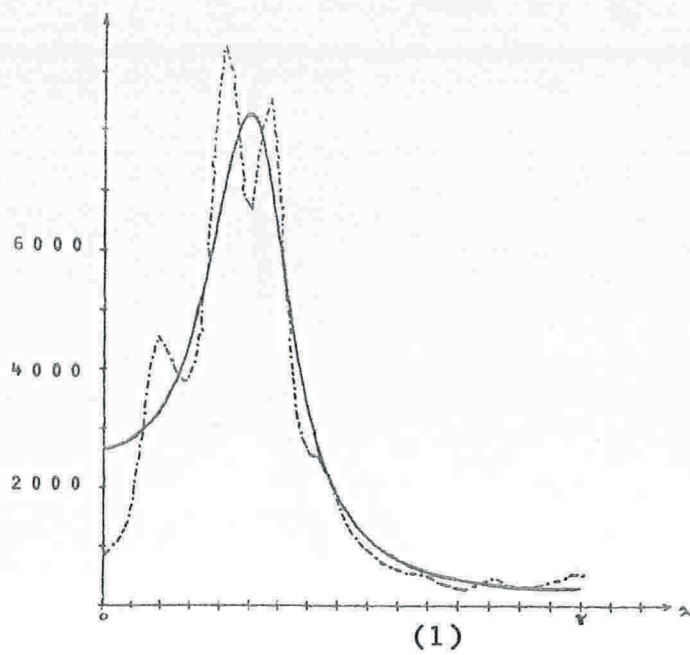


(b)



(c)

Espectros teórico e estimado para $t=320$





(1)

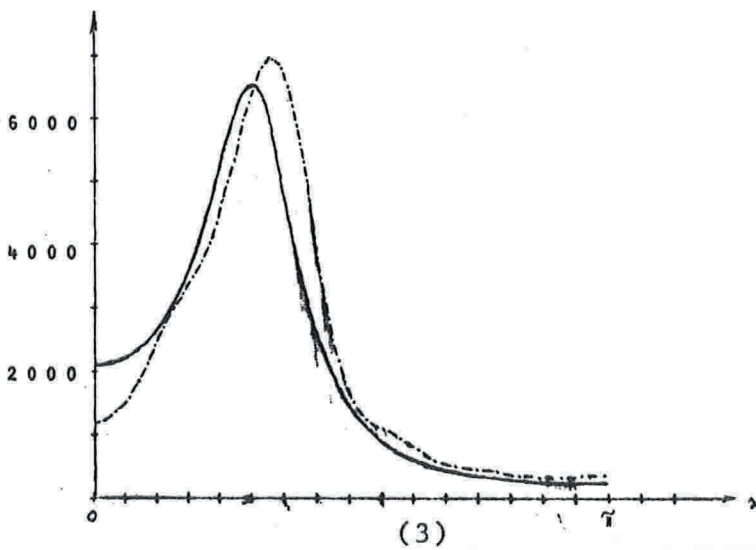
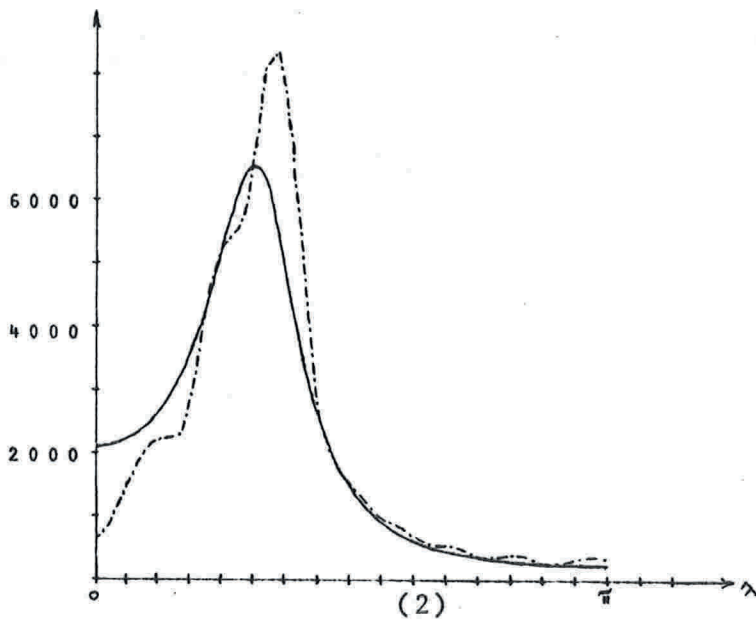
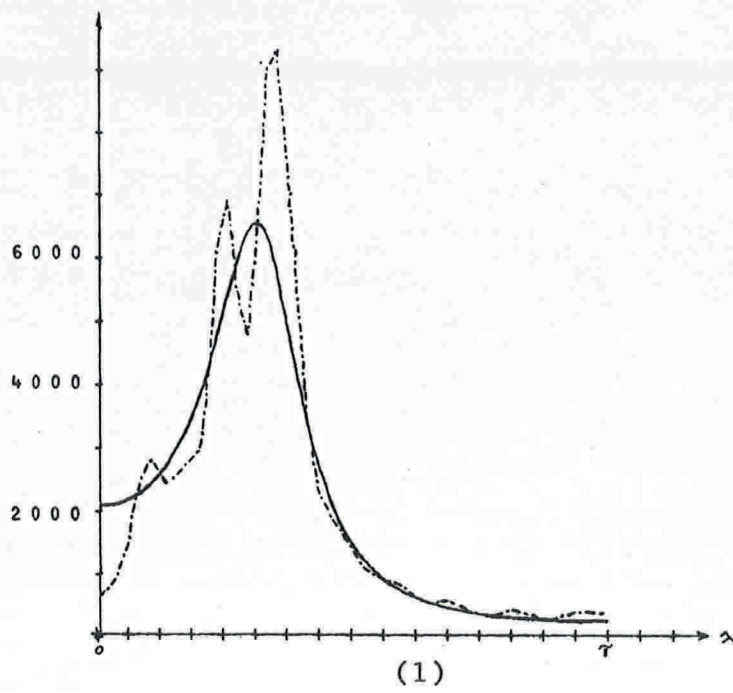


(2)



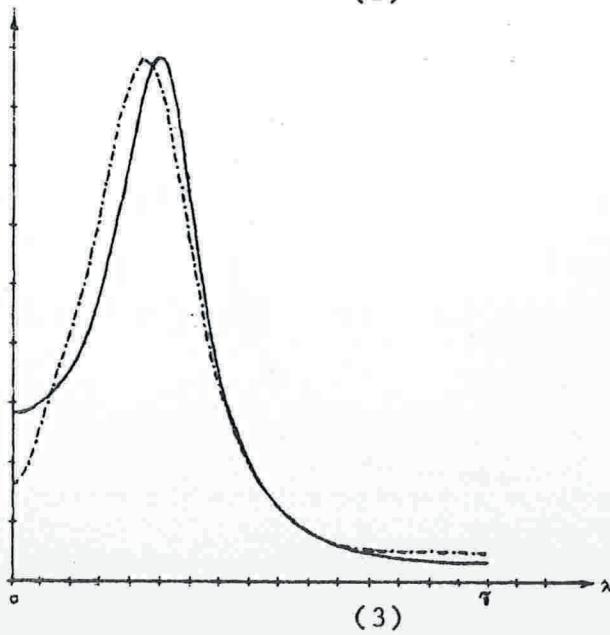
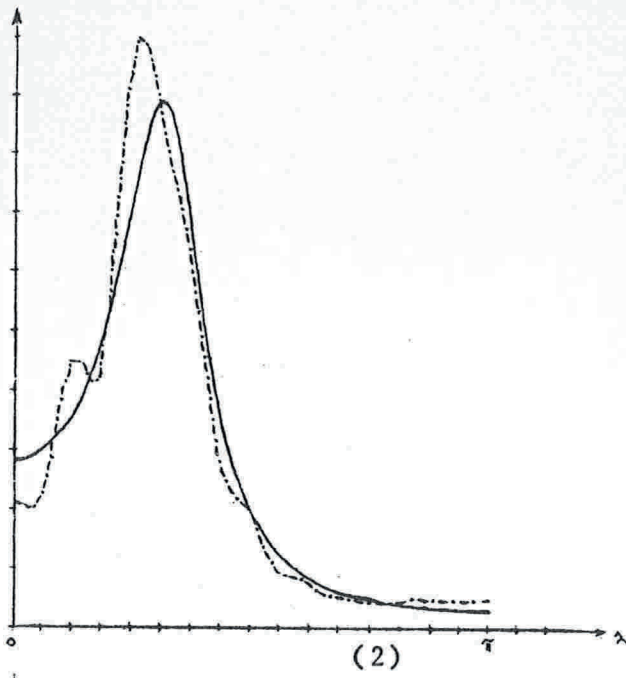
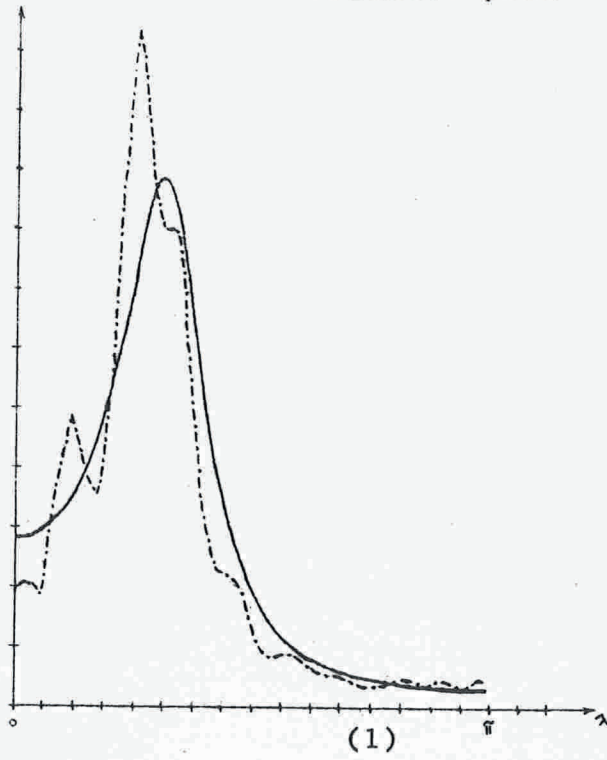
(3)

Espectros teórico e estimado para $t=380$





Espectros teórico e estimado para $t=270$



B I B L I O G R A F I A

- [1] BARTLETT, M.S. - "An Introduction to Stochastic Processes".
Cambridge, 1955.
- [2] BRILLINGER, D.R. e HATANAKA, M. - "An Harmonic Analysis of
Non-stationary Multivariate Economic Processes",
Econometrica, jan. 1969, 131-141.
- [3] BOX, G.E.P. e JENKINS, G. - "Time Series Analysis: Forecast-
ing and Control".
Holden Day, 1967.
- [4] CHOW, G. e LEVITAN, R. - "Spectral Properties of Non-station-
ary Systems of Linear Stochastic Difference Equations".
JASA, vol. 64, 1969, 581-590.
- [5] CRAMÉR, H. - "On Some Classes of Non-stationary Processes".
Proc. Fourth Berkeley Symposium Math. Statistic and Pro-
bability, 1960, 2, 57-78.
- [6] _____ - "On the Structure of Purely Non-determi-
nistic Stochastic Processes".
Arkiv Math, 1959, 4, 45-54.
- [7] _____ - "Remarques sur le Problème de Prediction pour
Certaines Classes de Processus Stochastiques".
Coloque sur le Calcul des Probabilités, 1959, 103-112.
- [8] CRAMÉR, H. e LEADBETTER, M.R. - "Stationary and Related
Stochastic Processes".
New York, Wiley, 1967.
- [9] DOOB, J.L. - "Stochastic Processes".
New York, Wiley, 1953.
- [10] GAIT, N. - "Ajustamento Sazonal de Séries Temporais".
Instituto de Matemática e Estatística, USP. São Paulo- 1975.

- [11] GRANGER, C.W.J. e HATANAKA, M. - "Spectral Analysis of Economic Time Series".
Princeton University Press - Princeton, New Jersey, 1964.
- [12] GRENANDER, U. e ROSENBLATT, M. - "Statistical Analysis of Stationary Time Series".
New York, Wiley, 1957.
- [13] HALMOS, P.R. - "Introduction to Hilbert Space and the Theory of Spectral Multiplicity".
New York, Chelsea, 1951.
- [14] LOÉVE, M. - "Probability Theory".
New York, Van Nostrand, 1955.
- [15] LOYNES, R.M. - "On the Concept of the Spectrum for Non-stationary Processes".
J. Roy. Stat. Soc. - Series B (Methodological) - vol. 30, 1, 1968, 1-30.
- [16] PAGE, C.H. - "Instantaneous Power Spectra".
J. Appl. Phys., 1952, 23, 103-106.
- [17] PARZEN, E. - "Mathematical Considerations in the Estimation of Spectra".
Technometrics, 1961, 3, 167-190.
- [18] _____ - "On Asymptotically Efficient Consistent Estimates of the Spectral Density Function of a Stationary Time Series".
J. Roy. Stat. Soc. - Serv. B, 1958, 20, 303-322.
- [19] PRIESTLEY, M.B. e TONG, H. - "On the Analysis of Bivariate Non-stationary Processes".
J. Roy. Stat. Soc., 1973, vol. 35, 2.
- [20] PRIESTLEY, M.B. - "Evolutionary Spectra and Non-stationary Processes".
J. Roy. Stat. Soc., B, 1965, 27, 204-237.

- [21] PRIESTLEY, M.B. - "Design Relations for Non-stationary Processes".
J.Roy. Stat. Soc., 1965, 28, 228-240.
- [22] _____ - "Basic Considerations in the Estimation of Spectra".
Technometrics, 1962, 4, 551-563.
- [23] RAO, M.M. - "Consistency and Limit Distributions of Estimators of Parameters in Explosive Stochastic Difference Equations".
Ann. Math. Stat., 1961, 32, 195-218.
- [24] STONE, M.H. - "On One Parameter Unitary Groups in Hilbert Space".
Ann. Math., 1932, 33, 643-648.