APLICAÇÃO DA ANÁLISE ESPECTRAL AO ESTUDO DO

MODELO DE DEFASAGENS DISTRIBUÍDAS

MARLI MIKAEL DA COSTA NEVES

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, MAIO DE 1977 -

AGRADECIMENTOS

Ao Professor Dr. Pedro Alberto Morettin, que ori
entou esta dissertação de forma paciente e amiga.

Aos colegas e amigos do Departamento de Estatís-
tica do Instituto de Matemática e Estatística da
Universidade de São Paulo, pelo incentivo constante
que deram à realização deste trabalho.

Ao sr. João Baptista Esteves de Oliveira, pelo e
ficiente trabalho de datilografia.

Í N D I C E

CAP.1 - INTRODUÇÃO	1
1.1 - Defasagens Distribuídas e suas Causas.	1
1.2 - Hipóteses Econômicas Geradoras do Modelo com Estrutura de Defasagens Geométrica, Exemplos de Aplicações	9
1.3 - A Estimação dos Coeficientes do Modelo	16
1.4 - Conceitos Iniciais e Objetivos	17
CAP.2 - TEORIA ESPECTRAL PARA SÉRIES TEMPORAIS ESTACIONÁRIAS . . .	23
2.1 - Introdução	23
2.2 - Distribuição Normal Complexa e Distribuição de Wishart.	23
2.3 - Representação Espectral dos Processos Estacionários de 2ª ordem.	26
2.4 - Análise Espectral Bivariada.	34
2.5 - Filtros.	37
2.6 - Função de Covariância Amostral	42
2.7 - Espectro Amostral - Propriedades - Distribuição. . .	45
2.8 - Periodograma	51
2.9 - Estimador do Espectro Cruzado, Coerência, Ganho e Ângulo de Fase	52
2.10 - Propriedades Assintóticas dos Estimadores - Distribuições Amostrais.	53
2.11 - Procedimento para Testar a Hipótese de Coerência nula	58
CAP.3 - MODELO DE DEFASAGENS DISTRIBUÍDAS: ESTIMAÇÃO	60
3.1 - Objetivos.	60
3.2 - Considerações Gerais	61
3.3 - O Enfoque Espectral para a Estimação do Vetor de Coeficientes	64
3.4 - A Estimação da Estrutura de Defasagens Geométrica. .	75
3.5 - A Escolha de m	77
3.6 - O Ajustamento para Médias Diferentes de Zero	78

CAP. 4 - MODELO DE DEFASAGENS DISTRIBUÍDAS: PROPRIEDADES DOS ESTIMADORES	79
4.1 - Propriedades Assintóticas do Estimador $\hat{a}'$	79
4.2 - Propriedades dos Estimadores $\hat{\alpha}$ e $\hat{\beta}$ da Estrutura de Defasagens Geométrica	88
4.3 - Quando o Processo Residual é o Ruído Branco	91
4.4 - A Validade do Modelo.	93
4.5 - A Extensão a Modelos Multivariados.	94
CAP. 5 - CONSIDERAÇÕES FINAIS.	97
5.1 - Modelos de Defasagens Distribuídas mais Gerais.	97
5.2 - Comparação entre Estimadores Obtidos no Domínio do Tempo e Estimadores Obtidos no Domínio da Frequência	100
5.3 - A Estimação do Vetor a' fazendo uso do Periodograma Alisado.	101
5.4 - O Problema da Computação.	104
BIBLIOGRAFIA	106

CAPÍTULO 1

INTRODUÇÃO

1.1 - DEFASAGENS DISTRIBUÍDAS E SUAS CAUSAS

Os modelos que explicam fenômenos que ocorrem na natureza são geralmente expressos por sistemas matemáticos. As variáveis que compõem estes sistemas podem classificar-se em dois grupos distintos: aquelas que são explicadas pelo modelo e que denominamos "endógenas", e aquelas que não são explicadas pelo modelo e que denominamos "exógenas". As exógenas ocorrem nas equações estabelecidas mas são determinadas independentemente.

As relações comportamentais envolvidas nos modelos descrevem como uma ou mais variáveis endógenas são explicadas pelos valores tomados por certas variáveis exógenas.

Muitas vezes, ocorre que os agentes descritos pelas variáveis endógenas, não reagem de modo completo e imediato à mudanças nas quantidades descritas pelas variáveis exógenas, mas sim de modo gradual, ou seja, a reação é es-

tendida sobre um período de tempo. Os *modelos de defasagens distribuídas* (distributed lag models) estão entre os que podem ser utilizados para explicar tal tipo de comportamento.

Para fixarmos as idéias, consideremos o caso em que uma só variável endógena y reage quando há mudanças numa só variável exógena x . Suponhamos também que após uma mudança em x esta permaneça constante neste novo nível. A alteração em x provoca um desequilíbrio no sistema, e como consequência poderemos observar um deslocamento progressivo em busca de um novo equilíbrio. Isto quer dizer que a reação de y não é imediata e pode ser descrita como um processo no tempo; parte desta reação sendo obtida após a decorrência de um período de tempo (mês, ano, etc.), outra parte durante um segundo período, e assim por diante até que, após um certo número de períodos, y esteja completamente ajustada ao novo valor de x . A figura (1.1) ilustra um possível processo de ajustamento, dado pela trajetória de ajustamento de y . A mudança em y por unidade de tempo, causada pela mudança em x , é dada por $\frac{dy(t)}{dt}$.

Um exemplo de um comportamento um pouco mais complexo é ilustrado na figura (1.2): seja y a demanda (por mês), e x o preço de algum bem não perecível. Se os compradores têm conhecimento antecipado da ocorrência de um aumento no preço ao tempo $t=0$, a demanda y pode crescer antes que os preços comecem a aumentar, declinando posteriormente talvez

FIGURA (1.1)

FIGURA (1.2)

mais abruptamente do que se não houvesse um acréscimo antecipado na demanda.

Existem três tipos de fatores que são responsáveis por adaptações não imediatas: tecnológicos, institucionais e psicológicos. Para citar um exemplo, suponhamos que uma indústria deseja aumentar sua produção para ajustá-la a um nível mais alto de vendas. Se a indústria não possui capacidade ociosa, um aumento da produtividade vai requerer uma modificação de ordem tecnológica (compra de equipamentos, aumento nas instalações, etc.) que só pode ser efetivada após um certo tempo.

De modo geral, a variável endógena y pode ser considerada como uma função linear dos valores prévios da variável exógena x , ou seja

$$y(t) = \sum_{s=0}^{\infty} a(s)x(t-s) + \epsilon(t) \quad (1.1)$$

onde $\epsilon(t)$ representa um erro aleatório.

Se supusermos que os coeficientes $a(s)$ são constantes desconhecidas não todas iguais a zero, e que $\epsilon(t)$ é uma variável aleatória independente de $x(t)$, tendo média zero e variância constante, o modelo dado em (1.1) é então denominado "Modelo Geral de Defasagens Distribuídas".

O fato de (1.1) conter um número infinito de parâmetros, faz com que a estimação dos mesmos seja impraticável sem que restrições sejam impostas ao modelo. Estas restrições devem se basear em alguma idéia que temos a respeito dos coeficientes e/ou em considerações de simplicidade, que aqui são entendidas como um instrumento para facilitar a estimação. Expressamos então nossa informação â priori numa forma analítica simples, isto é, escrevemos $a(s)$ como função de s e de um certo número de parâmetros.

É conveniente observar que a influência de uma alteração no valor da variável exógena no tempo t_0 , deve ser especialmente importante para valores da variável endógena em tempos imediatamente sucessivos. Sendo assim a sequência $a(0), a(1), \dots$, pode ser crescente em seus primeiros termos, mas deve decrescer continuamente a partir do máximo.

Irving Fisher, que foi um dos precursores do uso de modelos de defasagens distribuídas em Economia, propôs que os coeficientes em (1.1) fossem dados por:

$$a(s) = \begin{cases} \alpha(n+1-s), & s = 0, \dots, n \\ 0 & \text{caso contrário.} \end{cases} \quad (1.2)$$

A proposta de Fisher faz com que tenhamos que estimar apenas um parâmetro, mas apresenta o problema da estipu-

lação *à priori* de um valor para n . Se supusermos n conhecido a estimação de α torna-se muito fácil, pois (1.2) transforma (1.1) num modelo de regressão univariado (ver em [11] p.32).

Almon [1] adotou a suposição de que $a(s)$ seja escrito com uma função polinomial de s , sendo que o grau do polinômio deve ser escolhido *à priori*. Com isto Almon tornou mais flexível o esquema de Fisher.

Solow [36] sugeriu o uso da distribuição de Pascal, ou seja

$$a(s) = \binom{r+s-1}{s} (1-\lambda)^r \cdot \lambda^s, \quad (1.3)$$

a qual envolve apenas dois parâmetros r e λ . O esquema de Solow apresenta o inconveniente de conduzir a um problema de estimação com restrições não lineares (ver em [15]).

Dentre as várias hipóteses sobre a forma dos coeficientes de modelos de defasagens distribuídas, a que ganhou maior popularidade pelo êxito obtido em análises empíricas, é a idealizada por Koyck [27]. Partindo do modelo geral dado em (1.1) e estipulando um declínio geométrico dos coeficientes, ou seja,

$$a(s) = \alpha \lambda^s, \quad \text{onde } 0 \leq \lambda < 1, \quad (1.4)$$

ele chegou ao esquema de defasagens distribuídas dado por:

$$y(t) = \alpha \sum_{s=0}^{\infty} \lambda^s x(t-s) + \varepsilon(t), \quad 0 \leq \lambda < 1 \quad (1.5)$$

Da diferença entre as equações

$$y(t) = \alpha x(t) + \alpha \lambda x(t-1) + \alpha \lambda^2 x(t-2) + \dots + \varepsilon(t)$$

e

$$\lambda y(t-1) = \alpha \lambda x(t-1) + \alpha \lambda^2 x(t-2) + \dots + \lambda \varepsilon(t-1),$$

obteve

$$y(t) = \alpha x(t) + \lambda y(t-1) + \varepsilon(t) - \lambda \varepsilon(t-1) \quad (1.6)$$

A forma (1.6), em geral, facilita muito a estimação; o único inconveniente é que se em (1.5) os erros são variáveis a leatórias independentes, o mesmo não ocorre com relação a (1.6).

Um procedimento para estimar α e λ , e que foi usado por Koyck, é o método da regressão de mínimos quadrados de $y(t)$ em $x(t)$ e $y(t-1)$, onde $[\varepsilon(t) - \lambda \varepsilon(t-1)]$ é um erro aleatório composto. Para amostras pequenas sabe-se que o estimador do coeficiente de $y(t-1)$ é viciado. Koyck mostrou que até para amostras grandes o tratamento de mínimos quadrados leva a um estimador cujo vício não é desprezível, pois

$\epsilon(t-1)$ que é parte do erro aleatório composto, não é independente de $y(t-1)$. Assumindo que $\epsilon(t)$ seja um processo autoregressivo de 1ª ordem dado por:

$$\epsilon(t) = \xi \epsilon(t-1) + u(t), \quad (1.7)$$

ele determinou estimadores de α e λ assintoticamente não viesados, e consistentes no caso em que $\lambda = \xi$.

Baseando-se na consideração de que pode ser irreal assumir que os coeficientes sucessivos declinem imediatamente, particularmente quando o período de tempo t é pequeno, o modelo de Koyck pode ser estendido. Adota-se então um declínio geométrico a partir do k -ésimo lag, ou seja,

$$y(t) = \sum_{s=0}^{k-1} a(s)x(t-s) + \lambda a(k)x(t-k) + \lambda^2 a(k)x(t-k-1) + \dots + \epsilon(t) \quad (1.8)$$

Subtraindo-se então

$$\lambda y(t-1) = \lambda \sum_{s=0}^{k-1} a(s)x(t-1-s) + \lambda^2 a(k)x(t-1-k) + \dots + \lambda \epsilon(t-1)$$

de (1.8), chega-se a forma reduzida

$$\begin{aligned} y(t) = & a(0)x(t) + [a(1) - \lambda a(0)]x(t-1) + \dots + [a(k-1) - \lambda a(k-2)]x(t-k+1) + \\ & + \lambda y(t-1) + \epsilon(t) - \lambda \epsilon(t-1) \end{aligned} \quad (1.9)$$

Koyck utilizou este modelo para estudar a resposta

da capacidade devida a mudanças na produção em algumas indústrias americanas.

1.2 - HIPÓTESES ECONÔMICAS GERADORAS DO MODELO COM ESTRUTURA DE DEFASAGENS GEOMÉTRICAS. EXEMPLOS DE APLICAÇÕES

Cagan [7] e Nerlove [31, 32] utilizaram modelos econométricos que justificam o declínio geométrico para a forma de defasagens, isto é, mostraram que esta forma pode ser gerada quando estamos sob a suposição de algumas hipóteses econômicas básicas.

Cagan sugeriu um modelo conhecido como modelo de "expectativas adaptadas". A variável y é suposta ser função dos valores esperados de uma variável x , ou seja

$$y(t) = \alpha x^*(t) + u(t), \quad (1.10)$$

onde

$$x^*(t) - x^*(t-1) = \gamma[x(t-1) - x^*(t-1)], \quad 0 < \gamma < 1 \quad (1.11)$$

A expressão (1.11) é uma equação em diferenças finitas, e pode ser facilmente resolvida com relação a x_t^* , obtendo-se

$$x^*(t) = \sum_{j=0}^{\infty} \gamma(1-\gamma)^j x(t-j-1), \quad (1.12)$$

o que mostra que a hipótese de Cagan resulta numa forma de defasagens distribuídas onde os coeficientes declinam exponencialmente.

O procedimento de Cagan para estimar γ é bastante laborioso, pois consiste em experimentar diferentes γ 's construir a sequência de x^* associada a eles e escolher o γ que maximiza o coeficiente de correlação múltipla na equação (1.10).

Cagan formulou originalmente a hipótese de expectativas adaptadas, para introduzir as expectativas da taxa de inflação como uma variável explicativa da demanda de moeda nos países em processo de hiperinflação.

Para ilustrar a hipótese de expectativas adaptadas, vamos dar um exemplo bem geral. Admitamos que a quantidade de demanda $q(t)$ de um produto particular durante um período de tempo t , depende não somente do preço $p(t)$ deste produto, como também do preço $p^N(t)$ que os compradores consideram como normal para o mesmo. A demanda poderá ser escrita como:

$$q(t) = ap(t) + bp^N(t) + c + u(t),$$

onde c é uma constante. Admitamos também que o preço normal seja revisto em cada período em proporção ao erro associado

no período anterior, ou seja,

$$p^N(t) - p^N(t-1) = \gamma[p(t-1) - p^N(t-1)],$$

sendo γ um coeficiente numérico compreendido entre zero e um, característico da mesma maneira pela qual se forma o conceito de "preço normal". A equação acima conduz a

$$p^N(t) = \gamma \sum_{j=0}^{\infty} (1-\gamma)^j p(t-j-1),$$

o que mostra que o preço normal é uma média ponderada dos preços observados no passado, onde os coeficientes de ponderação formam uma progressão geométrica decrescente. Deste modo, a demanda pode ser expressa em função da série de preços observados segundo a relação:

$$q(t) = ap(t) + b\gamma \sum_{j=0}^{\infty} (1-\gamma)^j p(t-j-1) + c + u(t).$$

Nerlove combinando o modelo de expectativas adaptadas com o procedimento de redução de Koyck, formulou o modelo de "ajustamento parcial". Com isto, além de facilitar a estimação, conseguiu um esquema aplicável a um número bem grande de problemas. No modelo de ajustamento parcial, os valores da variável independente determinam o valor desejado da variável dependente, ou seja

$$y^*(t) = \alpha x(t) + u(t), \quad (1.13)$$

e apenas uma fração do ajustamento desejado é conseguida num período de tempo, isto é:

$$y(t) - y(t-1) = \gamma[y^*(t) - y(t-1)], \quad 0 < \gamma < 1 \quad (1.14)$$

De (1.13) e (1.14) chega-se a:

$$y(t) = \alpha \gamma x(t) + (1-\gamma)y(t-1) + \gamma u(t), \quad (1.15)$$

que é semelhante a forma reduzida (1.6), com a vantagem de não induzir correlação serial nos erros se estes originalmente são não correlacionados.

Nerlove utilizou o modelo de ajustamento parcial para estudar a resposta da oferta agrícola aos preços.

O exemplo abaixo, é dado com o objetivo de ilustrar o modelo de ajustamento parcial. Suponhamos que a demanda de um produto se estabelece em função de um nível de equilíbrio a longo prazo, que os compradores consideram adaptado ao preço que observam. Seja $q^E(t)$ esta demanda de equilíbrio, para a qual admitimos a relação

$$q^E(t) = ap(t) + b + u(t),$$

onde b é uma constante. A demanda real difere da de equilíbrio, pelo fato de que as instalações domésticas ou industriais em que o produto deve ser utilizado não estão perfeitamente adaptadas. Nestas condições, parece lógico admitir

que a variação de demanda $[q(t)-q(t-1)]$ depende do desvio entre esta demanda e a demanda de equilíbrio $[q(t-1)-q^E(t)]$, e estabelecer que:

$$q(t) - q(t-1) = \gamma[q^E(t) - q(t-1)] + v(t), \quad 0 < \gamma < 1.$$

Das duas igualdades precedentes chega-se a:

$$q(t) = (1-\gamma)q(t-1) + \gamma p(t) + b\gamma + \mu(t),$$

onde $\mu(t) = \gamma u(t) + v(t)$.

Vejamos agora alguns exemplos da utilização dos modelos aqui citados, na economia brasileira.

Para estudar a resposta da produção agrícola aos preços no Brasil, Pastore [34] utilizou os dados de produção, área cultivada e preços pagos aos produtores, existentes para os principais produtos agrícolas do país. Foram estimadas as funções de oferta para dez produtos individuais, que atingiam, no período 1962/64, mais de 70% do valor da produção agrícola global do país, a saber: algodão, amendoim, arroz, cana, cebola, feijão, fumo, mamona, mandioca e milho. Dentre os produtos mais importantes apenas o café foi abandonado, devido à sua característica de cultura permanente, e com um mecanismo de resposta aos preços mais complexo que dos demais produtos.

Observando que, devido a natureza da produção agrícola, é muito provável que a produção não responda de imediato aos estímulos dos preços relativos, Pastore achou viável a adoção da hipótese de ajustamento parcial para explicar o comportamento da oferta agrícola.

Denotando por $y^*(t)$ a oferta desejada de longo prazo no tempo t , e supondo que as expectativas são estáticas, isto é, que os preços esperados em t são iguais aos preços efetivamente verificados em $t-1$, o modelo pode ser expresso pelo sistema

$$y^*(t) = f(p(t-1), p^i(t-1), X) + u(t), \quad (1)$$

para $i=2,3,\dots,n$

$$[y(t) - y(t-1)] = \gamma[y^*(t) - y(t-1)], \quad (2)$$

onde $p(t-1)$ é o preço observado em $(t-1)$, $p^i(t-1)$ representa o preço do i -ésimo produto alternativo ou fator de produção, e X é um vetor de variáveis exógenas que afetam a produção independentemente dos preços. A forma mais frequentemente utilizada em análises empíricas é a que admite uma relação linear em (1) (com o objetivo de se conseguir estacionaridade através da eliminação de tendência, muitas vezes usam-se também relações lineares nos logaritmos para (1) e (2)).

As ofertas foram especificadas exprimindo a área cultivada em t (utilizada como uma aproximação da produção em t) em função dos preços pagos aos produtores em $(t-1)$, da área cultivada em $(t-1)$, e de uma tendência introduzida para captar os efeitos sistemáticos das variáveis que provocam o crescimento contínuo da produção, como o progresso tecnológico, o aumento do número de firmas no mercado, etc. Pastore justifica o porquê de algumas variáveis tais como, o preço dos fatores de produção e o preço de produtos substitutos mais próximos, não terem sido incluídos no modelo, e ainda a razão de ter sido utilizada a área cultivada para substituir a produção.

Deste modo, o modelo utilizado para a estimação da função de oferta de cada produto, pode ser obtido do sistema:

$$y^*(t) = a + bp(t-1) + ct + u(t) \quad (1)$$

$$y(t) - y(t-1) = \gamma[y^*(t) - y(t-1)] \quad (2)$$

Das equações (1) e (2) obtemos a forma reduzida do modelo, dada por:

$$y(t) = (1-\gamma)y(t-1) + a\gamma + b\gamma p(t-1) + c\gamma t + \gamma u(t).$$

O outro exemplo que iremos agora citar, foi extraído de um estudo sobre a política monetária brasileira [35] de

vido também a Pastore. Procurando estimar a função demanda de moeda, ele fez uso da hipótese de que a taxa esperada de inflação, que é uma variável explicativa da demanda de moeda, seja dada por um processo de expectativas adaptadas da forma

$$(1-bL)E(t) = (1-b)\pi(t-1),$$

onde $\pi(t)$ é a taxa de inflação no instante t , $E(t)$ é a taxa esperada de inflação e L é o operador defasagens (lag operator), e é tal que $Lx_t = x_{t-1}$. Esta última equação conduz a uma expressão em que a taxa esperada de inflação aparece como uma média móvel dos valores passados de π , com pesos geometricamente decrescentes.

1.3 - A ESTIMAÇÃO DOS COEFICIENTES DO MODELO

A estimação dos parâmetros da forma de defasagens geométrica, pelo sucesso que alcançou esta estrutura de defasagens, é a que tem merecido maior atenção por parte dos estudiosos do modelo.

Com a suposição de que em (1.6) os erros são não correlacionados serialmente, Koyck e Klein [37] determinaram estimadores assintoticamente não viciados e consistentes

tes para os parâmetros, utilizando uma combinação do procedimento de Aitken com o método de mínimos quadrados; sob a suposição de que os erros são dados por (1.7), os estimadores de Koyck e Klein são inconsistentes (exceto se $\xi = \lambda$). Fazendo uso de variáveis instrumentais, Leviatan [28] determinou estimadores consistentes no caso em que os erros obedecem a (1.7). Neste caso também, procedimentos de estimação em dois estágios foram propostos por Gupta [19] e Dhrymes [10].

Os métodos de estimação não estão restritos à forma geométrica, nem tão pouco se resumem nos aqui citados. Dhrymes [11] apresenta detalhadamente os principais, no contexto dos modelos com defasagens distribuídas.

Malinvaud [29] enfoca os principais problemas que prejudicam a estimação, que são: autocorrelação serial nos erros, escolha inapropriada do número de defasagens (lags) e multicolinearidade entre as variáveis defasadas.

1.4 - CONCEITOS INICIAIS E OBJETIVOS

Os conceitos descritos abaixo são importantes para que possamos formular os objetivos deste estudo.

Seja $\{X(t, \omega), t \in T, \omega \in \Omega\}$ um processo estocástico de

finido num espaço de probabilidade (Ω, A, P) . Para t fixo e ω variável, $X(t, \omega)$ é uma variável aleatória. Se t representa o tempo, então para ω fixo e t variável $X(t, \omega)$ é uma função do tempo.

Assim, um processo estocástico pode ser visto como uma coleção de variáveis aleatórias com t percorrendo T . Esta coleção é também denominada série temporal, e $X(t, \omega)$ para um ω fixado é uma realização da série temporal.

Denotaremos um processo estocástico por $\{X(t), t \in T\}$ omitindo a dependência em ω .

Em geral, T é um subconjunto de $Z = \{0, \pm 1, \pm 2, \dots\}$, e neste caso dizemos que o processo estocástico é de parâmetro discreto, ou um intervalo real e dizemos que o processo é de parâmetro contínuo. Porém T pode ser bem mais geral, e em muitas situações é um espaço vetorial n -dimensional.

Ao processo estocástico $\{X(t); t=0, \pm 1, \pm 2, \dots\}$ denominaremos sequência estocástica.

Um processo estocástico é multivariado de ordem $(k \times 1)$, e denotaremos por $\{X(t), t \in T\}$, se

$$X(t) = \begin{bmatrix} X_1(t) \\ X_2(t) \\ \vdots \\ X_k(t) \end{bmatrix}.$$

O processo estocástico $\{X(t), t \in T\}$ fica completamente descrito se

- a) para todo t pertencente a T , conhecermos a função distribuição $F(x, t) = P[X(t) \leq x]$ da variável $X(t)$;
- b) para todo par de elementos t_1, t_2 em T , conhecermos a função distribuição $F(x_1, x_2; t_1, t_2) = P[X(t_1) \leq x_1, X(t_2) \leq x_2]$, da variável bi-dimensional $[X(t_1), X(t_2)]$ e, em geral,
- c) para quaisquer $t_1, t_2, \dots, t_n$ pertencentes a T , conhecermos a função distribuição

$$F(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = P[X(t_1) \leq x_1, X(t_2) \leq x_2, \dots, X(t_n) \leq x_n]$$

da variável aleatória n -dimensional $[X(t_1), X(t_2), \dots, X(t_n)]$.

As funções distribuição $F(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)$ devem satisfazer às seguintes condições:

- i) $F(x_{j_1}, \dots, x_{j_n}; t_{j_1}, \dots, t_{j_n}) = F(x_1, \dots, x_n; t_1, \dots, t_n)$, onde $j_1, \dots, j_n$ é qualquer permutação de $1, 2, \dots, n$ (condição de simetria).
- ii) $F(x_1, \dots, x_m, \infty, \dots, \infty; t_1, \dots, t_m, t_{m+1}, \dots, t_n) = F(x_1, \dots, x_m; t_1, \dots, t_m)$, quaisquer que sejam $t_{m+1}, \dots, t_n$, se $m < n$ (condição de compatibilidade).

Reciprocamente, qualquer família de distribuições de dimensão finita satisfazendo a (i) e (ii) pode ser vista como definindo algum processo estocástico sobre T .

DEFINIÇÃO 1.1 - Um processo estocástico $\{X(t), t \in T\}$ é estritamente estacionário ou fortemente estacionário, se $[X(t_1), \dots, X(t_n)]$ e $[X(t_{1+s}), \dots, X(t_{n+s})]$ têm a mesma função distribuição n-dimensional de probabilidade, para todo $s \in T$ e para todo inteiro positivo n .

As leis de probabilidade que governam o processo estocástico são, portanto, invariantes por translação de tempo.

DEFINIÇÃO 1.2 - Seja $\{X(t), t \in T\}$ um processo estocástico de ordem $(k \times 1)$, tal que as médias $E[X_i(t)] = \mu_i(t)$ e as funções de covariância $E\{[X_i(s) - \mu_i(s)][X_j(t) - \mu_j(t)]\} = R_{ij}(s, t)$ existam e sejam finitas $\forall i, j = 1, \dots, k$. Este processo é denominado fracamente estacionário ou estacionário de 2ª ordem se:

- i) $\mu_j(t) = \mu_j, \forall j = 1, \dots, k$ (a média é uma constante independente de t);
- ii) $R_{ij}(s, t) = R_{ij}(s - t), \forall i, j = 1, \dots, k$ (R é uma função apenas da diferença $(s - t)$).

OBSERVAÇÃO - Um processo estritamente estacionário com momento de 2ª ordem finito, é também fracamente estacionário.

DEFINIÇÃO 1.3 - Um processo estocástico $\{X(t), t \in T\}$ real é denominado normal, se para qualquer conjunto finito

$$t_1, t_2, \dots, t_n \in T,$$

a distribuição conjunta de $X(t_1), \dots, X(t_n)$ é normal n-variá

da.

Neste trabalho nos propomos a estudar o modelo de defasagens distribuídas:

$$Y(t) = \sum_{s=-r_1}^{r_2} a(s)X(t-s) + \eta(t) \quad (1.16)$$

onde $X(t)$, $Y(t)$ e $\eta(t)$ são sequências estocásticas estacionárias (fracamente) de média zero, com a suposição de que $X(t)$ e $\eta(t)$ são independentes.

Dados T pares de observações de X e Y , com a suposição de que $\eta(t)$ é normal, veremos como o estimador de máxima verossimilhança do vetor

$$a' = [a(-r_1), a(-r_1+1), \dots, a(0), \dots, a(r_2)],$$

de coeficientes do modelo pode ser obtido, usando métodos da análise espectral, e estudaremos propriedades assintóticas deste estimador. Com a suposição adicional de que $X(t)$ é também normal, veremos qual a distribuição assintótica deste estimador.

Hannan [22,23,24] foi quem primeiro estudou processos para estimar os parâmetros de modelos de defasagens distribuídas, fazendo uso da análise espectral.

Além do modelo geral dado em (1.16), procuraremos

também estudar o caso particular em que a estrutura de defa-
sagens é geométrica.

Este trabalho, por apresentar uma técnica de esti-
mação ainda não suficientemente divulgada entre nossos pes-
quisadores, visará apenas um desenvolvimento teórico detalha-
do da mesma, sem usá-la em aplicações.

CAPÍTULO 2

TEORIA ESPECTRAL PARA SÉRIES

TEMPORAIS ESTACIONÁRIAS

2.1 - INTRODUÇÃO

Neste capítulo apresentamos as noções básicas necessárias à compreensão do estudo que temos em vista. Deste modo, na secção 2.2 apresentamos algumas distribuições contínuas multivariadas que mostram-se relevantes no estudo da distribuição amostral do estimador do vetor de coeficientes do modelo dado em (1.16), e nas demais secções, noções de teoria espectral fundamentais à estimação e ao estabelecimento das propriedades deste estimador.

2.2 - DISTRIBUIÇÃO NORMAL COMPLEXA E DISTRIBUIÇÃO DE WISHART

Se $X = [x_{jk}]$ é uma matriz complexa ($J \times k$), onde x_{jk} é o número complexo da j -ésima linha e k -ésima coluna, então representaremos:

- a) por $\bar{X} = [\bar{x}_{jk}]$ a matriz complexa conjugada de X .
- b) por $X' = [x_{kj}]$ a transposta de X .
- c) por $\text{Det}(X)$ o determinante de X , se X é uma matriz quadrada.
- d) por $\text{tr}(X)$ o traço de X , se X é uma matriz quadrada.
- e) se $x_{jk} = \text{Re}x_{jk} + i \text{Im}x_{jk}$, então:

$$\text{Re}X = [\text{Re}x_{jk}] \quad \text{e} \quad \text{Im}X = [\text{Im}x_{jk}]$$

DEFINIÇÃO 2.1 - Dizemos que X é hermetiana se $\bar{X}' = X$.

DEFINIÇÃO 2.2 - Se X é hermitiana, dizemos que X é positiva de
finita se

$$\sum_{j,k=1}^J \alpha_j \bar{\alpha}_k x_{jk} > 0,$$

quaisquer que sejam os números complexos α_j , $j=1, \dots, J$.

DEFINIÇÃO 2.3 - Se X é matriz quadrada, dizemos que é não sín-
gular se $\text{Det}(X) \neq 0$.

DEFINIÇÃO 2.4 - Seja X uma variável aleatória r -dimensional com componentes reais. Dizemos que X tem *distribuição normal multivariada* com média μ e matriz de covariâncias Σ , se a sua função densidade é dada por:

$$\frac{1}{(2\pi)^{r/2}} \cdot \frac{1}{[\det(\Sigma)]^{1/2}} \cdot e^{-1/2(X-\mu)' \Sigma^{-1}(X-\mu)}$$

onde Σ é hermitiana, positiva definida.

Notação: X é $N_r(\mu; \Sigma)$.

DEFINIÇÃO 2.5 - Seja X tal que

$$X = \text{Re}X + i\text{Im}X$$

Se o vetor de $2r$ componentes $\begin{bmatrix} \text{Re}X \\ \text{Im}X \end{bmatrix}$ tem distribuição normal multivariada

$$N_{2r} \left(\begin{bmatrix} \text{Re}\mu_X \\ \text{Im}\mu_X \end{bmatrix}; \frac{1}{2} \begin{bmatrix} \text{Re} \Sigma_{XX} & -\text{Im}\Sigma_{XX} \\ \text{Im} \Sigma_{XX} & \text{Re}\Sigma_{XX} \end{bmatrix} \right)$$

para algum vetor $(r \times 1)$ μ_X , e matriz Σ_{XX} hermitiana positiva definida, então X tem *distribuição normal multivariada complexa* com média μ_X e matriz de covariâncias Σ_{XX} .

Notação: X é $N_r^C(\mu_X, \Sigma_{XX})$.

DEFINIÇÃO 2.6 - Sejam $X_1, \dots, X_n$ variáveis aleatórias independentes cada uma $N_r(0, \Sigma_{XX})$; então a variável aleatória matricial (matriz $r \times r$)

$$W = \sum_{j=1}^n X_j X_j^*$$

tem distribuição de Wishart de dimensão r e com n graus de liberdade.

Notação: W é $W_r(n, \Sigma_{XX})$.

DEFINIÇÃO 2.7 - Se $X_1, \dots, X_n$ são variáveis aleatórias inde-

pendentes, cada uma $N_r^C(0, \Sigma_{XX})$, então a variável aleatória ma
tricial (matriz $r \times r$)

$$W = \sum_{j=1}^n x_j \bar{x}_j'$$

tem distribuição de Wishart complexa de dimensão r e com n graus de li
berdade.

Notação: W é $W_r^C(n, \Sigma_{XX})$.

A distribuição de Wishart complexa foi introduzida por Goodman [18]. Sua função densidade é dada por

$$[\pi^{r(r-1)/2} \prod_{j=1}^n \Gamma(n-j+1)]^{-1} (\text{Det} \Sigma_{XX})^{-n} (\text{Det} W)^{n-r} \exp[-\text{tr} \Sigma_{XX}^{-1} W]$$

para $n \geq r$ e W positiva definida. As propriedades da distribuição de Wishart podem ser vistas em Anderson [3], e a extensão ao caso complexo em Goodman [16,17].

2.3 - REPRESENTAÇÃO ESPECTRAL DOS PROCESSOS ESTACIONÁRIOS DE 2ª ORDEM

A idéia de que uma função analítica pode ser repre
sentada, sobre um intervalo fixado, por uma série de senos e cosenos, foi desenvolvida nos séculos XVIII e XIX por Lagrange, Euler e Fourier.

A análise destas funções é conhecida com o nome de Análise Harmônica ou de Fourier, e é de grande importância

no estudo das séries temporais.

Para compreendermos melhor esta importância, consideremos o processo:

$$X(t) = \sum_{j=1}^n (A_j \cos \lambda_j t + B_j \sin \lambda_j t), \quad (2.1)$$

onde $E(A_j) = E(B_j) = 0$ e $E(A_j^2) = E(B_j^2) = \sigma_j^2$, $\forall j = 1, \dots, n$.

Vamos calcular a média e a função de covariância do processo:

$$E[X(t)] = E\left[\sum_{j=1}^n A_j \cos \lambda_j t\right] + E\left[\sum_{j=1}^n B_j \sin \lambda_j t\right] = 0$$

Seja $C_j = \frac{A_j + iB_j}{2}$ e $\bar{C}_j$ o complexo conjugado de C_j

$$\begin{aligned} R(s, t) = E[X(s)X(t)] &= \sum_{j=1}^n \sum_{k=1}^n E[\bar{C}_j \bar{C}_k e^{i(\lambda_j t + \lambda_k s)} + C_j \bar{C}_k e^{-i(\lambda_j t - \lambda_k s)} \\ &\quad + \bar{C}_j C_k e^{i(\lambda_j t - \lambda_k s)} + C_j C_k e^{-i(\lambda_j t + \lambda_k s)}]. \end{aligned}$$

Para que o processo seja estacionário de 2ª ordem devemos ter $R(s, t) = R(s-t)$, do que segue

$$E[C_j C_k] = E[\bar{C}_j \bar{C}_k] = E[C_j \bar{C}_k] = 0, \quad \forall j \neq k,$$

ou seja,

$$\begin{cases} E(A_j A_k) = E(B_j B_k) = 0, & \forall j \neq k, \\ E(A_j B_k) = 0. \end{cases} \quad (2.2)$$

De (2.2) vemos que os elementos do vetor

$$A = (A_1, \dots, A_n)$$

são ortogonais aos pares, o mesmo ocorrendo com os elementos de $B = (B_1, \dots, B_n)$, e que todo elemento de A é ortogonal a todo elemento de B .

Com as suposições (2.2) a função de covariância fica:

$$R(s-t) = R(u) = \sum_{j=1}^n \sigma_j^2 \cos \lambda_j u. \quad (2.3)$$

A variância é dada por:

$$R(0) = \sum_{j=1}^n \sigma_j^2. \quad (2.4)$$

Assim, a decomposição de um processo estacionário de 2ª ordem numa superposição de ondas de seno e cosseno como em (2.1), nos garante as seguintes vantagens:

- I) A variância pode ser escrita como soma das variâncias de cada componente periódica. Este fato e falta de termos de covariância, nos permite determinar a importância relativa de cada componente na variação média

dia quadrática.

II) Se as variáveis A e B tivessem distribuição normal n-variada, então poderíamos considerar a variação média quadrática como sendo gerada por fontes independentes.

TEOREMA DA REPRESENTAÇÃO ESPECTRAL (Ver [37], p.36) - Todo processo real, estacionário de 2ª ordem, $\{X(t), t \in T\}$, pode ser representado na forma da seguinte integral de Fourier-Stieltjes:

$$X(t) = \int_0^\infty [\cos \lambda t dU(\lambda) + \sin \lambda t dV(\lambda)], \quad (2.5)$$

onde $U(\lambda)$ e $V(\lambda)$ são processos estocásticos de parâmetro contínuo λ , e têm as seguintes propriedades:

a) são processos com incrementos ortogonais, ou seja, para quaisquer $\lambda_1 < \lambda_2 \leq \lambda_3 < \lambda_4$

$$E\{[U(\lambda_2) - U(\lambda_1)][U(\lambda_4) - U(\lambda_3)]\} = 0 \quad (2.6)$$

$$E\{[V(\lambda_2) - V(\lambda_1)][V(\lambda_4) - V(\lambda_3)]\} = 0$$

b) para quaisquer $\lambda_1, \lambda_2, \lambda_3, \lambda_4$

$$E\{[U(\lambda_2) - U(\lambda_1)][V(\lambda_4) - V(\lambda_3)]\} = 0, \quad (2.7)$$

e simbolicamente escrevemos $E[dU(\lambda)dV(\lambda')] = 0, \forall \lambda, \lambda'.$

A propriedade a) nos garante a existência de uma função G monótona, não decrescente, determinada de modo único a menos de uma constante aditiva por:

$$E\{[U(\lambda_2) - U(\lambda_1)]^2\} = E\{[V(\lambda_2) - V(\lambda_1)]^2\} = G(\lambda_2) - G(\lambda_1), \quad (2.8)$$

e simbolicamente escrevemos: $E\{[dU(\lambda)]^2\} = E\{[dV(\lambda)]^2\} = dG(\lambda)$.

A representação (2.5) pode ser escrita como:

$$X(t) = \int_{-\infty}^{\infty} \left[\frac{e^{i\lambda t} + e^{-i\lambda t}}{2} dU(\lambda) + \frac{e^{i\lambda t} - e^{-i\lambda t}}{2i} dV(\lambda) \right] = \int_{-\infty}^{\infty} e^{i\lambda t} dZ(\lambda).$$

Assim

$$X(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dZ(\lambda), \quad (2.9)$$

onde

$$\begin{cases} dU(\lambda) = dU(-\lambda) \\ -dV(\lambda) = dV(-\lambda) \\ dZ(\lambda) = \frac{dU(\lambda) - idV(\lambda)}{2} \end{cases} \quad (2.10)$$

A representação (2.5) ou (2.9) é denominada *representação espectral do processo*.

Se $X(t)$ é um processo estacionário de 2ª ordem de média zero, a sua função de covariância é dada por:

$$\begin{aligned} R(s-t) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(\lambda s - \lambda' t)} E[dZ(\lambda) \overline{dZ(\lambda')}] = \\ &= \int_{-\infty}^{\infty} e^{i\lambda(s-t)} E[|dZ(\lambda)|^2] = \int_{-\infty}^{\infty} e^{i\lambda(s-t)} \frac{dG(\lambda)}{2}. \end{aligned}$$

Seja

$$\frac{dG(\lambda)}{2} = dF(\lambda), \quad \forall \lambda \quad (2.11)$$

portanto:

$$R(s-t) = \int_{-\infty}^{\infty} e^{i\lambda(s-t)} dF(\lambda) \quad (2.12)$$

e a variância de $X(t)$ é dada por:

$$R(0) = \int_{-\infty}^{\infty} dF(\lambda) \quad (2.13)$$

DEFINIÇÃO 2.8 - A função $F(\lambda)$, definida em (2.11), é denominada *função de distribuição espectral do processo* $X(t)$. Se $F(\lambda)$ é absolutamente contínua, a sua derivada $f(\lambda)$ é denominada *função de densidade espectral* ou *espectro do processo*. Neste caso, podemos escrever:

$$R(s-t) = R(u) = \int_{-\infty}^{\infty} e^{i\lambda u} f(\lambda) d\lambda. \quad (2.14)$$

Portanto, $R(u)$ e $f(\lambda)$ são pares de uma transformada de Fourier.

Se $\int_{-\infty}^{\infty} |R(u)| du < \infty$, então:

$$f(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R(u) e^{-i\lambda u} du. \quad (2.15)$$

Se $X(t)$ é uma sequência estacionária de 2^{a} ordem,

sua representação espectral, função de covariância e função de densidade espectral são dadas por:

$$X(t) = \int_{-\pi}^{\pi} e^{i\lambda t} dZ(\lambda),$$

$$R(u) = \int_{-\pi}^{\pi} e^{i\lambda u} f(\lambda) d\lambda,$$

$$f(\lambda) = \frac{1}{2\pi} \sum_{u=-\infty}^{\infty} e^{-i\lambda u} R(u),$$

respectivamente.

PROPRIEDADES - Se $X(t)$ é uma sequência estacionária, então:

- a) $f(\lambda) \geq 0$
- b) $f(\lambda) = f(-\lambda)$
- c) $f(\lambda) = f(\lambda + 2\pi)$.

Como vimos, podemos representar um processo estacionário de 2ª ordem no domínio do tempo, por meio da função de covariância, ou no domínio da frequência, por meio da função de densidade espectral. Esta última, apresenta a vantagem de caracterizar o processo em termos de contribuições aditivas não correlacionadas, tornando mais clara a análise do mesmo.

EXEMPLO 2.1 - Seja $\{\varepsilon(t), t=0, \pm 1, \dots\}$ sequência de variáveis aleatórias independentes e identicamente distribuídas, tal

que

$$E[\epsilon(t)] = 0$$

e

$$R_{\epsilon\epsilon}(u) = \begin{cases} \sigma^2 & \text{para } u=0 \\ 0 & \text{para } u \neq 0 \end{cases}$$

Esta sequência é denominada "ruído branco".

De acordo com (2.15):

$$f_{\epsilon\epsilon}(\lambda) = \frac{\sigma^2}{2\pi}, \quad -\pi \leq \lambda \leq \pi$$

Como $f_{\epsilon\epsilon}(\lambda)$ é uniforme sobre o domínio $(-\pi, \pi)$, todas as frequências neste domínio contribuem igualmente para a variância.

EXEMPLO 2.2 - Seja $\{X(t), t=0, \pm 1, \dots\}$ dada por

$$X(t) = \sum_{s=0}^{\infty} a(s)\epsilon(t-s)$$

onde $a(s)$ são números complexos e $\epsilon(t)$ é a sequência definida no exemplo 2.1.

$X(t)$ assim definida é denominada "representação de médias móveis".

É claro que: $E[X(t)] = 0$ e

$$R_{XX}(u) = \sum_{s=0}^{\infty} \sigma^2 a(s) a(s+u).$$

Assim:

$$\begin{aligned} f_{XX}(\lambda) &= \frac{\sigma^2}{2\pi} \sum_{u=-\infty}^{\infty} \sum_{s=0}^{\infty} a(s) a(s+u) e^{-i\lambda u} = \\ &= \frac{\sigma^2}{2\pi} |A(\lambda)|^2, \text{ onde } A(\lambda) = \sum_{s=0}^{\infty} a(s) e^{-i\lambda s}. \end{aligned}$$

Obs.: No que segue quando dissermos que um processo é estacionário, isto significará que o processo é estacionário de 2ª ordem.

2.4 - ANÁLISE ESPECTRAL BIVARIADA

Sejam $\{X(t), t \in \mathbb{Z}\}$ e $\{Y(t), t \in \mathbb{Z}\}$ duas sequências estacionárias tais que $E[X(t)] = E[Y(t)] = 0$, e cujas funções de covariância são respectivamente:

$$R_{XX}(u) = E[X(t) X(t+u)],$$

$$R_{YY}(u) = E[Y(t) Y(t+u)].$$

Como já vimos, podemos escrever:

$$X(t) = \int_{-\pi}^{\pi} e^{i\lambda t} dZ_X(\lambda), \quad Y(t) = \int_{-\pi}^{\pi} e^{i\lambda t} dZ_Y(\lambda),$$

com

$$E \left[dZ_X(\lambda) \overline{dZ_X(\lambda')} \right] = \begin{cases} dF_{XX}(\lambda), & \lambda = \lambda' \\ 0, & \lambda \neq \lambda' \end{cases}$$

$$E \left[dZ_Y(\lambda) \overline{dZ_Y(\lambda')} \right] = \begin{cases} dF_{YY}(\lambda), & \lambda = \lambda' \\ 0, & \lambda \neq \lambda' \end{cases}$$

De modo análogo, no caso bivariado nós temos:

$$E \left[\overline{dZ_X(\lambda)} dZ_Y(\lambda') \right] = \begin{cases} dF_{XY}(\lambda), & \lambda = \lambda' \\ 0, & \lambda \neq \lambda' \end{cases} \quad (2.16)$$

DEFINIÇÃO 2.9 - A função $F_{XY}(\lambda)$ é denominada *função de distribuição espectral cruzada* associada aos processos X e Y . Se $F_{XY}(\lambda)$ é absolutamente contínua, a sua derivada $f_{XY}(\lambda)$ é denominada *função de densidade espectral cruzada* ou *espectro cruzado*.

Por ser função complexa, f_{XY} pode ser escrita como:

$$f_{XY}(\lambda) = c(\lambda) - iq(\lambda), \quad (2.17)$$

onde $c(\lambda)$ e $q(\lambda)$ são funções reais.

DEFINIÇÃO 2.10 - $c(\lambda)$ é denominado *co-espectro* e $q(\lambda)$ *espectro-quadratura* associados aos processos $X(t)$ e $Y(t)$.

A função de covariância cruzada entre $X(t)$ e $Y(t)$ é dada por:

$$R_{XY}(u) = E[\overline{X(t)}Y(t+u)] = E\left[\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{-i\lambda t} e^{i\lambda'(t+u)} \overline{dZ_X(\lambda)} dZ_Y(\lambda')\right]$$

$$= \int_{-\pi}^{\pi} e^{-i\lambda t} e^{i\lambda(t+u)} E[\overline{dZ_X(\lambda)} dZ_Y(\lambda')] = \int_{-\pi}^{\pi} e^{i\lambda u} dF_{XY}(\lambda).$$

Se $dF_{XY}(\lambda) = f_{XY}(\lambda) d\lambda$, então

$$R_{XY}(u) = \int_{-\pi}^{\pi} e^{i\lambda u} f_{XY}(\lambda) d\lambda. \quad (2.18)$$

Se $\int_{-\infty}^{\infty} |R_{XY}(u)| du < \infty$, então:

$$f_{XY}(\lambda) = \frac{1}{2\pi} \sum_{u=-\infty}^{\infty} e^{-i\lambda u} R_{XY}(u). \quad (2.19)$$

Assim, $f_{XY}(\lambda)$ é a transformada de Fourier de $R_{XY}(u)$.

- PROPRIEDADES:
- a) $R_{XY}(-u) = R_{YX}(u)$,
 - b) $f_{XY}(-\lambda) = f_{YX}(\lambda) = \overline{f_{XY}(\lambda)}$,
 - c) $c(-\lambda) = c(\lambda)$, $q(-\lambda) = -q(\lambda)$.

OBSERVAÇÃO - Usando a representação (2.5) teremos:

$$X(t) = \int_0^{\infty} [\cos \lambda t dU_X(\lambda) + \sin \lambda t dV_X(\lambda)],$$

$$Y(t) = \int_0^{\infty} [\cos \lambda t dU_Y(\lambda) + \sin \lambda t dV_Y(\lambda)].$$

Portanto $R_{XY}(u) = \int_{-\pi}^{\pi} [\cos \lambda u dC(\lambda) + \sin \lambda u dQ(\lambda)]$, onde:

$$dC(\lambda) = \frac{dU_X(\lambda)dU_Y(\lambda)}{2} = \frac{dV_X(\lambda)dV_Y(\lambda)}{2},$$

$$dQ(\lambda) = \frac{dU_X(\lambda)dV_Y(\lambda)}{2} = \frac{-dV_X(\lambda)dU_Y(\lambda)}{2}.$$

Vemos, então, que o co-espectro mede a covariância entre as duas componentes cosseno, ou seja, componentes *que estão em fase*, e o espectro-quadratura a covariância entre as componentes seno e cosseno, ou seja, componentes $\frac{\pi}{2}$ *fora de fase*.

2.5 - FILTROS

Muitas vezes algumas operações matemáticas são usadas no estudo de uma série temporal com o objetivo de facilitar a sua análise.

Entre estas operações, podemos destacar a classe das que são lineares e invariantes no tempo. Assim, seja A uma operação cujo domínio consiste de séries $X(t)$, $t \in \mathbb{Z}$ e cujo contradomínio consiste de séries $Y(t)$, $t \in \mathbb{Z}$. Por meio desta operação uma série $X(t)$ é transformada numa série $Y(t)$ e escrevemos então: $Y(t) = A[X(t)]$.

Dizemos que A é *linear* se para séries $X_1(t)$, $t \in \mathbb{Z}$ e $X_2(t)$, $t \in \mathbb{Z}$, pertencentes ao domínio de A , tivermos:

$$\begin{aligned} A[\alpha_1 X_1(t) + \alpha_2 X_2(t)] &= \alpha_1 A[X_1(t)] + \alpha_2 A[X_2(t)] = \\ &= \alpha_1 Y_1(t) + \alpha_2 Y_2(t), \end{aligned} \quad (2.20)$$

onde α_1 e α_2 são constantes.

Dizemos que A é *invariante no tempo* se

$$A[T^u[X(t)]] = A[X(t+u)] = Y(t+u) \quad (2.21)$$

para $t, u \in \mathbb{Z}$, onde T^u é o operador translação.

DEFINIÇÃO 2.11 - Um operador A com as propriedades dadas por (2.20) e (2.21) é denominado *filtro linear*.

EXEMPLO 2.3 - Uma classe usual de filtros lineares é dada por

$$Y(t) = \sum_{s=-\infty}^{\infty} a(s) X(t-s). \quad (2.22)$$

Dizemos que o filtro linear cuja função de impulso é $a(s)$, transforma a série $X(t)$ na série $Y(t)$.

$X(t)$ é a *entrada* do sistema e $Y(t)$ é a *saída*.

Suponhamos que $Y(t)$ seja uma sequência estacionária, e que

$$\sum_{s=-\infty}^{\infty} |a(s)| < \infty.$$

A função de covariância de $Y(t)$ fica então:

$$R_{YY}(u) = \sum_{s, r=-\infty}^{\infty} a(r) a(s) R_{XX}(u+r-s). \quad (2.23)$$

Os efeitos da filtragem da série $X(t)$ são dificilmente reconhecidos pelas expressões (2.22) e (2.23). Vamos então considerar a transformada de Fourier para os dois membros da expressão (2.23).

$$\frac{1}{2\pi} \sum_{u=-\infty}^{\infty} R_{YY}(u) e^{-i\lambda u} = \frac{1}{2\pi} \sum_{u=-\infty}^{\infty} \sum_{r,s=-\infty}^{\infty} a(r) a(s) R_{XX}(u+r-s) e^{-i\lambda u}.$$

Portanto:

$$f_{YY}(\lambda) = \frac{1}{2\pi} \sum_{r=-\infty}^{\infty} a(r) e^{i\lambda r} \sum_{s=-\infty}^{\infty} a(s) e^{-i\lambda s} \sum_{u=-\infty}^{\infty} R_{XX}(u) e^{-i\lambda u},$$

ou seja,

$$f_{YY}(\lambda) = |A(\lambda)|^2 f_{XX}(\lambda), \quad (2.24)$$

onde

$$A(\lambda) = \sum_{s=-\infty}^{\infty} a(s) e^{-i\lambda s}. \quad (2.25)$$

$A(\lambda)$ é denominada *função de transferência do filtro*.

A expressão (2.24) mostra que o espectro em cada frequência, é atenuado ou amplificado pelo fator $|A(\lambda)|^2$.

Podemos também escrever:

$$A(\lambda) = \Lambda_1(\lambda) + i\Lambda_2(\lambda) = G(\lambda) e^{i\phi(\lambda)}, \quad (2.26)$$

onde $G(\lambda) = [\Lambda_1^2(\lambda) + \Lambda_2^2(\lambda)]^{1/2}$ (2.27), representa o *ganho do fil*

tro, e

$$\phi(\lambda) = \text{arc tg} \left[\frac{\Lambda_2(\lambda)}{\Lambda_1(\lambda)} \right], \quad (2.28)$$

representa o *ângulo de fase* do filtro.

Se em (2.22) substituirmos $X(t)$ por sua representa
ção espectral, teremos

$$\begin{aligned} Y(t) &= \sum_{s=-\infty}^{\infty} a(s) \int_{-\pi}^{\pi} e^{i\lambda(t-s)} dZ_X(\lambda) = \int_{-\pi}^{\pi} \left[\sum_{s=-\infty}^{\infty} a(s) e^{-i\lambda s} \right] e^{i\lambda t} dZ_X(\lambda) = \\ &= \int_{-\pi}^{\pi} \Lambda(\lambda) e^{i\lambda t} dZ_X(\lambda) = \int_{-\pi}^{\pi} e^{i[\lambda t + \phi(\lambda)]} G(\lambda) dZ_X(\lambda). \end{aligned}$$

Portanto:

$$Y(t) = \int_{-\pi}^{\pi} e^{i[\lambda t + \phi(\lambda)]} G(\lambda) dZ_X(\lambda) = \int_{-\pi}^{\pi} e^{i\lambda t} \Lambda(\lambda) dZ_X(\lambda). \quad (2.29)$$

Assim, este filtro linear leva a uma mudança na am
plitude de onda, dada por $G(\lambda)$, e a um deslocamento no tem-
po, dado por $\frac{\phi(\lambda)}{\lambda}$.

Sejam $X(t)$ e $Y(t)$ duas sequências estacionárias tais
que:

$$Y(t) = \sum_{s=-\infty}^{\infty} a(s) X(t-s) + \eta(t), \quad (2.30)$$

com $\sum_{s=-\infty}^{\infty} a^2(s) < \infty$, onde $\eta(t)$ é sequência estacionária, tem mé-

dia zero e não é correlacionada com $X(t)$.

$\eta(t)$ responde por uma relação não linear entre as sequências $X(t)$ e $Y(t)$, ou pela influência de outro processo no comportamento de $Y(t)$.

A função de covariância cruzada entre $X(t)$ e $Y(t)$ é dada por

$$R_{XY}(u) = \sum_{s=-\infty}^{\infty} a(s)R_{XX}(u),$$

logo:

$$f_{XY}(\lambda) = A(\lambda)f_{XX}(\lambda) \quad (2.31)$$

e

$$f_{YY}(\lambda) = |A(\lambda)|^2 f_{XX}(\lambda) + f_{\eta\eta}(\lambda), \quad (2.32)$$

onde

$$A(\lambda) = \sum_{s=-\infty}^{\infty} a(s)e^{-i\lambda s}$$

é a função de transferência.

De (2.31) segue que:

$$A(\lambda) = \frac{f_{XY}(\lambda)}{f_{XX}(\lambda)} \quad \text{e} \quad |A(\lambda)| = \frac{[c^2(\lambda) + q^2(\lambda)]^{1/2}}{f_{XX}(\lambda)}.$$

A expressão

$$\gamma(\lambda) = \left[\frac{c^2(\lambda) + q^2(\lambda)}{f_{XX}(\lambda) f_{YY}(\lambda)} \right]^{1/2} = \frac{|f_{XY}(\lambda)|}{\sqrt{f_{XX}(\lambda) f_{YY}(\lambda)}}, \quad (2.33)$$

é denominada *coerência* na frequência λ .

O significado da coerência é análogo ao do coeficiente de correlação, ou seja, é uma medida adimensional do grau de correlação linear entre as duas séries, como função da frequência. As propriedades da correlação cruzada entre duas séries podem ser descritas pela coerência e pelo ângulo de fase.

De (2.32) e (2.33) segue que:

$$f_{\eta\eta}(\lambda) = f_{YY}(\lambda) [1 - \gamma^2(\lambda)]. \quad (2.34)$$

Podemos notar que $0 \leq \gamma^2(\lambda) \leq 1$.

Quando $\gamma^2(\lambda) = 1$ o espectro de η se anula, e existe uma perfeita correlação entre $X(t)$ e $Y(t)$ na frequência λ .

2.6 - FUNÇÃO DE COVARIÂNCIA AMOSTRAL

Queremos estimar a função de covariância $R_{XX}(u)$ da sequência estacionária $X(t)$, baseados em T observações

$$X(1), \dots, X(T).$$

Um possível estimador é dado por:

$$\hat{R}_{XX}(u) = \begin{cases} \frac{1}{T} \sum_{t=1}^{T-|u|} (X(t) - \bar{X})(X(t+|u|) - \bar{X}), & \text{se } |u| < T \\ 0 & , \text{ se } |u| \geq T \end{cases} \quad (2.35)$$

onde

$$\bar{X} = \frac{1}{T} \sum_{t=1}^T X(t).$$

No que segue, assumiremos que $E[X(t)] = 0$. Nestas condições:

$$\begin{aligned} E[\hat{R}_{XX}(u)] &= \frac{1}{T} \sum_{t=1}^{T-|u|} E[X(t)X(t+|u|)] = \\ &= \frac{1}{T} \sum_{t=1}^{T-|u|} R_{XX}(u) = \begin{cases} R_{XX}(u) [1 - \frac{|u|}{T}], & \text{se } |u| < T \\ 0 & , \text{ se } |u| \geq T \end{cases} \end{aligned} \quad (2.36)$$

Logo, $\hat{R}_{XX}(u)$ é um estimador viciado de $R_{XX}(u)$.

Se

$$R'_{XX}(u) = \begin{cases} \frac{1}{T-|u|} \sum_{t=1}^{T-|u|} X(t)X(t+u), & |u| < T \\ 0 & , |u| \geq T \end{cases}$$

fosse usado para estimar $R_{XX}(u)$, teríamos um estimador não viciado. Dois motivos nos levam a preferir $\hat{R}_{XX}(u)$ a $R'_{XX}(u)$ (ver [33], p.174):

- 1º) $\hat{R}_{XX}(u)$ conduz a menor erro médio quadrático,
 2º) $\hat{R}_{XX}(u)$ é positivo definido enquanto que $R'_{XX}(u)$ não o é.

A expressão para covariância é dada por (ver [25], p.176):

$$\begin{aligned} \text{Cov}[\hat{R}_{XX}(u_1), \hat{R}_{XX}(u_2)] &= \\ &= \frac{1}{T^2} \sum_{v=-T-u_1-1}^{T-u_2-1} \left\{ \phi(v) [R_{XX}(v)R_{XX}(v+u_2-u_1)] \right. \\ &\quad \left. + R_{XX}(v-u_1)R_{XX}(v+u_2) + K_4(-v, u_1, u_2) \right\}, \end{aligned} \quad (2.37)$$

onde

$$\phi(v) = \begin{cases} T-u_1-v, & 1-T+u_1 \leq v \leq 0 \\ T-u_1, & 0 \leq v \leq u_1-u_2 \\ T-u_2-v, & u_1-u_2 \leq v \leq T-u_2-1, u_2 \geq u_1 \end{cases}$$

e a função $K_4(-v, u_1, u_2)$ representa o cumulante de 4ª ordem da sequência $X(t)$. Se $X(t)$ é normal, $K_4 = 0$. Para outros processos mostra-se que, ao obtermos propriedades dos estimadores da autocovariância, a contribuição dada por K_4 pode ser desprezada (ver [4], p.27). Assim, daqui para frente, omitiremos este termo.

É bastante difícil computarmos a covariância por meio de (2.37). Quanto T é suficientemente grande, temos a seguinte aproximação (ver [5], pp.253-256):

$$\begin{aligned}
 \lim_{T \rightarrow \infty} T [\text{cov}(\hat{R}_{XX}(u_1), \hat{R}_{XX}(u_2))] &= \\
 &= \sum_{v=-\infty}^{\infty} [R_{XX}(v)R_{XX}(v+u_2-u_1) + R_{XX}(v+u_2)R_{XX}(v-u_1)] = \\
 &= 4\pi \int_{-\pi}^{\pi} f_{XX}^2(\omega) \cos \omega u_1 \cos \omega u_2 d\omega, \quad (2.38)
 \end{aligned}$$

logo:

$$\text{cov}[\hat{R}_{XX}(u_1), \hat{R}_{XX}(u_2)] \sim \frac{1}{T} \sum_{v=-\infty}^{\infty} [R_{XX}(v)R_{XX}(v+u_2-u_1) + R_{XX}(v+u_2)R_{XX}(v-u_1)] \quad (2.39)$$

e

$$\text{var}[\hat{R}_{XX}(u)] \sim \frac{1}{T} \sum_{v=-\infty}^{\infty} [R_{XX}^2(v) + R_{XX}(v+u)R_{XX}(v-u)]. \quad (2.40)$$

2.7 - ESPECTRO AMOSTRAL - PROPRIEDADES - DISTRIBUIÇÃO

Um estimador para $f_{XX}(\lambda)$ é dado por:

$$\tilde{f}_{XX}(\lambda) = \frac{1}{2\pi} \sum_{u=-T}^T \hat{R}_{XX}(u) e^{-i\lambda u}. \quad (2.41)$$

A média deste estimador é:

$$E[\tilde{f}_{XX}(\lambda)] = \begin{cases} \frac{1}{2\pi} \sum_{u=-T}^T \left(1 - \frac{|u|}{T}\right) R_{XX}(u) e^{-i\lambda u}, & \text{se } |u| < T \\ 0, & \text{se } |u| \geq T \end{cases} \quad (2.42)$$

Para T suficientemente grande, temos:

$$E[\tilde{f}_{XX}(\lambda)] = f_{XX}(\lambda). \quad (2.43)$$

Apesar de $\tilde{f}_{XX}(\lambda)$ ser assintoticamente não viciado, ele não é consistente, no sentido de que sua variância não diminui quando T cresce.

• Estimadores com menor variância podem ser obtidos fazendo-se uso de "funções de peso" (lag window).

Se $h_M(u)$ é uma função peso, ela satisfaz às seguintes condições:

- a) $h_M(0) = 1$,
- b) $h_M(u) = h_M(-u)$,
- c) $h_M(u) = 0$, $|u| > M$, $M \leq T$.

Assim, um estimador de $f_{XX}(\lambda)$, denominado *estimador suavizado*, é dado por:

$$\hat{f}_{XX}(\lambda) = \frac{1}{2\pi} \sum_{u=-M}^M h_M(u) \hat{R}_{XX}(u) e^{-i\lambda u}. \quad (2.44)$$

A transformada de Fourier de $h_M(u)$ é denominada *núcleo* ou *janela espectral*, e é dada por:

$$H_M(\lambda) = \frac{1}{2\pi} \sum_{u=-\infty}^{\infty} h_M(u) e^{-i\lambda u}.$$

Muitas funções peso têm sido sugeridas. A tabela 1 destaca algumas das mais empregadas, com suas respectivas janelas espectrais.

TABELA 1 - FUNÇÕES PESO E NÚCLEOS

REPRESENTAÇÃO	FUNÇÃO PESO	JANELA ESPECTRAL
Retangular	$h_M(u) = \begin{cases} 1, & u \leq M \\ 0, & u > M \end{cases}$	$H_M(\lambda) = 2M \left(\frac{\text{sen } 2\pi\lambda M}{2\pi\lambda M} \right)$
Tukey- Hanning	$h_M(u) = \begin{cases} \frac{1}{2} \left(1 + \cos \frac{\pi u}{M} \right), & u \leq M \\ 0, & u > M \end{cases}$	$\begin{aligned} H_M(\lambda) &= M \left\{ \frac{\text{sen } 2\pi\lambda M}{2\pi\lambda M} + \frac{1}{2} \frac{\text{sen } 2\pi M \left(\lambda + \frac{1}{2} \right)}{2\pi M \left(\lambda + \frac{1}{2} \right)} \right. \\ &\quad \left. + \frac{1}{2} \frac{\text{sen } 2\pi M \left(\lambda - \frac{1}{2} \right)}{2\pi M \left(\lambda - \frac{1}{2} \right)} \right\} = \\ &= M \left(\frac{\text{sen } 2\pi\lambda M}{2\pi\lambda M} \right) \left(\frac{1}{1 - (2\pi\lambda M)^2} \right) \end{aligned}$
Parzen	$h_M(u) = \begin{cases} 1 - 6 \left(\frac{ u }{M} \right)^2 + 6 \left(\frac{ u }{M} \right)^3, & u \leq \frac{M}{2} \\ 2 \left(1 - \frac{ u }{M} \right)^3, & \frac{M}{2} < u \leq M \\ 0, & u > M \end{cases}$	$H_M(\lambda) = \frac{3}{4} M \left(\frac{\text{sen } 2\pi\lambda M/2}{\pi\lambda M/2} \right)^4$

FONTE: [25].

É fácil ver que:

$$\hat{f}_{XX}(\lambda) = \int_{-\pi}^{\pi} \tilde{f}_{XX}(\omega) H_M(\lambda - \omega) d\omega \quad (2.45)$$

De (2.45) e (2.43) segue que:

PROPRIEDADE 2.1

Para T suficientemente grande

$$E[\hat{f}_{XX}(\lambda)] \sim \int_{-\pi}^{\pi} f_{XX}(\omega) H_M(\lambda - \omega) d\omega = \bar{f}_{XX}(\lambda). \quad (2.46)$$

$\bar{f}_{XX}(\lambda)$ é denominado *espectro suavizado* de $X(t)$.

Gostaríamos de encontrar uma função peso, de tal modo que:

- a) o vício $V(\lambda) = E[\hat{f}_{XX}(\lambda)] - f_{XX}(\lambda) \sim \bar{f}_{XX}(\lambda) - f_{XX}(\lambda)$, seja pequeno (condição de fidelidade);
- b) a variância de $\hat{f}_{XX}(\lambda)$ seja pequena (condição de estabilidade).

Para que possamos reduzir o vício do estimador, $H_M(\alpha)$ deve estar altamente concentrada em torno de $\alpha = 0$. Muitas medidas desta concentração têm sido propostas, e recebem o nome de *larguras de faixa*.

Aqui, usaremos a sugerida por Parzen, que é dada por:

$$b_M = \frac{1}{H_M(0)} = \frac{2\pi}{\sum_{u=-M}^M h_M(u)}. \quad (2.47)$$

A coluna (1) da tabela 2, mostra a largura de faixa para as funções peso consideradas anteriormente.

Verificamos que quanto maior for M , menor é a largura de faixa. Apesar da função peso retangular apresentar menor largura de faixa, não podemos considerá-la como sendo

TABELA 2

REPRESENTAÇÕES	(1) LARGURA DE FAIXA	(2) VARIÂNCIA/ $f_{XX}^2(\lambda)$	(3) GRAUS DE LIBERDADE
Retangular	$\pi/M + \frac{1}{2}$	$2,00 \cdot M/T$	T/M
Tukey-Hanning	$2\pi/M$	$0,75 \cdot M/T$	$2,7 \cdot T/M$
Parzen	$8\pi/3M$	$0,54 \cdot M/T$	$3,7 \cdot T/M$

Observação: Nas frequências zero e $\pm\pi$ a razão variância/ $f_{XX}^2(\lambda)$ é o dobro daquela indicada na coluna (2) e o nº de graus de liberdade é a metade do indicado na coluna (3).

a melhor, antes de examinarmos as variâncias dos respectivos estimadores.

Usando a expressão 2.38 obtemos

PROPRIEDADE 2.2 -

$$\lim_{T \rightarrow \infty} T \text{ cov}[\hat{f}_{XX}(\lambda), \hat{f}_{XX}(\lambda')] = \pi \int_{-\pi}^{\pi} f_{XX}^2(\omega) [H_M(\omega-\lambda) + H_M(\omega+\lambda)] \times [H_M(\omega-\lambda') + H_M(\omega+\lambda')] d\omega, \quad (2.48)$$

e

$$\lim_{M \rightarrow \infty} \lim_{T \rightarrow \infty} M^{-1} T \text{ cov}[\hat{f}_{XX}(\lambda), \hat{f}_{XX}(\lambda')] = \begin{cases} cf_{XX}^2(\lambda), & \lambda = \pm\lambda' \neq 0, \pm\pi \\ 2cf_{XX}^2(\lambda), & \lambda = \lambda' = 0, \lambda = \pm\lambda' = \pm\pi, \\ 0, & \lambda \neq \lambda' \end{cases} \quad (2.49)$$

onde

$$c = \begin{cases} 0,75 & \text{se a função peso usada for a de Tukey-Hanning} \\ 0,54 & \text{se a função peso usada for a de Parzen.} \end{cases}$$

PROPRIEDADE 2.3 -

Para T e M suficientemente grandes, temos:

$$\text{var}[\hat{f}_{XX}(\lambda)] \sim (\delta_{\gamma}, \psi_{T,M}) f_{XX}^2(\lambda), \quad (2.50)$$

onde

$$\psi_{T,M} = \frac{2\pi}{T} \int_{-\pi}^{\pi} H_M^2(\omega) d\omega = \frac{cM}{T} \quad (2.51)$$

e

$$\delta_{\lambda} = \begin{cases} 2, & \lambda = 0, \pm \pi \\ 1, & \lambda \neq 0, \pm \pi \end{cases}$$

Na coluna (2) da tabela 2, temos a razão entre a variância e $f_{XX}^2(\lambda)$, para as três janelas espectrais consideradas. A variância do estimador de Parzen é cerca de 96% da variância do estimador de Tukey-Hanning, usando-se a mesma largura de faixa. Para que isso possa ser obtido, o número de "lags" M usado no cálculo do estimador de Tukey-Hanning, deve ser igual a $\frac{3}{4}$ do número de lags usado no de Parzen.

Assim, diminuindo-se M conseguimos diminuir a variância, mas aumentamos o vício do estimador. Existe portanto um compromisso entre as condições a) e b). Quando T cresce, é razoável que M também cresça. No entanto, se o aumento relativo em M for bem menor do que em T , nós podemos me-

lhorar a estabilidade do nosso estimador ao mesmo tempo que conseguimos reduzir o vício.

PROPRIEDADE 2.4 - (ver [25], pp.253-254).

Para T suficientemente grande, mostra-se que:

$$\frac{\hat{v}f_{XX}(\lambda)}{f_{XX}(\lambda)} \sim \chi^2_v, \text{ onde } v = 2 \frac{\{E[f_{XX}(\lambda)]\}^2}{\text{var}[f_{XX}(\lambda)]} \sim \frac{2}{\delta_\lambda \psi_{T,M}}.$$

A coluna (3) da tabela 2, mostra o número de graus de liberdade da distribuição de $\frac{\hat{v}f_{XX}(\lambda)}{f_{XX}(\lambda)}$, quando consideramos os três estimadores suavizados dados pelas três funções peso usadas anteriormente.

2.8 - PERIODOGRAMA

Seja $X(t)$ sequência estacionária. Com base nas observações $X(1), \dots, X(T)$, podemos computar a transformada de Fourier finita:

$$d_X^T(\lambda) = \sum_{t=1}^T X(t) e^{-i\lambda t}, \quad -\infty < \lambda < +\infty.$$

A estatística

$$I_{XX}(\lambda) = \frac{1}{2\pi T} |d_X^T(\lambda)|^2, \quad -\infty < \lambda < +\infty, \quad (2.52)$$

introduzida por Schuster [11], é denominada *periodograma* dos valores $X(1), \dots, X(T)$.

Quando T é suficientemente grande, podemos mostrar que: (ver [13], p.105)

$$\int_{-\pi}^{\pi} H_M(\lambda - \omega) I_{XX}(\omega) d\omega = \frac{1}{2\pi} \sum_{u=-M}^M h_M(u) \hat{R}_{XX}(u) e^{-i\lambda u} = \hat{f}_{XX}(\lambda). \quad (2.53)$$

Assim, o estimador suavizado é uma média ponderada do periodograma, com peso concentrado na vizinhança λ .

Vemos então, que podemos calcular o nosso estimador de dois modos: determinando a função de covariância amostral, ou por meio do periodograma. Para uma série temporal de comprimento T , teríamos que fazer cerca de T^2 operações para determinarmos o espectro estimado usando o periodograma ponderado. Em 1965, Cooley e Tukey [9] descreveram um algoritmo para calcular os coeficientes de Fourier, com uma considerável economia no tempo de cálculo. Este algoritmo é denominado "fast Fourier transform", e reduz o número de operações da ordem de T^2 para ordem de $T \log_2 T$.

Neste trabalho, o espectro será estimado por meio da soma ponderada das covariâncias amostrais.

2.9 - ESTIMADOR DO ESPECTRO CRUZADO, COERÊNCIA, GANHO E ÂNGULO DE FASE

Um estimador da função de covariância cruzada entre $X(t)$ e $Y(t)$, se estes têm média zero, é dado por:

$$\hat{R}_{XY}(u) = \frac{1}{T} \sum_{t=1}^{T-|u|} X(t)Y(t+u). \quad (2.54)$$

Assim, podemos estimar o espectro cruzado, o co-espectro e o espectro-quadratura, respectivamente por:

$$\hat{f}_{XY}(\lambda) = \hat{c}(\lambda) - i\hat{q}(\lambda) = \frac{1}{2\pi} \sum_{u=-M}^M h_M(u) \hat{R}_{XY}(u) e^{-i\lambda u}, \quad (2.55)$$

$$\hat{c}(\lambda) = \frac{1}{2\pi} \sum_{u=-M}^M h_M(u) \hat{R}_{XY}(u) \cos \lambda u, \quad (2.56)$$

$$\hat{q}(\lambda) = \frac{1}{2\pi} \sum_{u=-M}^M h_M(u) \hat{R}_{XY}(u) \sin \lambda u. \quad (2.57)$$

Os estimadores da coerência, ganho e ângulo de fase são dados respectivamente, por:

$$\hat{\gamma}(\lambda) = \frac{|\hat{f}_{XY}(\lambda)|}{[\hat{f}_{XX}(\lambda)\hat{f}_{YY}(\lambda)]^{1/2}}, \quad (2.58)$$

$$\hat{G}(\lambda) = \frac{|\hat{f}_{XY}(\lambda)|}{\hat{f}_{XX}(\lambda)}, \quad (2.59)$$

$$\hat{\phi}(\lambda) = \text{tg}^{-1} \left[\frac{-\hat{q}(\lambda)}{\hat{c}(\lambda)} \right]. \quad (2.60)$$

2.10 - PROPRIEDADES ASSINTÓTICAS DOS ESTIMADORES - DISTRIBUIÇÕES AMOSTRAIS

De modo análogo ao da secção 2.6 (ver também [13])

p.375, [24] cap.9 e [25], podemos mostrar que para T suficientemente grande, valem as seguintes propriedades:

PROPRIEDADE 2.5 -

$$\begin{aligned} E[\hat{f}_{XY}(\lambda)] &\sim \int_{-\pi}^{\pi} H_M(\lambda-\omega) f_{XY}(\omega) d\omega, \\ E[\hat{c}(\lambda)] &\sim \int_{-\pi}^{\pi} H_M(\lambda-\omega) c(\omega) d\omega, \\ E[\hat{q}(\lambda)] &\sim \int_{-\pi}^{\pi} H_M(\lambda-\omega) q(\omega) d\omega; \end{aligned} \quad (2.61)$$

PROPRIEDADE 2.6 -

$$\begin{aligned} \text{cov}[\hat{f}_{XX}(\lambda), \hat{f}_{XY}(\lambda')] &\sim \int_{-\pi}^{\pi} J_{T,M}(\omega, \lambda, \lambda') f_{XX}(\omega) f_{XY}(\omega) d\omega, \\ \text{cov}[\hat{f}_{XX}(\lambda), \hat{f}_{YY}(\lambda')] &\sim \int_{-\pi}^{\pi} J_{T,M}(\omega, \lambda, \lambda') |f_{XY}(\omega)|^2 d\omega, \end{aligned} \quad (2.62)$$

onde

$$J_{T,M}(\omega, \lambda, \lambda') = \frac{1}{\pi T} [\{ H_M(\omega-\lambda') + H_M(\omega+\lambda') \} [H_M(\omega-\lambda) + H_M(\omega+\lambda)]];$$

PROPRIEDADE 2.7 -

$$\begin{aligned} \text{cov}[\hat{f}_{XY}(\lambda), \hat{f}_{XY}(\lambda')] &\sim \frac{1}{2\pi T} \int_{-\pi}^{\pi} [f_{XX}(\omega) f_{YY}(\omega) H_M(\omega+\lambda') \\ &\quad + f_{XY}^2(\omega) H_M(\omega-\lambda')] H_M(\omega-\lambda) d\omega, \\ \text{cov}[\hat{f}_{XY}(\lambda), \hat{f}_{XY}^*(\lambda')] &\sim \frac{1}{2\pi T} \int_{-\pi}^{\pi} [f_{XX}(\omega) f_{YY}(\omega) H_M(\omega-\lambda') \\ &\quad + f_{XY}^2(\omega) H_M(\omega+\lambda')] H_M(\omega-\lambda) d\omega, \end{aligned} \quad (2.63)$$

onde $\hat{f}_{XY}^*(\lambda')$ é o complexo conjugado de $\hat{f}_{XY}(\lambda)$.

PROPRIEDADE 2.8 -

Se os espectros são aproximadamente constantes sobre a largura da faixa e $\lambda = \lambda'$, a matriz de covariâncias assintóticas de $\hat{f}_{XX}(\lambda)$, $\hat{f}_{YY}(\lambda)$, $\hat{c}(\lambda)$, $\hat{q}(\lambda)$ é dada por

$$\delta_\lambda \psi_{T,M} \begin{bmatrix} f_{XX}(\lambda) & c^2(\lambda) + q^2(\lambda) & f_{XX}(\lambda)c(\lambda) & f_{XX}(\lambda)q(\lambda) \\ & f_{YY}(\lambda) & f_{YY}(\lambda)c(\lambda) & f_{YY}(\lambda)q(\lambda) \\ & & \frac{1}{2}[f_{XX}(\lambda)f_{YY}(\lambda) + c^2(\lambda) - q^2(\lambda)] & c(\lambda)q(\lambda) \\ & & & \frac{1}{2}[f_{XX}(\lambda)f_{YY}(\lambda) - c^2(\lambda) + q^2(\lambda)] \end{bmatrix}, \quad (2.64)$$

onde $\psi_{T,M}$ e δ_λ são dados por (2.51).

Na matriz dada em (2.64), o elemento da 1ª linha e 1ª coluna representa a variância assintótica de $\hat{f}_{XX}(\lambda)$, o elemento da 1ª linha e 2ª coluna representa a covariância assintótica entre $\hat{f}_{XX}(\lambda)$ e $\hat{f}_{YY}(\lambda)$, o elemento da 1ª linha e 3ª coluna representa a covariância assintótica entre $f_{XX}(\lambda)$ e $c(\lambda)$, e assim por diante.

Se os espectros amostrais não variam muito ao redor de suas médias e os espectros são aproximadamente constantes sobre a largura de faixa, por meio da expansão de Taylor (ver [13], p.132), podemos verificar as seguintes propriedades:

PROPRIEDADE 2.9 -

$$E[\hat{\gamma}(\lambda) - \gamma(\lambda)] \sim \delta_{\lambda} \psi_{T,M} [1 - \gamma^2(\lambda)]^2 / 4\gamma(\lambda),$$

$$E[\hat{G}(\lambda) - G(\lambda)] \sim \delta_{\lambda} \psi_{T,M} [1 - \gamma^2(\lambda)] / 4\gamma^2(\lambda),$$

$$E[\hat{\phi}(\lambda) - \phi(\lambda)] \sim 0.$$

PROPRIEDADE 2.10 -

$$\text{Var}[\hat{\gamma}(\lambda)] \sim \delta_{\lambda} \psi_{T,M} [1 - \gamma^2(\lambda)]^2 / 2,$$

$$\text{Var}[\hat{G}(\lambda)] \sim \delta_{\lambda} \psi_{T,M} [1 - \gamma^2(\lambda)] \frac{f_{YY}(\lambda)}{2f_{XX}(\lambda)},$$

$$\text{Var}[\hat{\phi}(\lambda)] \sim \delta_{\lambda} \psi_{T,M} [\gamma^{-2}(\lambda) - 1] / 2.$$

As propriedades 2.9 e 2.10 são válidas para $|f_{XY}|^2 > 0$.

Com base na teoria desenvolvida por Goodman em [16] para variáveis aleatórias normais complexas, podemos determinar a distribuição aproximada para o estimador do espectro bivariado. As propriedades abaixo seguem como extensão desta teoria.

Sejam $X(t)$ e $Y(t)$ dois processos estacionários normais de média zero, e sejam

$$\hat{f}(\lambda) = \begin{bmatrix} \hat{f}_{XX}(\lambda) & \hat{c}(\lambda) - i\hat{q}(\lambda) \\ \hat{c}(\lambda) + i\hat{q}(\lambda) & \hat{f}_{YY}(\lambda) \end{bmatrix}$$

e

$$f(\lambda) = \begin{bmatrix} f_{XX}(\lambda) & c(\lambda) - iq(\lambda) \\ c(\lambda) + iq(\lambda) & f_{YY}(\lambda) \end{bmatrix},$$

então:

PROPRIEDADE 2.11 -

$$n\hat{f}(\lambda) \sim W_2^C(n, f(\lambda)), \text{ para } 0 < |\lambda| < \pi$$

$$n\hat{f}(\lambda) \sim W_2(n, f(\lambda)), \text{ para } |\lambda| = 0, \pi,$$

onde $n = \frac{1}{\psi_{T,M}}$;

PROPRIEDADE 2.12 -

A estatística

$$n_\lambda^2 \text{Det}(\hat{f}(\lambda)) / \text{Det}(\hat{f}(\lambda)) = [n_\lambda \hat{f}_{\eta\eta}(\lambda) / f_{\eta\eta}(\lambda)] [n_\lambda \hat{f}_{XX}(\lambda) / f_{XX}(\lambda)],$$

onde

$$\hat{f}_{\eta\eta}(\lambda) = \hat{f}_{YY}(\lambda) [1 - \gamma^2(\lambda)],$$

$$f_{\eta\eta}(\lambda) = f_{YY}(\lambda) [1 - \gamma^2(\lambda)],$$

$$n_\lambda = \frac{x_\lambda}{\psi_{T,M}}, \quad x_\lambda = \frac{1}{\delta_\lambda},$$

pode ser vista como o produto de duas variáveis qui-quadrado independentes, com $(n_\lambda - x_\lambda)$ e n_λ graus de liberdade, respectivamente. Portanto:

$$n_{\lambda} \frac{\hat{f}_{\eta\eta}(\lambda)}{f_{\eta\eta}(\lambda)} \sim \chi^2_{(n_{\lambda} - x_{\lambda})}$$

$$n_{\lambda} \frac{\hat{f}_{XX}(\lambda)}{f_{\eta\eta}(\lambda)} \sim \chi^2_{n_{\lambda}}$$

PROPRIEDADE 2.13 -

Quando n_{λ} cresce, a distribuição conjunta de $n_{\lambda} \hat{f}_{XX}(\lambda)$, $n_{\lambda} \hat{f}_{YY}(\lambda)$, $n_{\lambda} \hat{c}(\lambda)$ e $n_{\lambda} \hat{q}(\lambda)$, converge para uma distribuição normal multivariada.

2.11 - PROCEDIMENTO PARA TESTAR A HIPÓTESE DE COERÊNCIA NULA

Muitas vezes é de interesse testar a hipótese de que a coerência entre duas sequências é zero. O procedimento de teste que descrevemos abaixo pode ser visto com maiores detalhes em [13] pp.135-137.

Suponha que computamos $(L+1)$ estimadores de $\gamma(\lambda)$ no intervalo $(0, \pi)$, tais que

$$\text{cov}[\hat{\gamma}(\lambda_j), \hat{\gamma}(\lambda_k)] = 0, \quad \lambda_j = \frac{\pi j}{L}, \quad j=0, 1, \dots, L.$$

Seja

$$\Delta_L = \left\{ \sum_{j=0}^L \{ [\hat{\gamma}(\lambda_j) - p_{n_j}] / [(L+1)q_{n_j}]^{1/2} \} \right\}, \quad (2.65)$$

onde

$$p_n = \frac{(\pi)^{1/2} \Gamma(n)}{2\Gamma(n+1/2)}, \quad q_n = \frac{1}{n} - \pi \left[\frac{\Gamma(n)}{2\Gamma(n+1/2)} \right]^2, \quad n > 2$$

(Γ denota a função gama).

e

$$n_j = \frac{1}{\delta_{\lambda_j} \psi_{T,M}} = \frac{x_{\lambda_j}}{\psi_{T,M}}.$$

Sob a hipótese de coerência zero, segue do teorema do limite central que Δ_L é assintoticamente normalmente distribuída, com média zero e variância 1.

Assim, rejeitamos $H_0: \gamma = 0$, ao nível de significância α , se $|\Delta_L| > z_c$, onde z_c é tal que

$$1 - \phi(z_c) = \alpha,$$

$$\phi(z_c) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_c} e^{-t^2/2} dt.$$

CAPÍTULO 3

MODELO DE DEFASAGENS DISTRIBUÍDAS: ESTIMAÇÃO

3.1 - OBJETIVOS

Neste capítulo nos propomos a estudar o modelo de defasagens distribuídas

$$Y(t) = \sum_{s=-r_1}^{r_2} a(s)X(t-s) + \eta(t), \quad (3.1)$$

supondo que $X(t)$, $Y(t)$ e $\eta(t)$ são sequências estocásticas estacionárias de média zero, que $\eta(t)$ é normal, e que $X(t)$ e $\eta(t)$ são independentes. Dados T pares de observações $(X(t), Y(t)), t=1, \dots, T$, veremos como a estimação de

$$a' = (a(-r_1), \dots, a(0), \dots, a(r_2))$$

pode ser feita com base na teoria espectral.

A técnica espectral na estimação de modelos de defasagens distribuídas, foi introduzida por Hannan [22], no caso em que os coeficientes $a(s)$ decrescem exponencialmente. Este caso particular também iremos abordar aqui.

As sequências $X(t)$ e $Y(t)$ podem ter média diferente de zero. Na secção 3.6 veremos como é feito o ajuste para este caso.

3.2 - CONSIDERAÇÕES GERAIS

a) De (3.1) segue que

$$\begin{aligned} \hat{R}_{\eta\eta}(u) &= \frac{1}{T} \sum_{t=1}^{T-|u|} \eta(t)\eta(t+u) = \frac{1}{T} \sum_{t=1}^{T-|u|} \left[Y(t) - \sum_{s=-r_1}^{r_2} a(s)X(t-s) \right] \\ &\times \left[Y(t+u) - \sum_{j=-r_1}^{r_2} a(j)X(t+u-j) \right] = \hat{R}_{YY}(u) + \sum_{j,s=-r_1}^{r_2} a(j)a(s)\hat{R}_{XX}(u-j+s) \\ &- \sum_{s=-r_1}^{r_2} a(s) \left[\hat{R}_{XY}(u+s) + \hat{R}_{XY}(-u+s) \right]. \end{aligned}$$

Assim:

$$\hat{f}_{\eta\eta}(\lambda) = \hat{f}_{YY}(\lambda) + \sum_{j,s=-r_1}^{r_2} a(j)a(s)\hat{g}_{XX}(\lambda) - 2 \sum_{s=-r_1}^{r_2} a(s)\text{Re}[\hat{g}_{XY}(\lambda)], \quad (3.2)$$

onde

$$\hat{g}_{XX}(\lambda) = \frac{1}{2\pi} \sum_{u=-M}^M h_M(u) \hat{R}_{XX}(u-j+s) e^{-i\lambda u} \quad (3.3)$$

e

$$\hat{g}_{XY}(\lambda) = \frac{1}{2\pi} \sum_{u=-M}^M h_M(u) \hat{R}_{XY}(u+s) e^{-i\lambda u}. \quad (3.4)$$

Como

$$\begin{aligned} E[\hat{g}_{XX}(\lambda)] &\sim \frac{1}{2\pi} \sum_{u=-M}^M h_M(u) R_{XX}(u+s-j) e^{-i\lambda u} = \\ &= \int_{-\pi}^{\pi} H_M(\lambda-\omega) f_{XX}(\omega) e^{i\omega(s-j)} d\omega, \end{aligned} \quad (3.5)$$

e como $f_{XX}(\omega) e^{i\omega(s-j)}$ oscila quando ω varia em $(-\pi, \pi)$, notamos que necessitaremos de um número de lags M bem maior para obtermos um estimador fiel a $f_{XX}(\omega) e^{i\omega(s-j)}$, do que aquele que seria necessário para estimar $f_{XX}(\omega)$. Vemos então, que a fidelidade do nosso estimador fica prejudicada pelo fato de $h_M(u)$ não estar centrada em $\hat{R}_{XX}(0)$. Por este motivo, é conveniente substituírmos $\hat{g}_{XX}(\lambda)$ por $\hat{f}_{XX}(\lambda) e^{i\lambda(s-j)}$.

Podemos verificar que:

$$E[\hat{g}_{XX}(\lambda) - \hat{f}_{XX}(\lambda) e^{i\lambda(s-j)}] \sim \int_{-\pi}^{\pi} H_M(\lambda-\omega) f_{XX}(\omega) [e^{i\omega(s-j)} - e^{i\lambda(s-j)}] d\omega,$$

e portanto quando H torna-se mais concentrada ao redor de $\lambda = \omega$, o erro da nossa aproximação converge para zero.

Usando argumento semelhante ao dado acima, substituiremos também $\hat{g}_{XY}(\lambda)$ por $\hat{f}_{XY}(\lambda) e^{i\lambda s}$. Deste modo a expressão (3.2) fica:

$$\hat{f}_{\eta\eta}(\lambda) = \hat{f}_{YY}(\lambda) + \sum_{j,s=-r_1}^{r_2} a(j)a(s) \hat{f}_{XX}(\lambda) e^{i\lambda(s-j)} - 2 \sum_{s=-r_1}^{r_2} a(s) \text{Re}[\hat{f}_{XY}(\lambda) e^{i\lambda s}] \quad (3.6)$$

b) De (3.1) segue também que:

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^{T-|u|} Y(t)X(t-s+u) &= \frac{1}{T} \sum_{t=1}^{T-|u|} \sum_{s=-r_1}^{r_2} a(s)X(t-s)X(t-s+u) \\ &+ \frac{1}{T} \sum_{t=1}^{T-|u|} \eta(t)X(t-s+u), \end{aligned}$$

e portanto:

$$\hat{R}_{XY}(u-s) = \sum_{s=-r_1}^{r_2} a(s) \hat{R}_{XX}(u) + \hat{R}_{X\eta}(u-s).$$

Usando o mesmo argumento que usamos em a) temos:

$$\hat{f}_{X\eta}(\lambda) = \hat{f}_{XY}(\lambda) - \hat{f}_{XX}(\lambda) \sum_{s=-r_1}^{r_2} a(s) e^{-i\lambda s}. \quad (3.7)$$

c) Sabemos que o periodograma dos valores $X(1), \dots, X(T)$, é dado por:

$$\begin{aligned} I_{XX}(\lambda) &= \frac{1}{2\pi T} \left| \sum_{t=1}^T X(t) e^{-i\lambda t} \right|^2 = \frac{1}{2\pi T} \left[\sum_{t=1}^T X(t) e^{-i\lambda t} \right] \\ &\times \left[\sum_{s=1}^T X(s) e^{i\lambda s} \right] = \frac{1}{2\pi} \sum_{u=-T+1}^{T-1} \hat{R}_{XX}(u) e^{-i\lambda u}. \end{aligned}$$

Seja então $I_{XY}(\lambda)$, definido como

$$I_{XY}(\lambda) = \frac{1}{2\pi T} \left[\sum_{t=1}^T X(t) e^{-i\lambda t} \right] \left[\sum_{s=1}^T Y(s) e^{-i\lambda s} \right] = \frac{1}{2\pi T} \sum_{u=-T+1}^{T-1} \hat{R}_{XY}(u) e^{-i\lambda u},$$

o *periodograma cruzado* dos valores $X(t)$, $t=1, \dots, T$ e $Y(t)$, $t=1, \dots, T$.

De (3.1), usando o argumento dado em a) decorre que:

$$I_{\eta\eta}(\lambda) = I_{YY}(\lambda) + \sum_{j,s=-r_1}^{r_2} a(j)a(s)I_{XX}(\lambda)e^{i\lambda(s-j)} - 2 \sum_{s=-r_1}^{r_2} a(s)\text{Re}[I_{XY}(\lambda)e^{i\lambda s}]. \quad (3.8)$$

3.3 - O ENFOQUE ESPECTRAL PARA A ESTIMAÇÃO DO VETOR DE COEFICIENTES

Por (3.1), a distribuição conjunta de $X(t)$ e $Y(t)$ nas observações amostrais $[(X(1),Y(1)), \dots, (X(T),Y(T))]$, pode ser transformada na distribuição conjunta dos elementos residuais $[\eta(1), \dots, \eta(T)]$.

Lembrando que $\eta(t)$ é normal com média zero, e que temos apenas um registro amostral $[(X(1),Y(1)), \dots, (X(T),Y(T))]$, se pensarmos em determinar o estimador de máxima verossimilhança para o vetor de parâmetros a' , devemos considerar a função de verossimilhança. Esta, para os T pares de observações dependentes é dada por:

$$f(a' | \eta_T) = \frac{\exp(-1/2 \eta_T' \Gamma^{-1} \eta_T)}{(2\pi)^{T/2} \text{Det}(\Gamma)^{1/2}},$$

onde

$$\eta_T' = (\eta(1), \dots, \eta(T)),$$

e Γ é a matriz de covariâncias positiva definida dada por:

$$\Gamma = \begin{bmatrix} R_{\eta\eta}(0) & R_{\eta\eta}(1) \dots R_{\eta\eta}(T-1) \\ R_{\eta\eta}(-1) & R_{\eta\eta}(0) \dots \\ \vdots & \\ R_{\eta\eta}(1-T) & \dots \dots \dots R_{\eta\eta}(0) \end{bmatrix}.$$

Queremos determinar o vetor a' que maximize f ou o logaritmo de f , dado por

$$\log_e [f(a' | \eta_T)] = -\frac{T}{2} \log_e 2\pi - \frac{1}{2} \log_e \text{Det}(\Gamma) - \frac{1}{2} \eta_T' \Gamma^{-1} \eta_T \quad (3.9)$$

Como Γ geralmente é desconhecida, não podemos determinar diretamente o estimador de máxima verossimilhança. Se $\eta(t)$ fosse o ruído branco, Γ seria diagonal e precisaríamos apenas estimar $R_{\eta\eta}(0) = \sigma^2$. Neste caso a maximização convencional de $\log_e [f(a' | \eta_T)]$ levaria ao melhor estimador linear não viciado, que seria também o estimador de mínimos quadrados não ponderado. Como, em geral, $\eta(t)$ é uma sequência autocorrelacionada, a estimação de Γ é um pouco mais complicada. Para isto podemos usar a regressão linear de mínimos quadrados para estimar a' , computar os resíduos pela equação de regressão e usá-los para estimar Γ . O estimador de Γ será estatisticamente consistente, desde que r_1 e r_2 tenham sido especificados corretamente na regressão inicial, e substituído em (3.9) fornecerá um estimador para a' .

Uma outra alternativa é estimar $R_{\eta\eta}(u)$ por meio de

$$\hat{R}_{\eta\eta}(u) = \frac{1}{2m} \sum_{j=-m+1}^m \hat{f}_{\eta\eta}(\lambda_j) e^{i\lambda_j u}, \quad \lambda_j = \frac{\pi j}{m},$$

onde $\frac{\pi}{m}$ é o intervalo de frequência entre os estimadores das ordenadas do espectro. O estimador de Γ , assim obtido, substituído em (3.9) fornecerá um estimador de a' . Este último método para estimar Γ é independente da especificação de r_1 e r_2 . O enfoque espectral devido a Hannan para a estimação de um modelo com defasagens distribuídas, está baseado neste procedimento. Este enfoque foi revisto por Fishman em [13] quando o modelo está sujeito às restrições apresentadas na secção 3.1, e é o que procuraremos abordar com detalhes a seguir.

Considerando a matriz Γ , observamos que ela é simétrica e que todos os elementos da diagonal principal são iguais. Ela é portanto uma forma de Toeplitz, e pode ser aproximadamente diagonalizada por uma transformação que é independente de seus elementos. Para isto seja W uma matriz $T \times T$ onde o elemento da h -ésima linha e k -ésima coluna é dado por

$$\omega_{hk} = \frac{1}{\sqrt{T}} e^{i\lambda_h(k-n)}, \quad \lambda_h = \frac{\pi(h-n)}{n}, \quad n = [T/2].$$

Tomando $D = W\Gamma W^*$, onde $W^* = \bar{W}'$ é a matriz transposta conjugada

da de W , o elemento da h -ésima linha e ℓ -ésima coluna de D é dado por:

$$d_{h\ell} = \frac{1}{T} \sum_{r=1}^T \sum_{s=1}^T R_{\eta\eta}(s-r) e^{i\lambda_h(r-n)} e^{-i\lambda_\ell(s-n)}.$$

Fazendo $r - n = k$ e $s - n = j$ temos:

$$d_{h\ell} = \frac{1}{T} \sum_{j=-n+1}^{T-n} \sum_{k=-n+1}^{T-n} R_{\eta\eta}(j-k) e^{-i(\lambda_\ell j - \lambda_h k)}.$$

Vamos supor, para facilidade de exposição, que $T = 2n$. Assim:

$$d_{h\ell} = \frac{1}{T} \sum_{j,k=-n+1}^n R_{\eta\eta}(k-j) e^{i(\lambda_\ell j - \lambda_h k)},$$

e substituindo $R_{\eta\eta}(k-j)$ por sua representação espectral te mos:

$$\begin{aligned} d_{h\ell} &= \frac{1}{T} \int_{-\pi}^{\pi} f_{\eta\eta}(\lambda) \left[\sum_{j=-n+1}^n e^{i(\lambda_\ell - \lambda)j} \sum_{k=-n+1}^n e^{i(\lambda - \lambda_h)k} \right] d\lambda = \frac{1}{T} \int_{-\pi}^{\pi} f_{\eta\eta}(\lambda) e^{\frac{i(\lambda_\ell - \lambda)}{2}} \\ &\quad \times e^{\frac{i(\lambda - \lambda_h)}{2}} \left[\sum_{j=-n+1}^n e^{i(\lambda_\ell - \lambda)(j-1/2)} \sum_{k=-n+1}^n e^{i(\lambda - \lambda_h)(k-1/2)} \right] d\lambda = \\ &= \frac{1}{T} \int_{-\pi}^{\pi} f_{\eta\eta}(\lambda) e^{\frac{i(\lambda_\ell - \lambda_h)}{2}} \left[\sum_{j=-n+1/2}^{n-1/2} e^{i(\lambda_\ell - \lambda)j} \sum_{k=-n+1/2}^{n-1/2} e^{i(\lambda - \lambda_h)k} \right] d\lambda. \end{aligned}$$

Como

$$\sum_{j=-n+1/2}^{n-1/2} e^{i\lambda_j} = \frac{e^{i\lambda(-n+\frac{1}{2})} - e^{i\lambda(n+\frac{1}{2})}}{1 - e^{i\lambda}} = \frac{e^{\frac{i\lambda}{2}} [e^{-i\lambda n} - e^{i\lambda n}]}{e^{\frac{i\lambda}{2}} \left[e^{-\frac{i\lambda}{2}} - e^{\frac{i\lambda}{2}} \right]} = \frac{\text{sen } n\lambda}{\text{sen } \frac{\lambda}{2}},$$

vem

$$\begin{aligned} d_{h\ell} &= \frac{1}{T} \int_{-\pi}^{\pi} f_{\eta\eta}(\lambda) e^{\frac{i(\lambda_{\ell} - \lambda_h)}{2}} \left[\frac{\text{sen}(\lambda_{\ell} - \lambda) \text{sen } n(\lambda_h - \lambda)}{\text{sen} \frac{1}{2}(\lambda_{\ell} - \lambda) \text{sen} \frac{1}{2}(\lambda_h - \lambda)} \right] d\lambda = \\ &= \int_{-\pi}^{\pi} f_{\eta\eta}(\lambda) e^{\frac{i(\lambda_{\ell} - \lambda_h)}{2}} \left[\frac{\text{sen } n(\lambda_{\ell} - \lambda) \text{sen } n(\lambda_h - \lambda)}{T \text{sen} \frac{1}{2}(\lambda_{\ell} - \lambda) \text{sen} \frac{1}{2}(\lambda_h - \lambda)} \right] d\lambda. \end{aligned}$$

Notamos que:

i) se $\lambda_{\ell} = \lambda_h = \lambda$, temos:

$$d_{hh} = \int_{-\pi}^{\pi} f_{\eta\eta}(\lambda_h) \lim_{\lambda \rightarrow \lambda_h} \left[\frac{\text{sen } n(\lambda_h - \lambda)}{T \text{sen} \frac{1}{2}(\lambda_h - \lambda)} \right]^2 d\lambda,$$

mas

$$\lim_{\lambda \rightarrow \lambda_h} \left[\frac{\text{sen } n(\lambda_h - \lambda)}{T \text{sen} \frac{1}{2}(\lambda_h - \lambda)} \right]^2 = \lim_{\lambda \rightarrow \lambda_h} \left[\frac{\text{sen } n(\lambda_h - \lambda)}{T n(\lambda_h - \lambda)} \cdot \frac{n(\lambda_h - \lambda)}{\text{sen} \frac{1}{2}(\lambda_h - \lambda)} \cdot \frac{\frac{1}{2}(\lambda_h - \lambda)}{\frac{1}{2}(\lambda_h - \lambda)} \right]^2 = \left(\frac{2n}{T} \right)^2 = 1.$$

logo:

$$d_{hh} = 2\pi f_{\eta\eta}(\lambda_h).$$

ii) se $\lambda = \lambda_h \neq \lambda_{\ell}$, temos:

$$d_{h\ell} = \int_{-\pi}^{\pi} f_{\eta\eta}(\lambda_h) e^{\frac{i(\lambda_{\ell} - \lambda_h)}{2}} \left[\frac{\text{sen } n(\lambda_{\ell} - \lambda_h)}{\text{sen} \frac{1}{2}(\lambda_{\ell} - \lambda_h)} \right] \cdot \lim_{\lambda \rightarrow \lambda_h} \left[\frac{\text{sen } n(\lambda_h - \lambda)}{T \text{sen} \frac{1}{2}(\lambda_h - \lambda)} \right] d\lambda,$$

mas

$$\lim_{\lambda \rightarrow \lambda_h} \left[\frac{\text{sen } n(\lambda_h - \lambda)}{T \text{sen} \frac{1}{2}(\lambda_h - \lambda)} \right] = 1,$$

logo

$$\begin{aligned} d_{h\ell} &= \int_{-\pi}^{\pi} f_{\eta\eta}(\lambda_h) e^{\frac{i(\lambda_\ell - \lambda_h)}{2}} \cdot \frac{\text{sen } n(\lambda_\ell - \lambda_h)}{\text{sen} \frac{1}{2}(\lambda_\ell - \lambda_h)} d\lambda = \\ &= 2\pi f_{\eta\eta}(\lambda_h) \left[\cos \frac{(\lambda_\ell - \lambda_h)}{2} - i \text{sen} \frac{(\lambda_\ell - \lambda_h)}{2} \right] \cdot \frac{\text{sen } n(\lambda_\ell - \lambda_h)}{\text{sen} \frac{1}{2}(\lambda_\ell - \lambda_h)} \\ &= 2\pi f_{\eta\eta}(\lambda_h) \text{sen}(\lambda_\ell - \lambda_h) \left[\frac{\cos \frac{1}{2}(\lambda_\ell - \lambda_h)}{\text{sen} \frac{1}{2}(\lambda_\ell - \lambda_h)} - i \right] = \\ &= 2\pi f_{\eta\eta}(\lambda_h) \text{sen} \pi(\ell - h) \left[\frac{\cos \frac{1}{2}(\lambda_\ell - \lambda_h)}{\text{sen} \frac{1}{2}(\lambda_\ell - \lambda_h)} - i \right] = 0, \end{aligned}$$

pois $(\ell - h) \in \mathbb{Z}$.

iii) Se $\lambda \neq \lambda_\ell \neq \lambda_u \neq \lambda$, então quando T torna-se grande, a expressão

$$\frac{\text{sen } n(\lambda_\ell - \lambda_h) \cdot \text{sen } n(\lambda_h - \lambda)}{T \text{sen} \frac{1}{2}(\lambda_\ell - \lambda) \cdot \text{sen} \frac{1}{2}(\lambda_h - \lambda)}$$

converge para zero.

Deste modo, pelo que mostramos em i), ii) e iii) temos:

$$\lim_{T \rightarrow \infty} d_{h\ell} = \begin{cases} 2\pi f_{\eta\eta}(\lambda_h), & h = \ell \\ 0, & h \neq \ell, \end{cases}$$

ou seja, D é, assintoticamente, uma matriz diagonal cujo h -ésimo elemento é $2\pi f_{\eta\eta}(\lambda_h)$. Consequentemente, D^{-1} é, assintoticamente, também diagonal e seu h -ésimo elemento é

$$[2\pi f_{\eta\eta}(\lambda_h)]^{-1}.$$

O elemento da ℓ -ésima linha e h -ésima coluna da matriz $W \cdot W^*$ é:

$$\frac{1}{T} e^{i(\lambda_\ell - \lambda_h)n} \sum_{k=1}^{2n} e^{i(\lambda_h - \lambda_\ell)k} = \begin{cases} 1, & \text{se } h = \ell \\ 0, & \text{se } h \neq \ell, \end{cases}$$

o que mostra que W é unitária, ou seja $W \cdot W^* = I$. Assim,

$$D^{-1} = (W \Gamma W^*)^{-1} = W \Gamma^{-1} W^*,$$

$$\Gamma^{-1} = W^* D^{-1} W,$$

e, assintoticamente,

$$\eta_T' \Gamma^{-1} \eta_T = \eta_T' W^* D^{-1} W \eta_T = \frac{1}{2\pi T} \sum_{\ell=1}^{2n} \sum_{t=1}^{2n} \sum_{v=1}^{2n} \eta_v e^{-i(t-n)(\lambda_v - \lambda_\ell)} f_{\eta\eta}^{-1}(\lambda_t) \eta_\ell.$$

Como $\lambda_v = \frac{\pi(v-n)}{n}$, $\lambda_\ell = \frac{\pi(\ell-n)}{n}$, $\lambda_t = \frac{\pi(t-n)}{n}$, fazendo $v-n = u$, $\ell-n = r$ e $t-n = j$, teremos:

$$\lambda_u = \frac{\pi u}{n}, \quad \lambda_r = \frac{\pi r}{n} \quad \text{e} \quad \lambda_j = \frac{\pi j}{n}.$$

Deste modo, assintoticamente,

$$\begin{aligned} \eta_T' \Gamma^{-1} \eta_T &= \frac{1}{2\pi T} \sum_{j=-n+1}^n \left[\sum_{r=-n+1}^n \eta_{r+n} e^{ij\lambda_r} \sum_{u=-n+1}^n \eta_{u+n} e^{-ij\lambda_u} \right] f_{\eta\eta}^{-1}(\lambda_j) = \\ &= \sum_{j=-n+1}^n I_{\eta\eta}(\lambda_j) f_{\eta\eta}^{-1}(\lambda_j), \end{aligned} \quad (3.10)$$

onde

$$I_{\eta\eta}(\lambda_j) = \frac{1}{2\pi T} \left| \sum_{r=-n+1}^n \eta_{r+n} e^{-i\lambda_j r} \right|^2.$$

Substituindo (3.8) em (3.10), obtemos:

$$\begin{aligned} \eta_T' \Gamma^{-1} \eta_T &= \sum_{j=-n+1}^n f_{\eta\eta}^{-1}(\lambda_j) \left\{ I_{YY}(\lambda_j) + \sum_{k,s=-r_1}^{r_2} a(k)a(s) I_{XX}(\lambda_j) e^{i\lambda_j(s-k)} \right. \\ &\quad \left. - 2 \sum_{s=-r_1}^{r_2} a(s) \operatorname{Re} \left[I_{XY}(\lambda_j) e^{i\lambda_j s} \right] \right\}. \end{aligned}$$

Observando que $I_m[I_{XY}(\lambda)] = -I_m[I_{XY}(-\lambda)]$, temos:

$$\begin{aligned} \eta_T' \Gamma^{-1} \eta_T &= \sum_{j=-n+1}^n f_{\eta\eta}^{-1}(\lambda_j) \left[I_{YY}(\lambda_j) + \sum_{k,s=-r_1}^{r_2} a(k)a(s) I_{XX}(\lambda_j) e^{i\lambda_j(s-k)} \right. \\ &\quad \left. - 2 \sum_{s=-r_1}^{r_2} a(s) I_{XY}(\lambda_j) e^{i\lambda_j s} \right]. \end{aligned}$$

Maximizar (3.9) com relação a a' , corresponde a minimizar $\eta_T' \Gamma^{-1} \eta_T$ com relação a a' . Assim, calculando

$$\frac{\partial \eta_T' \Gamma^{-1} \eta_T}{\partial a'}$$

e igualando a zero, chegamos a um sistema de $r_1 + r_2 - 1$ equações da forma

$$\sum_{s=-r_1}^{r_2} \hat{a}(s) \sum_{j=-n+1}^n f_{\eta\eta}^{-1}(\lambda_j) I_{XX}(\lambda_j) e^{i\lambda_j(\ell-s)} = \sum_{j=-n+1}^n f_{\eta\eta}^{-1}(\lambda_j) I_{XY}(\lambda_j) e^{i\lambda_j \ell}, \quad (3.11)$$

$\ell = -r_1, -r_1 + 1, \dots, 0, \dots, r_2 - 1, r_2$, onde $\hat{a}(s)$ é o estimador de $a(s)$.

Como $I_{XX}(\lambda_j)$ e $I_{XY}(\lambda_j)$ não são estimadores consistentes, em (3.11) nós os substituiremos pelos estimadores suavizados $\hat{f}_{XX}(\lambda_j)$ e $\hat{f}_{XY}(\lambda_j)$.

Se $\frac{\pi}{n}$ for consideravelmente menor do que a largura de faixa do estimador, que é $b_M = \frac{1}{H_M(0)}$, e se o espectro é relativamente constante sobre a largura da faixa, não esperamos muita variação entre as ordenadas adjacentes do espectro. Deste modo ainda teremos um quadro representativo do espectro, se n for substituído por um número menor m .

Com as substituições que citamos acima, as equações (3.11) podem ser escritas na forma:

$$\hat{\mathbf{a}}' \cdot \hat{\mathbf{H}} = \hat{\mathbf{b}}', \quad (3.12)$$

onde o elemento da ℓ -ésima linha e s -ésima coluna de $\hat{\mathbf{H}}$ é da
do por

$$\hat{h}_{\ell s} = \frac{1}{2m} \sum_{j=-m+1}^m f_{\eta\eta}^{-1}(\lambda_j) \hat{f}_{XX}(\lambda_j) e^{i\lambda_j(\ell-s)}, \quad (3.13)$$

e o elemento da ℓ -ésima coluna de $\hat{\mathbf{b}}'$ é dado por:

$$\hat{b}_{\ell} = \frac{1}{2m} \sum_{j=-m+1}^m f_{\eta\eta}^{-1}(\lambda_j) f_{XY}(\lambda_j) e^{i\lambda_j \ell}, \quad (3.14)$$

$$\lambda_j = \frac{\pi j}{m}.$$

Resta ainda determinar um estimador consistente pa
ra $f_{\eta\eta}(\lambda)$. Lembrando que

$$f_{\eta\eta}(\lambda_j) = f_{YY}(\lambda_j) [1 - \gamma_{XY}^2(\lambda_j)], \quad (3.15)$$

onde $\gamma_{XY}(\lambda_j)$ é a coerência na frequência λ_j (definida pela
expressão (2.33)), estimaremos $f_{\eta\eta}(\lambda_j)$ por

$$\hat{f}_{\eta\eta}(\lambda_j) = \hat{f}_{YY}(\lambda_j) [1 - \hat{\gamma}_{XY}^2(\lambda_j)], \quad (3.16)$$

onde $\hat{\gamma}_{XY}(\lambda_j)$ é a coerência amostral em λ_j (definida pela ex
pressão (2.58)).

Assim, substituindo-se (3.16) em (3.12), obtemos a
forma final das equações de estimação.

Convém observar que na expressão (3.13) o peso de cada coeficiente é dado pela razão

$$\frac{\hat{f}_{XX}(\lambda_j)}{\hat{f}_{\eta\eta}(\lambda_j)} = \frac{\hat{f}_{XX}(\lambda_j)}{\hat{f}_{YY}(\lambda_j) [1 - \gamma_{XY}^2(\lambda_j)]}. \quad (3.17)$$

Portanto, cada peso é função da coerência e da razão

$$\frac{\hat{f}_{XX}(\lambda_j)}{\hat{f}_{YY}(\lambda_j)}.$$

A razão $\frac{f_{XX}(\lambda)}{f_{\eta\eta}(\lambda)}$ é denominada quociente sinal/ruído. Se esta razão é constante para todo λ em $(-\pi, \pi)$, isto é, se

$$f_{\eta\eta}^{-1}(\lambda) f_{XX}(\lambda) = K, \quad \lambda \in (-\pi, \pi), \quad (3.18)$$

então de (3.11) decorre que:

$$\sum_{s=-r_1}^{r_2} \hat{a}(s) \cdot k \left[\frac{1}{2m} \sum_{j=-m+1}^m e^{i\lambda_j(\ell-s)} \right] = \frac{1}{2m} \sum_{j=-m+1}^m f_{\eta\eta}^{-1}(\lambda_j) \hat{f}_{XY}(\lambda_j) e^{i\lambda_j \ell},$$

e portanto:

$$\hat{a}(s) = \frac{1}{2m} \sum_{j=-m+1}^m f_{XX}^{-1}(\lambda_j) \hat{f}_{XY}(\lambda_j) e^{i\lambda_j \ell}. \quad (3.19)$$

Deste modo, se pudermos supor que o espectro do processo re residual se aproxima do espectro do processo $X(t)$, poderemos estimar sucessivos $a(\ell)$'s sem que precisemos recomputar os anteriores. Isto é desejável, particularmente quando r_1 e r_2 são desconhecidos *à priori*, e não temos nenhum conhecimento

que nos permita fazer boas atribuições para eles.

3.4 - A ESTIMAÇÃO DA ESTRUTURA DE DEFASAGENS GEOMÉTRICA

Como vimos no capítulo 1, a suposição de que os coeficientes do modelo satisfaçam à restrição

$$a(s) = \alpha \cdot \beta^s, \quad s=1,2,\dots, \quad |\beta| < 1, \quad (3.20)$$

além de facilitar a estimação, em alguns casos é decorrência de considerações econômicas. Se (3.20) for substituído em (3.1), obtemos:

$$Y(t) = \alpha \sum_{s=-r_1}^{r_2} \beta^s X(t-s) + \eta(t). \quad (3.21)$$

De (3.21) decorre que:

$$Y(t) = \alpha X(t) + \beta Y(t-1) + \eta(t) - \beta \eta(t-1). \quad (3.22)$$

Como vemos, $Y(t)$ é uma transformação de $X(t)$ produzida por um filtro linear onde a função de transferência é dada por:

$$B(\lambda) = \sum_{s=0}^{\infty} a(s) e^{-i\lambda s} = \alpha (1 - \beta e^{-i\lambda})^{-1} = \alpha \cdot A(\lambda), \quad A(\lambda) = (1 - \beta e^{-i\lambda})^{-1} \quad (3.23)$$

Assim, $f_{XY}(\lambda) = B(\lambda) f_{XX}(\lambda) = \alpha A(\lambda) \cdot f_{XX}(\lambda)$, e o logaritmo da função de verossimilhança, ignorando termos cons-

tantes, é aproximadamente:

$$(2m)^{-1} \sum_{j=-m+1}^m \left\{ f_{\eta\eta}^{-1}(\lambda_j) \left[\hat{f}_{YY}(\lambda_j) + \alpha^2 |A(\lambda)|^2 \hat{f}_{XX}(\lambda_j) - \alpha A^*(\lambda_j) \hat{f}_{XY}(\lambda_j) - \alpha A(\lambda_j) \hat{f}_{XY}^*(\lambda_j) \right] \right\}. \quad (3.24)$$

Maximizando (3.24) com relação a α e β obtemos:

$$\begin{aligned} \frac{\hat{\alpha}}{2m} \sum_{j=-m+1}^m \frac{\hat{f}_{XX}(\lambda_j)}{\hat{f}_{ZZ}(\lambda_j)} &= \frac{1}{2m} \sum_{j=-m+1}^m \frac{\hat{f}_{XY}(\lambda_j)}{\hat{f}_{ZZ}(\lambda_j)} - \frac{\hat{\beta}}{2m} \sum_{j=-m+1}^m \frac{\hat{f}_{XY}(\lambda_j)}{\hat{f}_{ZZ}(\lambda_j)} e^{-i\lambda_j}, \\ \frac{\hat{\alpha}}{2m} \sum_{j=-m+1}^m \frac{\hat{f}_{XX}(\lambda_j) e^{i\lambda_j} \hat{A}^*(\lambda_j)}{\hat{f}_{ZZ}(\lambda_j)} &= \frac{1}{2m} \sum_{j=-m+1}^m \frac{\hat{f}_{XY}(\lambda_j) e^{i\lambda_j} \hat{A}^*(\lambda_j)}{\hat{f}_{ZZ}(\lambda_j)} \\ &\quad - \frac{\hat{\beta}}{2m} \sum_{j=-m+1}^m \frac{\hat{f}_{XY}(\lambda_j) \hat{A}^*(\lambda_j)}{\hat{f}_{ZZ}(\lambda_j)}, \end{aligned}$$

onde

$$\hat{f}_{ZZ}(\lambda_j) = \frac{f_{\eta\eta}(\lambda_j)}{|\hat{A}^*(\lambda_j)|^2}$$

e $\hat{A}^*(\lambda_j)$ é o complexo conjugado de $\hat{A}(\lambda_j)$.

Rearranjando os termos destas equações, obtemos:

$$\frac{1}{2m} \sum_{j=-m+1}^m \frac{\hat{f}_{XY}(\lambda_j)}{\hat{f}_{ZZ}(\lambda_j)} = \frac{1}{2m} \sum_{j=-m+1}^m \frac{\hat{f}_{XX}(\lambda_j) \hat{\alpha} + \hat{f}_{XY}(\lambda_j) e^{-i\lambda_j} \hat{\beta}}{\hat{f}_{ZZ}(\lambda_j)},$$

$$\frac{1}{2m} \sum_{j=-m+1}^m \frac{\hat{f}_{XY}(\lambda_j) e^{i\lambda_j \hat{A}^*(\lambda_j)}}{\hat{f}_{ZZ}(\lambda_j)} = \frac{1}{2m} \sum_{j=-m+1}^m \frac{\hat{\alpha} \hat{f}_{XX}(\lambda_j) e^{i\lambda_j \hat{A}^*(\lambda_j)} + \hat{f}_{XY}(\lambda_j) \hat{A}^*(\lambda_j) \hat{\beta}}{\hat{f}_{ZZ}(\lambda_j)} \quad (3.25)$$

ou ainda

$$\frac{1}{2m} \sum_{j=-m+1}^m \hat{f}_{ZZ}^{-1}(\lambda_j) \hat{f}_{XY}(\lambda_j) \begin{bmatrix} 1 \\ \hat{A}^*(\lambda_j) e^{i\lambda_j} \end{bmatrix} = \hat{H} \cdot \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix}, \quad (3.26)$$

onde

$$\hat{H} = \frac{1}{2m} \sum_{j=-m+1}^m \hat{f}_{ZZ}^{-1}(\lambda_j) \begin{bmatrix} \hat{f}_{XY}(\lambda_j) & \hat{f}_{XY}(\lambda_j) e^{-i\lambda_j} \\ \hat{A}^*(\lambda_j) \hat{f}_{XX}(\lambda_j) e^{i\lambda_j} & \hat{A}^*(\lambda_j) \hat{f}_{XY}(\lambda_j) \end{bmatrix}$$

Observamos que a segunda equação de (3.25) não é linear em $\hat{\beta}$. Portanto este sistema só pode ser resolvido iterativamente. Um estimador preliminar $\tilde{\beta}$ de β , é então usado para estimar $\hat{f}_{ZZ}(\lambda_j)$ e $\hat{A}(\lambda_j)$; em geral $\tilde{\beta}$ é o estimador de Leviatan [28] ou de Klein [26]. Este procedimento pode ser visto com maiores detalhes em [11], p.285 ou [22].

3.5 - A ESCOLHA DE m

Sabemos que as ordenadas do espectro são tomadas sobre o intervalo $(-\pi, \pi)$. Um problema importante é o da determinação do espaçamento adequado entre duas ordenadas adjacentes do espectro. Assim, se $\frac{\pi j}{m}$ é este espaçamento, a

pergunta que se faz é: como determinar m de modo a dar uma representação adequada à estrutura espectral? É razoável supor que o espectro de interesse é aproximadamente constante sobre a largura de faixa da janela espectral considerada. Como as larguras de faixa das janelas espectrais de Tukey-Hanning e Parzen são, respectivamente, $\frac{2\pi}{M}$ e $\frac{2,6\pi}{M}$, então tomar $m = \frac{M}{2}$ deve resultar numa boa descrição do espectro. Esta escolha apresenta ainda a vantagem de reduzir o número de parâmetros.

3.6 - O AJUSTAMENTO PARA MÉDIAS DIFERENTES DE ZERO

Se $X(t)$ e $Y(t)$ não possuem média igual a zero, devemos considerar a especificação:

$$Y(t) = \mu + \sum_{s=-r_1}^{r_2} a(s)X(t-s) + \eta(t),$$

e para estimar μ usamos:

$$\hat{\mu} = \bar{Y} - \bar{X} \sum_{s=-r_1}^{r_2} \hat{a}(s),$$

onde $\bar{X}$ e $\bar{Y}$ são as médias amostrais de X e Y , respectivamente.

No modelo de declínio geométrico tomamos

$$\hat{\mu} = \bar{Y} - \frac{\hat{\alpha}\bar{X}}{1-\hat{\beta}}.$$

CAPÍTULO 4

MODELO DE DEFASAGENS DISTRIBUÍDAS:

PROPRIEDADES DOS ESTIMADORES

4.1 - PROPRIEDADES ASSINTÓTICAS DO ESTIMADOR $\hat{a}'$

No capítulo 3 vimos que o estimador de máxima verossimilhança do vetor a' de coeficientes do modelo (3.1), quando se supõe que $X(t)$, $Y(t)$ e $\eta(t)$ são sequências estocásticas estacionárias de média zero, que $\eta(t)$ é normal, e que $X(t)$ e $\eta(t)$ são independentes, pode ser obtido através das equações

$$\sum_{s=-r_1}^{r_2} \hat{a}(s) \sum_{j=-n+1}^{\eta} \hat{f}_{\eta\eta}^{-1}(\lambda_j) \hat{f}_{XX}(\lambda_j) e^{i\lambda_j(\ell-s)} = \sum_{j=-n+1}^{\eta} \hat{f}_{\eta\eta}^{-1}(\lambda_j) \hat{f}_{XY}(\lambda_j) e^{i\lambda_j \ell},$$
$$\ell = -r_1, -r_1+1, \dots, 0, \dots, r_2-1, r_2. \quad (4.1)$$

Com o objetivo de facilitar a exposição das propriedades assintóticas deste estimador, iremos inicialmente supor $f_{\eta\eta}(\lambda)$ conhecido, ou seja, não faremos uso das equações de estimação na forma (4.1) mas sim na forma (3.12), quando

ainda não havíamos substituído $f_{\eta\eta}(\lambda)$ por seu estimador. Também, no caso em que o quociente sinal/ruído é constante, não substituiremos nas equações de estimação (ver (3.19)) $f_{XX}(\lambda)$ por seu estimador.

No final desta secção veremos qual o efeito produzido quando efetuamos estas substituições.

Em (3.14) vamos substituir $\hat{f}_{XY}(\lambda_j)$ por sua expressão dada em (3.7). Temos então:

$$\begin{aligned}\hat{b}_\ell &= \frac{1}{2m} \sum_{j=-m+1}^m f_{\eta\eta}^{-1}(\lambda_j) \left[\hat{f}_{X\eta}(\lambda_j) + \hat{f}_{XX}(\lambda_j) \sum_{s=-r_1}^{r_2} a(s) e^{-i\lambda_j s} \right] e^{i\lambda_j \ell} \\ &= \frac{1}{2m} \sum_{j=-m+1}^m f_{\eta\eta}^{-1}(\lambda_j) \hat{f}_{X\eta}(\lambda_j) e^{i\lambda_j \ell} + \frac{1}{2m} \sum_{j=-m+1}^m \hat{f}_{\eta\eta}^{-1}(\lambda_j) \hat{f}_{XX}(\lambda_j) \sum_{s=-r_1}^{r_2} a(s) e^{i\lambda_j(\ell-s)} \\ &= \hat{c}_\ell + \sum_{s=-r_1}^{r_2} \left[\frac{1}{2m} \sum_{j=-m+1}^m f_{\eta\eta}^{-1}(\lambda_j) \hat{f}_{XX}(\lambda_j) e^{i\lambda_j(\ell-s)} \right] a(s) \\ &= \hat{c}_\ell + \sum_{s=-r_1}^{r_2} a(s) \cdot \hat{h}_{\ell s},\end{aligned}$$

$$\text{onde} \quad \hat{c}_\ell = \frac{1}{2m} \sum_{j=-m+1}^m f_{\eta\eta}^{-1}(\lambda_j) \hat{f}_{X\eta}(\lambda_j) e^{i\lambda_j \ell} \quad (4.2)$$

Logo,

$$\hat{b}' = \hat{c}' + \hat{a}' \cdot \hat{H},$$

e por (3.12)

$$(\hat{a}' - a') \cdot \hat{H} = \hat{c}' \quad (4.3)$$

Vamos primeiramente mostrar que a matriz H com os elementos dados por

$$h_{s\ell} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_{\eta\eta}^{-1}(\omega) f_{XX}(\omega) e^{i\omega(\ell-s)},$$

é o limite em probabilidade da matriz $\hat{H}$. Assim, se em (3.13) substituirmos $\hat{f}_{XX}(\lambda_j)$ pela sua expressão, temos:

$$\hat{h}_{\ell s} = \frac{1}{4\pi m} \sum_{j=-m+1}^m f_{\eta\eta}^{-1}(\lambda_j) \sum_{u=-M}^M h_M(u) \hat{R}_{XX}(u) e^{-i\lambda_j(\ell-s-u)}.$$

Como $f_{\eta\eta}(\lambda)$ é positiva, segue por um teorema dado por Wiener (ver [12], pp.159-160), que existe uma série de Fourier absolutamente convergente tal que:

$$f_{\eta\eta}^{-1}(\lambda) = \frac{1}{2\pi} \sum_{v=-\infty}^{\infty} \mu_v e^{-i\lambda v}, \quad \sum_{v=-\infty}^{\infty} |\mu_v| < \infty$$

Então,

$$\hat{h}_{\ell s} = \left(\frac{1}{2\pi}\right)^2 \sum_{v=-\infty}^{\infty} \mu_v \sum_{u=-\infty}^{\infty} h_M(u) \hat{R}_{XX}(u) \left[\frac{1}{2m} \sum_{j=-m+1}^m e^{i\lambda_j(\ell-s-u-v)} \right]$$

e como

$$\sum_{j=-m+1}^m e^{i\lambda_j(\ell-s-u-v)} = \begin{cases} 2m, & \text{se } \ell-s-u=v \\ 0, & \text{se } \ell-s-u \neq v \end{cases}$$

vem:

$$\hat{h}_{\ell s} = \left(\frac{1}{2\pi} \right)^2 \sum_{u=-M}^M \mu_{\ell-s-u} h_M(u) \hat{R}_{XX}(u).$$

Pelo que vimos na secção 2.6 do capítulo 2, podemos concluir que $\hat{R}_{XX}(u)$ converge em média quadrática para $R_{XX}(u)$, portanto podemos afirmar que $\hat{R}_{XX}(u)$ converge para a probabilidade para $R_{XX}(u)$. Sendo assim, o limite em probabilidade de $\hat{h}_{\ell s}$ é dado por:

$$p \lim_{T \rightarrow \infty} \hat{h}_{s\ell} = \left(\frac{1}{2\pi} \right)^2 \sum_{u=-M}^M \mu_{\ell-s-u} h_M(u) R_{XX}(u).$$

Como, também,

$$\lim_{M \rightarrow \infty} h_M(u) = 1, \quad \forall u,$$

$\hat{h}_{s\ell}$ converge em probabilidade para

$$h_{s\ell} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_{\eta\eta}^{-1}(\omega) f_{XX}(\omega) e^{i\omega(\ell-s)} d\omega, \quad (4.4)$$

pois

$$\begin{aligned} \lim_{M \rightarrow \infty} \left(\frac{1}{2\pi} \right)^2 \sum_{u=-M}^M \mu_{\ell-s-u} h_M(u) R_{XX}(u) &= \left(\frac{1}{2\pi} \right)^2 \sum_{u=-\infty}^{\infty} \mu_{\ell-s-u} \int_{-\infty}^{\infty} f_{XX}(\omega) e^{i\omega u} d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[(2\pi)^{-1} \sum_{u=-\infty}^{\infty} \mu_{\ell-s-u} e^{-i\omega(\ell-s-u)} \right] f_{XX}(\omega) e^{i\omega(\ell-s-u)} e^{i\omega u} d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f_{\eta\eta}^{-1}(\omega) f_{XX}(\omega) e^{i\omega(\ell-s)} d\omega. \end{aligned}$$

Para a demonstração das propriedades que aparecem a seguir, além das suposições que já fizemos a respeito das sequências $X(t)$, $Y(t)$ e $\eta(t)$, acrescentaremos a de que $X(t)$ seja também sequência normal.

PROPRIEDADE 4.1 - Para T , M e m suficientemente grandes, $\hat{c}'$ tem distribuição aproximadamente normal, com vetor de médias zero e matriz de covariâncias $T^{-1}H$.

De (4.2) segue que:

$$E(\hat{c}_\ell) = \frac{1}{2m} \sum_{j=-m+1}^m f_{\eta\eta}^{-1}(\lambda_j) E[\hat{f}_{X\eta}(\lambda_j)] e^{i\lambda_j \ell}.$$

Como, para T e M suficientemente grandes

$$E[\hat{f}_{X\eta}(\lambda)] \sim f_{X\eta}(\lambda) = 0,$$

então $\hat{c}'$, assintoticamente, tem vetor de médias zero.

Os elementos da matriz de covariâncias de $\hat{c}'$, são:

$$\text{cov}(\hat{c}_\ell, \hat{c}_s) = \left(\frac{1}{2m}\right)^2 \sum_{j,h=-m+1}^m f_{\eta\eta}^{-1}(\lambda_j) f_{\eta\eta}^{-1}(\lambda_h) \text{cov}[\hat{f}_{X\eta}(\lambda_j) \hat{f}_{X\eta}(\lambda_h)] e^{i(\lambda_j \ell + \lambda_h s)}$$

Pela propriedade (2.7) e lembrando que $f_{X\eta}(\lambda) = 0$, temos:

$$\lim_{T \rightarrow \infty} T \text{cov}[\hat{f}_{X\eta}(\lambda_j) \hat{f}_{X\eta}(\lambda_h)] = 2\pi \int_{-\pi}^{\pi} f_{XX}(\omega) f_{\eta\eta}(\omega) H_M(\omega + \lambda_h) H_M(\omega - \lambda_j) d\omega,$$

e portanto:

$$\lim_{T \rightarrow \infty} T \text{ cov } (\hat{c}_\ell, \hat{c}_s) = 2\pi \int_{-\pi}^{\pi} f_{XX}(\omega) f_{\eta\eta}(\omega) \left[\left(\frac{1}{2m} \right)^2 \sum_{j,h=-m+1}^m f_{\eta\eta}^{-1}(\lambda_j) H_M(\omega - \lambda_j) \right. \\ \left. f_{\eta\eta}^{-1}(\lambda_h) \cdot H_M(\omega + \lambda_h) e^{i(\lambda_j \ell + \lambda_h s)} \right] d\omega.$$

Quando m cresce, temos:

$$\lim_{m \rightarrow \infty} \frac{1}{2m} \sum_{j=-m+1}^m f_{\eta\eta}^{-1}(\lambda_j) H_M(\omega - \lambda_j) e^{i\lambda_j \ell} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_{\eta\eta}^{-1}(\lambda) H_M(\omega - \lambda) e^{i\lambda \ell} d\lambda.$$

Quando M cresce, temos:

$$\lim_{M \rightarrow \infty} \int_{-\pi}^{\pi} f_{\eta\eta}^{-1}(\lambda) H_M(\omega - \lambda) e^{i\lambda \ell} d\lambda = f_{\eta\eta}^{-1}(\lambda) e^{i\omega \ell}.$$

Assim, os termos da matriz de covariâncias assintóticas de $\hat{c}'$ são:

$$\frac{1}{2\pi T} \int_{-\pi}^{\pi} f_{\eta\eta}^{-1}(\omega) f_{XX}(\omega) e^{i\omega(\ell-s)} d\omega. \quad (4.5)$$

Por outro lado, como os estimadores do espectro cruzado suavizado são assintoticamente normalmente distribuídos, $\hat{c}'$ tem distribuição aproximadamente normal.

Pelo que vimos, $\hat{c}'$ tem, assintoticamente, distribuição normal multivariada com vetor de médias zero e matriz de covariâncias $T^{-1}H$.

PROPRIEDADE 4.2 - Para T , m e M suficientemente grandes, $(\hat{a}' - a')$ tem distribuição aproximadamente normal, com vetor

de médias zero e matriz de covariâncias $T^{-1}H^{-1}$.

De (4.3) decorre que $(\hat{a}' - a') = \hat{c}'H^{-1}$. Como a distribuição limite de $\hat{c}'$ é normal multivariada com vetor de médias zero e matriz de covariâncias $T^{-1}H$, e como $\hat{H}$ converge em probabilidade para H , podemos concluir por um teorema devido a Slutsky (ver [8], pp.254-255) que $(\hat{a}' - a')$ tem a mesma distribuição limite que $\hat{c}'H^{-1}$. Assim $(\hat{a}' - a')$ tem distribuição aproximadamente normal multivariada, com vetor de médias zero e matriz de covariâncias $T^{-1}H^{-1}$.

Da propriedade 4.2 decorre que:

- i) a' é assintoticamente não viciado.
- ii) A matriz $T^{-1}H^{-1}$ de covariâncias assintóticas de $\hat{a}'$, é a matriz de covariâncias do melhor estimador linear não viciado (ver [14], cap.7).

Como estimador da matriz de covariâncias usaremos $T^{-1}\hat{H}^{-1}$.

Examinando (4.4), concluímos que a autocorrelação na sequência $X(t)$ causa uma inflação em H , se $\eta(t)$ for o ruído branco. Poderíamos supor que filtrando as sequências $X(t)$ e $Y(t)$ conseguiríamos diminuir os elementos da matriz de covariâncias. Porém, se o mesmo filtro é aplicado às duas séries, a razão $\frac{f_{XX}(\lambda)}{f_{\eta\eta}(\lambda)}$ não é afetada e consequentemente as covariâncias também.

PROPRIEDADE 4.3 - Se o quociente sinal/ruído é constante sobre o intervalo $(-\pi, \pi)$, os elementos de $\hat{a}'$ são, assintoticamente, independentes.

No caso em que $f_{XX}(\lambda) \cdot f_{\eta\eta}^{-1}(\lambda) = k$, $\forall \lambda \in (-\pi, \pi)$, os elementos da matriz de covariância de $\hat{a}'$ são dados por:

$$\begin{aligned} \text{cov}[\hat{a}(\ell), \hat{a}(s)] &= \left(\frac{1}{2m} \right)^2 \sum_{j, h=-m+1}^m f_{XX}^{-1}(\lambda_j) f_{XX}^{-1}(\lambda_h) \\ &\quad \times \text{cov}[\hat{f}_{X\eta}(\lambda_j) \hat{f}_{X\eta}(\lambda_h)] e^{i(\lambda_j \ell + \lambda_h s)}, \end{aligned}$$

e, como antes,

$$\begin{aligned} \lim_{M \rightarrow \infty} \lim_{m \rightarrow \infty} \lim_{T \rightarrow \infty} T \text{cov}[\hat{a}(\ell), \hat{a}(s)] &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f_{\eta\eta}(\omega) f_{XX}^{-1}(\omega) e^{i\omega(\ell-s)} d\omega = \\ &= \begin{cases} f_{\eta\eta}(\omega) f_{XX}^{-1}(\omega), & \text{se } \ell = s \\ 0, & \ell \neq s \end{cases} \quad (4.6) \end{aligned}$$

Portanto, os elementos de $\hat{a}'$ são assintoticamente independentes.

Sob a suposição (3.18), como estimador de $\text{cov}[\hat{a}(s), \hat{a}(\ell)]$, usaremos:

$$\frac{1}{2mT} \sum_{j=-m+1}^m \hat{f}_{XX}^{-1}(\lambda_j) \hat{f}_{\eta\eta}(\lambda_j) e^{i\lambda_j(\ell-s)}. \quad (4.7)$$

Sabemos que a matriz de covariâncias assintóticas

de $\hat{a}'$ é $T^{-1}H^{-1}$. Queremos saber o que acontece com os elementos desta matriz quando substituímos $f_{\eta\eta}(\lambda)$ pelo seu estimador consistente $\hat{f}_{\eta\eta}(\lambda)$.

Supondo que as flutuações amostrais são suficientemente pequenas para permitir o uso da expansão da série de Taylor truncada, quando formos determinar a variância da razão dos estimadores, e se T/M é suficientemente grande para que possamos ignorar contribuições de variância de ordem mais alta, temos

$$\text{var}[\hat{f}_{\eta\eta}^{-1}(\lambda)\hat{f}_{XX}(\lambda)] \sim \text{var}\left\{f_{\eta\eta}^{-1}(\lambda)f_{XX}(\lambda) + f_{\eta\eta}^{-1}(\lambda)[\hat{f}_{XX}(\lambda) - f_{XX}(\lambda)] - f_{\eta\eta}^{-2}(\lambda)f_{XX}(\lambda)[\hat{f}_{\eta\eta}(\lambda) - f_{\eta\eta}(\lambda)]\right\}.$$

Assim, para estimadores que são aproximadamente constantes sobre a largura de faixa temos:

$$\text{var}[\hat{f}_{\eta\eta}^{-1}(\lambda)\hat{f}_{XX}(\lambda)] \sim 2\delta_{\lambda}\psi_{T,M}f_{\eta\eta}^{-2}(\lambda)f_{XX}^2(\lambda),$$

pois

$$\text{var}[\hat{f}_{XX}(\lambda)] = \delta_{\lambda}\psi_{T,M}\hat{f}_{XX}(\lambda)$$

(δ_{λ} e $\psi_{T,M}$ estão definidos em (2.51)).

Como

$$\text{var}[f_{\eta\eta}^{-1}(\lambda)\hat{f}_{XX}(\lambda)] \sim \delta_\lambda \psi_{T,M} f_{\eta\eta}^{-2}(\lambda) f_{XX}^2(\lambda),$$

a substituição de $f_{\eta\eta}(\lambda)$ por $\hat{f}_{\eta\eta}(\lambda)$ não invalida o resultado básico, mas torna a convergência mais lenta.

No caso da razão sinal/ruído ser constante, a expressão (3.19) fornece o estimador apropriado. Como $f_{XX}(\lambda)$ é geralmente desconhecido, nós o substituiremos pelo seu estimador, e portanto:

$$\hat{a}(\ell) = \frac{1}{2m} \sum_{j=-m+1}^m \hat{f}_{XX}^{-1}(\lambda_j) \hat{f}_{XY}(\lambda_j) e^{i\lambda_j \ell}. \quad (4.8)$$

Usando a aproximação de Taylor, podemos mostrar que os elementos da matriz de covariâncias assintóticas são dados por:

$$\begin{cases} f_{\eta\eta}(\omega) f_{XX}^{-1}(\omega) + O(T^{-1}), & \text{se } \ell = s \\ O(T^{-1}) & , \text{ se } \ell \neq s. \end{cases}$$

4.2 - PROPRIEDADES DOS ESTIMADORES $\hat{\alpha}$ E $\hat{\beta}$ DA ESTRUTURA DE

DEFASAGENS GEOMÉTRICA

Vimos, no capítulo 3, que quando se especifica que

$$Y(t) = \alpha \sum_{s=-r_1}^{r_2} \beta^s X(t-s) + \eta(t),$$

os estimadores de máxima verossimilhança de α e β são obtidos através de (3.26).

Para o estudo das propriedades assintóticas de $\hat{\alpha}$ e $\hat{\beta}$, é conveniente supor $A(\lambda)$ conhecida.

Sabemos que:

$$\hat{H} \begin{pmatrix} \hat{\alpha} - \alpha \\ \hat{\beta} - \beta \end{pmatrix} = f_{ZZ}^{-1}(\lambda_j) \begin{bmatrix} \hat{\alpha} \hat{f}_{XX}(\lambda_j) - \alpha \hat{f}_{XX}(\lambda_j) + \hat{\beta} \hat{f}_{XY}(\lambda_j) e^{-i\lambda_j} - \beta \hat{f}_{XY}(\lambda_j) e^{-i\lambda_j} \\ A^*(\lambda_j) [\hat{f}_{XY}(\lambda_j) e^{i\lambda_j} \hat{\alpha} - \hat{f}_{XX}(\lambda_j) e^{i\lambda_j} \alpha + \hat{f}_{XY}(\lambda_j) \hat{\beta} - \hat{f}_{XY}(\lambda_j) \beta] \end{bmatrix},$$

onde

$$f_{ZZ}(\lambda_j) = \frac{f_{\eta\eta}(\lambda_j)}{|A(\lambda_j)|^2}.$$

Substituindo

$$\hat{\alpha} \hat{f}_{XX}(\lambda_j) = \hat{f}_{XY}(\lambda_j) (1 - \hat{\beta} e^{-i\lambda_j}),$$

temos:

$$\hat{H} \begin{pmatrix} \hat{\alpha} - \alpha \\ \hat{\beta} - \beta \end{pmatrix} = f_{ZZ}^{-1}(\lambda_j) \begin{bmatrix} \hat{f}_{XY}(\lambda_j) - \alpha \hat{f}_{XX}(\lambda_j) - \beta \hat{f}_{XY}(\lambda_j) e^{-i\lambda_j} \\ A^*(\lambda_j) e^{i\lambda_j} [\hat{f}_{XY}(\lambda_j) (1 - \hat{\beta} e^{-i\lambda_j}) + \hat{f}_{XY}(\lambda_j) \hat{\beta} e^{-i\lambda_j} - \alpha \hat{f}_{XX}(\lambda_j) - \beta \hat{f}_{XY}(\lambda_j) e^{-i\lambda_j}] \end{bmatrix}$$

Lembrando que aproximadamente

$$\hat{f}_{XY}(\lambda_j) = \alpha \hat{f}_{XX}(\lambda_j) + \hat{\beta} \hat{f}_{XY}(\lambda_j) e^{-i\lambda_j} + \hat{f}_{X\eta}(\lambda_j) A^{-1}(\lambda_j), \quad (4.9)$$

chega-se a

$$\hat{H} \begin{bmatrix} \hat{\alpha} - \alpha \\ \hat{\beta} - \beta \end{bmatrix} = \hat{c}, \quad (4.10)$$

onde

$$\hat{c} = \frac{1}{2m} \sum_{j=-m+1}^m f_{ZZ}^{-1}(\lambda_j) \begin{bmatrix} \hat{f}_{XZ}(\lambda_j) \\ A^*(\lambda_j) \hat{f}_{XZ}(\lambda_j) e^{i\lambda_j} \end{bmatrix},$$

$$\hat{f}_{XZ}(\lambda_j) = \frac{\hat{f}_{X\eta}(\lambda_j)}{A(\lambda_j)}.$$

Fazendo uso de (4.9) podemos mostrar que a matriz $\hat{H}$ pode ser escrita como:

$$\hat{H} = \frac{1}{2m} \sum_{j=-m+1}^m f_{ZZ}^{-1}(\lambda_j) \begin{bmatrix} \hat{f}_{XX}(\lambda_j) & [\alpha A(\lambda_j) \hat{f}_{XX}(\lambda_j) + \hat{f}_{X\eta}(\lambda_j)] e^{-i\lambda_j} \\ A^*(\lambda_j) \hat{f}_{XX}(\lambda_j) e^{i\lambda_j} & A^*(\lambda_j) [\alpha A(\lambda_j) \hat{f}_{XX}(\lambda_j) + \hat{f}_{X\eta}(\lambda_j)] \end{bmatrix},$$

e de modo análogo ao descrito na secção 4.1, mostrar que esta matriz converge em probabilidade para

$$\begin{aligned} H &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f_{ZZ}^{-1}(\lambda) \begin{bmatrix} f_{XX}(\lambda) & \alpha f_{XX}(\lambda) A(\lambda) e^{-i\lambda} \\ f_{XX}(\lambda) A^*(\lambda) e^{i\lambda} & \alpha f_{XX}(\lambda) |A(\lambda)|^2 \end{bmatrix} d\lambda = \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f_{ZZ}^{-1}(\lambda) f_{XX}(\lambda) U(\lambda) \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix} d\lambda, \end{aligned} \quad (4.11)$$

onde

$$U(\lambda) = \begin{bmatrix} 1 & A(\lambda)e^{-i\lambda} \\ A^*(\lambda)e^{i\lambda} & |A(\lambda)|^2 \end{bmatrix}.$$

Fazendo uso deste fato e com raciocínio análogo ao da secção anterior, e supondo aqui também que $X(t)$ é normal, pode-se mostrar que:

PROPRIEDADE 4.4 - Para T, M e m suficientemente grandes,

$$\begin{pmatrix} \hat{\alpha} - \alpha \\ \hat{\beta} - \beta \end{pmatrix}$$

tem distribuição aproximadamente normal, com vetor de médias zero e matriz de covariâncias

$$T^{-1} \begin{bmatrix} 1 & 0 \\ 0 & \alpha^{-1} \end{bmatrix} H^{-1}.$$

O resultado acima ainda é válido quando $f_{ZZ}(\lambda)$ e $A(\lambda)$ são substituídos em $\hat{H}$ pelos seus estimadores consistentes.

4.3 - QUANDO O PROCESSO RESIDUAL É O RUÍDO BRANCO

Neste caso, como

$$f_{\eta\eta}(\lambda) = \frac{\sigma^2}{2\pi} = \frac{R_{\eta\eta}(0)}{2\pi},$$

de (3.11) segue que:

$$\frac{1}{2m} \sum_{s=-r_1}^{r_2} \hat{a}(s) \sum_{j=-m+1}^m \hat{f}_{XX}(\lambda_j) e^{i\lambda_j(\ell-s)} = \frac{1}{2m} \sum_{j=-m+1}^m \hat{f}_{XY}(\lambda_j) e^{i\lambda_j \ell} \quad (4.12)$$

Substituindo $\hat{f}_{XX}(\lambda_j)$ e $\hat{f}_{XY}(\lambda_j)$ por suas expressões, temos

$$\begin{aligned} \frac{1}{2\pi} \sum_{s=-r_1}^{r_2} \hat{a}(s) \sum_{u=-M}^M h_M(u) \hat{R}_{XX}(u) \left[\frac{1}{2m} \sum_{j=-m+1}^m e^{i\lambda_j(\ell-s-u)} \right] = \\ = \frac{1}{2\pi} \sum_{u=-M}^M h_M(u) \hat{R}_{XY}(u) \left[\frac{1}{2m} \sum_{j=-m+1}^m e^{i\lambda_j(\ell-u)} \right]. \end{aligned}$$

Assim,

$$\sum_{s=-r_1}^{r_2} \hat{a}(s) \hat{R}_{XX}(\ell-s) h_M(\ell-s) = R_{XY}(\ell) h_M(\ell).$$

Se M é suficientemente grande, a função peso é próxima da unidade, e portanto:

$$\sum_{s=-r_1}^{r_2} \hat{a}(s) \hat{R}_{XX}(\ell-s) \sim \hat{R}_{XY}(\ell).$$

A inversa da matriz de covariâncias assintóticas de $T^{1/2} \hat{\mathbf{a}}$, tem elementos dados por:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f_{\eta\eta}^{-1}(\omega) f_{XX}(\omega) e^{i\omega(\ell-s)} d\omega = \frac{R_{XX}(\ell-s)}{R_{\eta\eta}(0)}.$$

Vemos portanto, que se o espectro residual é uniforme, o problema de estimar $f_{\eta\eta}(\lambda)$ desaparece. É pois inte

ressante, antes de estimar $\hat{a}'$, inspecionar

$$\hat{f}_{\eta\eta}(\lambda) = \hat{f}_{YY}(\lambda)[1 - \hat{\gamma}_{XY}^2(\lambda)].$$

Se $\hat{f}_{\eta\eta}(\lambda)$ é relativamente constante, podemos assumir $f_{\eta\eta}(\lambda)$ como sendo uniforme. Em alguns casos, é possível filtrar $Y(t)$ de tal forma que o espectro residual seja constante. Se o mesmo filtro é aplicado a $X(t)$, então podemos usar o estimador de mínimos quadrados nas duas séries para estimar o modelo de defasagens distribuídas original.

4.4 - A VALIDADE DO MODELO

Para testar a validade do modelo com os dados, devemos testar a hipótese de que a função de coerência $\gamma_{X\eta}(\lambda)$ é zero. O procedimento do teste foi descrito no capítulo 2, secção 2.11.

A rigor, a série temporal deveria ser dividida em duas partes: uma para a estimação dos parâmetros, e a outra para a validação do modelo. Como, em geral, as séries econômicas são pequenas, quase sempre usamos os mesmos dados para as duas coisas.

A coerência amostral é calculada usando a série temporal $X(t)$ e a série residual calculada por:

$$\hat{\eta}(t) = Y(t) - \hat{\mu} - \sum_{s=-r_1}^{r_2} \hat{a}(s)X(t-s).$$

O número de observações fica aqui reduzido de T para $T - r_1 - r_2$.

No caso da estrutura de defasagens geométrica, considera-se

$$\varepsilon(t) = \eta(t) - \beta\eta(t-1),$$

e testa-se a hipótese de que $\gamma_{X\eta}(\lambda)$ é zero. Como $\varepsilon(t)$ é uma transformação de $\eta(t)$ produzida por um filtro que preserva a associação linear de $\eta(t)$ com $X(t)$, se alguma existe, não ocorre nenhuma perda de generalidade. A série residual é calculada por

$$\varepsilon(t) = Y(t) - \hat{\alpha}X(t) - \hat{\beta}Y(t-1) - \hat{\mu}(1-\hat{\beta}),$$

e o número de observações neste caso é $T-1$.

Outros enfoques para testar a validade, não tão simples quanto o aqui apresentado, foram sugeridos por Hannan [22], pp.220-223, Hamon e Hannan [20] e Wallis (ver [13] p. 166).

4.5 - A EXTENSÃO A MODELOS MULTIVARIADOS

Em certos tipos de comportamento, a variável endógena não pode ser explicada pelos valores tomados por apenas uma variável exógena. Assim, o modelo bivariado dado em

(3.1), usado nestes casos, conduz a resultados errôneos (ver [13], p.74). Considera-se então a especificação:

$$Y(t) = \sum_{s=-r_1}^{r_2} a'(s)X(t-s) + \eta(t), \quad (4.13)$$

onde $X(t)$ é um vetor ($q \times 1$) de variáveis exógenas, $a'(s)$ é um vetor ($1 \times q$) de coeficientes.

Em geral, um comportamento econômico não pode ser explicado por apenas uma equação, como em (4.13), mas sim por um sistema de equações. Deste modo, temos o conjunto de p equações

$$Y(t) = \sum_{s=-r_1}^{r_2} B(s)X(t-s) + \eta(t), \quad (4.14)$$

onde $Y(t)$ é um vetor ($p \times 1$) de variáveis endógenas, $X(t)$ é um vetor ($q \times 1$) de variáveis exógenas, $\eta(t)$ é um vetor ($p \times 1$) de resíduos, e $B(s)$ é a matriz ($p \times q$) de coeficientes do modelo.

Supondo que todas as componentes em (4.14) são normais e fracamente estacionárias, e ainda que

$$E[\eta(t)] = 0, E[X(t)\eta(s)] = 0, \quad t, s=0, \pm 1, \dots, \pm \infty,$$

Fishman [13], pp.172-180 estimou a matriz de coeficientes B e estudou as propriedades assintóticas do estimador, utilizando as técnicas de análise espectral. O procedimento por

- 96 -

ele utilizado não é mais que uma extensão daquele aqui apre
sentado no caso do modelo bivariado.

- I -

CAPÍTULO 5

CONSIDERAÇÕES FINAIS

5.1 - MODELOS DE DEFASAGENS DISTRIBUÍDAS MAIS GERAIS

Do estudo do modelo

$$Y(t) = \sum_{s=-r_1}^{r_2} a(s)X(t-s) + \eta(t),$$

desenvolvido nos capítulos 3 e 4, algumas conclusões podem ser tiradas:

- a) Sob a suposição de que $\eta(t)$, $X(t)$ e $Y(t)$ são sequências estocásticas, estacionárias de 2ª ordem, de média zero, e de que a sequência $\eta(t)$ é normal e independente de $X(t)$, podemos determinar o estimador de máxima verossimilhança de a' (vetor de coeficientes do modelo), e verificar que este estimador é assintoticamente não viciado com matriz de covariâncias $T^{-1}H^{-1}$;
- b) Se às suposições dadas em a), adicionarmos a hipóte-

se de que $X(t)$ é normal, podemos verificar que a distribuição assintótica de $\hat{a}'$ (estimador de máxima verossimilhança do vetor a') é normal multivariada.

Uma sugestão para um possível estudo seria então o da determinação da distribuição assintótica de $\hat{a}'$ sob condições menos restritas a respeito da sequência $X(t)$. Poder-se-ia, por exemplo, focar o problema sob a suposição de que esta sequência seja estacionária de 2ª ordem, mas não necessariamente normal.

Tendo em vista a utilização da análise espectral no estudo de modelos mais gerais, é interessante citarmos aqui alguns desenvolvimentos neste sentido, devidos a Hannan.

Em [22] ele apresenta os estimadores de máxima verossimilhança bem como as propriedades destes estimadores no caso do modelo geométrico de defasagens, impondo a $X(t)$ as condições

$$\begin{aligned} p \lim_{n \rightarrow \infty} d_n^2 &= p \lim_{n \rightarrow \infty} \sum_{t=1}^n X^2(t) = \infty, \\ p \lim_{n \rightarrow \infty} \frac{\max_{-n \leq j \leq n} X_j^2}{d_n^2} &= 0, \\ p \lim_{n \rightarrow \infty} \frac{\sum_{t=1}^n X(t)X(t+h)}{d_n^2} &= \rho(h), \end{aligned} \tag{5.1}$$

que são menos restritivas que a asserção de estacionaridade de 2ª ordem para $X(t)$. Conclui, então, que

$$d_n \left[\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} - \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \right]$$

é assintoticamente normal com vetor de médias zero e matriz de covariâncias

$$2\pi \int_{-\pi}^{\pi} f_{ZZ}^{-1}(\omega) u(\theta) u^*(\theta) dF_{XX}(\theta),$$

onde

$$u(\theta) = \begin{bmatrix} 1 \\ \frac{2e^{-i\theta}}{1-\beta e^{-i\theta}} \end{bmatrix},$$

e $u^*(\theta)$ é a transposta conjugada de $u(\theta)$.

Em [23] Hannan considera o modelo

$$Y(t) = \sum_{s=-r_1}^{r_2} B(s)X(t-s) + \eta(t),$$

onde $Y(t)$ e $\eta(t)$ são vetores de p componentes, $X(t)$ é vetor de q componentes, e portanto $B(s)$ é matriz $p \times q$, e discute uma teoria para estimar $B(s)$ baseada na análise espectral. Com a suposição de que os resíduos $\eta(t)$ são gerados por um processo linear independente de $X(t)$, e sob algumas condições adicionais de regularidade, Hannan mostra que, assinto

ticamente, a distribuição conjunta dos estimadores é normal multivariada. O método de estimação não é eficiente, a menos que os espectros dos processos $X(t)$ e $\eta(t)$ sejam os mesmos. Por eficiência, entende-se que a distribuição conjunta dos estimadores seja assintoticamente normal, com matriz de covariâncias que é aquela do melhor estimador linear não viado.

5.2 - COMPARAÇÃO ENTRE ESTIMADORES OBTIDOS NO DOMÍNIO DO TEMPO E ESTIMADORES OBTIDOS NO DOMÍNIO DA FREQUÊNCIA

A maioria dos estudos comparativos entre os vários estimadores, aparecem em conexão com a forma de defasagens geométrica, dada a sua importância. Assim, no caso particular em que o processo residual $\eta(t)$ é normal, Amemiya e Fuller [2] mostraram que os estimadores de máxima verossimilhança (no domínio do tempo) e os estimadores de Hannan (no domínio da frequência) possuem a mesma distribuição assintótica. Mostraram também que os estimadores de Hannan são assintoticamente mais eficientes que os de Klein. O próprio Hannan em [23] verificou que os seus estimadores de máxima verossimilhança para α e β são assintoticamente mais eficientes que os estimadores de Leviatan. Para a definição de eficiência assintótica relativa ver [11].

Meyer e Cargill em [30] apresentam um estudo comparativo, resultante de experimentos de simulação, de três estimadores: o de Almon, o estimador de mínimos quadrados usual (OLS) e o estimador ineficiente de Hannan (citado em 5.1). Com o objetivo de atenuar algumas das desvantagens de um experimento de simulação, cada método de estimação foi comparado para amostras de tamanho 50 e 100, usando-se várias formas da distribuição de defasagens, e várias estruturas para o processo residual. Os resultados indicam um melhor desempenho (avaliado por cinco critérios) do estimador OLS sobre os estimadores de Hannan e Almon, para qualquer tamanho de amostra e especificação do modelo, e um melhor desempenho do estimador de Hannan sobre o de Almon. Quando o número de lags é especificado incorretamente, existe uma tendência de aumento dos vieses no caso do estimador de Almon e no OLS, enquanto que o de Hannan não sofre influência. O estimador OLS, mesmo em presença de correlação serial marcante nos resíduos, apresenta uma melhor descrição da distribuição de defasagens do que o de Almon e o de Hannan.

5.3 - A ESTIMAÇÃO DO VETOR a' FAZENDO USO DO PERIODOGRAMA ALISADO

Na secção 3.3 do capítulo 3, quando procurávamos estimar o vetor a' , chegamos às equações (3.11), onde apare-

cem os periodogramas $I_{XX}(\lambda)$ e $I_{XY}(\lambda)$. Pelo fato do periodograma não ser um estimador consistente do espectro, em (3.11) nós substituímos $I_{XX}(\lambda)$ e $I_{XY}(\lambda)$ por $\hat{f}_{XX}(\lambda)$ e $\hat{f}_{XY}(\lambda)$, que são os estimadores suavizados do espectro, obtidos, respectivamente, através de uma média ponderada da função de covariância amostral de $X(t)$ e de uma média ponderada da função de covariância amostral cruzada de $X(t)$ e $Y(t)$. É claro que outros estimadores, que gozem da propriedade de consistência, podem ser obtidos e usados para substituir $I_{XX}(\lambda)$ e $I_{XY}(\lambda)$.

Brillinger [6] propôs para a estimação do espectro, o uso de uma média aritmética simples de ordenadas do periodograma. Supondo que $s(T)$ é um inteiro tal que $2\pi \frac{s(T)}{T}$ está próximo de λ , ele mostrou que as $(2M+1)$ ordenadas adjacentes do periodograma $I_{XX}\left[\frac{2\pi(s(T)+j)}{T}\right]$, $j=0, \pm 1, \dots, \pm M$, são variáveis aleatórias assintoticamente independentes, cada uma com distribuição limite

$$f_{XX}(\lambda) \frac{\chi^2_2}{2}, \text{ se } 2[s(T)+j] \not\equiv 0 \pmod{T},$$

$$j=0, \pm 1, \dots, \pm M,$$

Isto lhe sugeriu o uso de:

$$\tilde{I}_{XX}(\lambda) = \begin{cases} \frac{1}{2M+1} \sum_{j=-M}^M I_{XX}\left(\frac{2\pi[s(T)+j]}{T}\right), & \text{se } \lambda \not\equiv 0 \pmod{\pi} \\ \frac{1}{M} \sum_{j=1}^M I_{XX}\left(\lambda + \frac{2\pi j}{T}\right), & \text{se } \lambda \equiv 0 \pmod{\pi} \text{ e } T \text{ é par} \\ \frac{1}{M} \sum_{j=1}^M I_{XX}\left(\lambda - \frac{\pi}{T} + \frac{2\pi j}{T}\right), & \text{se } \lambda \equiv 0 \pmod{\pi} \text{ e } T \text{ é ímpar} \end{cases} \quad (5.2)$$

para estimar $f_{XX}(\lambda)$, denominado "periodograma alisado" de $X(t)$.

Se $X(t)$ é uma sequência estacionária normal, pode-se mostrar que com $\lambda - \frac{2\pi s(T)}{T} = O(T^{-1})$, o estimador dado em (5.2) é assintoticamente não viciado e que

$$\text{Var}[\tilde{I}_{XX}(\lambda)] = \begin{cases} \frac{f_{XX}^2(\lambda)}{2M+1} + O(T^{-1}), & \text{se } \lambda \not\equiv 0 \pmod{\pi} \\ \frac{f_{XX}^2(\lambda)}{M} + O(T^{-1}), & \text{se } \lambda \equiv 0 \pmod{\pi} \end{cases}$$

Portanto, a variância decresce quando M cresce. Porém, aqui também, um aumento de M acarreta um aumento no viés do estimador ([6]p.414), de tal sorte que um valor de compromisso para M deve ser escolhido.

Para estimar $f_{XY}(\lambda)$, Brillinger tomou:

$$\tilde{I}_{XY}(\lambda) = \begin{cases} \frac{1}{2M+1} \sum_{j=-M}^M I_{XY}\left(\frac{2\pi(s(T)+j)}{T}\right), & \text{se } \lambda \not\equiv 0 \pmod{\pi} \\ \frac{1}{M} \sum_{j=1}^M \operatorname{Re} I_{XY}\left(\lambda + \frac{2\pi j}{T}\right), & \text{se } \lambda \equiv 0 \pmod{\pi} \text{ e } T \text{ é par} \\ \frac{1}{M} \sum_{j=1}^M \operatorname{Re} I_{XY}\left(\lambda - \frac{\pi}{T} + \frac{2\pi j}{T}\right), & \text{se } \lambda \equiv 0 \pmod{\pi} \text{ e } T \text{ é ímpar,} \end{cases} \quad (5.3)$$

sendo que $\frac{2\pi s(T)}{T} \rightarrow \lambda$ quando $T \rightarrow \infty$. $I_{XY}(\lambda)$ é também assintoticamente não viciado, e sua variância decresce quando M cresce.

Assim, um problema a ser estudado é o da determinação das propriedades e da distribuição assintótica de $\hat{a}$, quando $\tilde{I}_{XX}(\lambda)$ e $\tilde{I}_{XY}(\lambda)$ são usados em (3.11) no lugar de $I_{XX}(\lambda)$ e $I_{XY}(\lambda)$, respectivamente.

5.4 - O PROBLEMA DA COMPUTAÇÃO

O que nos propusemos neste trabalho, foi apresentar um desenvolvimento teórico à respeito do modelo de defasagens distribuídas dado em (3.1). No entanto, não podemos nos esquecer que todo desenvolvimento teórico nesta área visa principalmente a aplicação prática.

Tendo em vista o vasto campo de aplicação deste modelo, principalmente em Econometria, e lembrando também que o estimador aqui apresentado apresenta um comportamento as-

sintótico dos mais desejáveis, seria de interesse que um programa de computação fosse feito visando estas aplicações, pois não é do nosso conhecimento a existência de algum.

(B)

BIBLIOGRAFIA

- [1] - Almon, S., "The Distributed Lag Between Capital Appropriation and Expenditures", *Econometrica*, vol. 33, N° 1, 1965, pp. 178-196.
- [2] - Amemiya, T. e W. Fuller, "A Comparative Study of Alternative Estimators in A Distributed Lag Model", *Econometrica*, vol. 35, 1967, pp. 509-529.
- [3] - Anderson, T.W., *An Introduction to Multivariate Statistical Analysis*, John Wiley and Sons, New York, 1958.
- [4] - Bartlett, M.S., "On the Theoretical Specification and Sampling Properties of Autocorrelated Time Series", *Journal of the Royal Statistical Society, Serie B*, Vol. 8, 1946, pp. 27-41.
- [5] - ———, *An Introduction to Stochastic Processes*, Cambridge University Press, London, 1961.
- [6] - Brillinger, D.R. *Time Series: Data Analysis and Theory*, Holt Rinehart and Winston, Inc., 1975.
- [7] - Cagan, P., "The Monetary Dynamics of Hyper Inflation", in M. Friedman ed. *Studies in the Quantity Theory of Money*, University of Chicago Press, 1956.
- [8] - Cramér, H., *Mathematical Methods of Statistics*, Princeton University Press, Princeton, New Jersey, 1946.
- [9] - Cooley, J.M. e Tukey, J.M., "An Algorithm for the Machine Calculation of Fourier Series", *Mathematics of Computation*, Vol. 19, 1965, pp. 297-301.
- [10] - Dhrymes, P.J., "Efficient Estimation of Distributed Lags with Autocorrelated Errors", *International Economic Review*, vol. 10, N° 1, 1969, pp. 47-67

- [11] - Dhrymes, P.J., *Distributed Lags. Problems of Estimation and Formulation*, San Francisco, Holden-Day, 1971.
- [12] - Doob, J.L., *Stochastic Processes*, John Wiley and Sons, New York, 1953.
- [13] - Fishman, G.S., *Spectral Methods in Econometrics*, Harvard University Press, Cambridge, 1969.
- [14] - Grenander, U. e M. Roseblatt, *Statistical Analysis of Stationary Time Series*, John Wiley and Sons, New York, 1957.
- [15] - Griliches, Z., "Distributed Lags: A Survey", *Econometrica*, vol. 35 (1), 1967, pp. 16-49.
- [16] - Goodman, N.R., "On the Joint Estimation of the Spectra, Coespectrum and Quadrature Spectrum of a Two-Dimensional Stationary Gaussian Process", Ph.D. Thesis, Princeton University, 1957.
- [17] - ———, "The Distribution of the Determinant of a Complex Wishart Distributed Matrix", *Annals of Mathematical Statistics*, vol. 34, n° 1, 1963, pp. 178-185.
- [18] - ———, "Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution" (An Introduction), *Annals of Mathematical Statistics*, vol. 34, n° 1, 1963, pp. 152-177.
- [19] - Gupta, Y.P., "Least Squares Variant of Dhrymes Two-Step Estimation Procedure of the Distributed Lag Model", *International Economic Review*, vol. 10, N° 1, 1969, pp. 112-117.
- [20] - Hamon, B.V., e E.J. Hannan, "Estimating Relations Between Time Series", *J. Geophysical Res.*, vol. 68, n° 21, 1963.
- [21] - Hannan, E.J., "Regression for Time Series", in *Proceedings of the Symposium on Time Series Analysis*, Brown University, 1962, M. Roseblatt (ed), John Wiley and Sons, New York, 1963.
- [22] - ———, "The Estimation Relationships Involving Distributed Lags", *Econometrica*, vol. 33, n° 1, 1965, pp. 206-224.
- [23] - ———, "The Estimation of a Lagged Regression Relation", *Biometrika*, vol. 54, N° 3 e 4, 1967, pp. 409-418.
- [24] - Jenkins, G.M., "Cross-Spectral Analysis and Estimation of Open Loop

- Transfer Functions", in *Proceedings of the Symposium on Time Series Analysis*, Brown University, 1962, M. Robenblatt (ed), John Wiley and Sons, New York, 1963, pp.267-276.
- [25] - Jenkins, G.M. and Watts, D.G., *Spectral Analysis and Its Applications*, Holden Day Inc., 1968.
- [26] - Klein, L. "The Estimation of Distributed Lags" *Econometrica*, vol. 26, Nº 4, 1958, pp.553-565.
- [27] - Koyck, L.M., *Distributed Lags and Investment Analysis*, Amsterdam, North-Holland Publishing Co., 1954.
- [28] - Leviatan, N., "Consistent Estimation of Distributed Lags", *International Economic Review*, vol. 4, Nº 1, 1963, pp.44-52.
- [29] - Malinvaud, E., *Statistical Methods of Econometrics*, Rand Mc Nally, Chicago, 1966.
- [30] - Meyer, R.A. e T.F. Cargill, "Some Time and Frequency Domain Distributed Lag Estimators: A Comparative Monte Carlo Study", *Econometrica*, vol. 42, Nº 6, 1974, pp.1031-1044.
- [31] - Nerlove, M. "Estimates of the Elasticities of Supply of Selected Agricultural Commodities", *Journal of Farm Economics*, vol. 38, 1956, pp.496-509.
- [32] - ———, *The Dynamics of Supply: Estimation of Farmer's Response to Price*, John Hopkins Press, Baltimore, 1958.
- [33] - Parzen, E., "Mathematical Considerations in the Estimation of Spectra", *Technometrics*, vol. 3, Nº 2, 1961, pp.167-190.
- [34] - Pastore, A.C. "A resposta da Produção Agrícola aos Preços no Brasil, Tese de Doutorado, USP, 1968.
- [35] - ———, "Aspectos da Política Monetária Recente no Brasil", *Revista de Estudos Econômicos*, IPE-USP, Vol. 3, Nº 3, 1973, pp.7 a 58.
- [36] - Solow, R.M., "On a Family of Lag Distribution", *Econometrica*, vol. 28, Nº 2, 1960, pp.393-406.
- [37] - Yaglom, A.M., *An Introduction to the Theory of Stationary Random Functions*, Prentice-Hall, Inc., Englewood Cliffs, N. Jersey, 1962.