AJUSTAMENTO SAZONAL

DE SÉRIES TEMPORAIS

NAZIRA GAIT

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

ORIENTADOR:

Prof.Dr. PEDRO ALBERTO MORETTIN

- SÃO PAULO, março de 1975 -

NOTA INICIAL

Este trabalho não poderia ter sido realizado sem o apoio e amizade dos colegas do Departamento de Estatística, do Instituto de Matemática e Estatística da Universidade de São Paulo, entre os quais devo ressaltar:

Prof. Dr. Lindo Fava, Chefe do Departamento, que tornou possível minha contratação neste Instituto.

Prof. Dr. Pedro Alberto Morettin, que tomou a seu encargo a difícil tarefa de orientação do trabalho e a realizou com paciência e bom humor.

Sr. João Baptista Esteves de Oliveira, secretário do Departamento, que realizou um excelente trabalho de datilografia.

Dedico este trabalho a meus pais, por todo o amor de que são fonte inesgotável.

I N D I C E

CAPÍTULO I - INTRODUÇÃO	1
CAPÍTULO II - TEORIA ESPECTRAL.	12
2.1 - Introdução	12
2.2 - Descrição da ST no domínio do tempo. . .	13
2.3 - Descrição da ST no domínio de frequência	16
2.4 - Exemplos de Espectro e Auto-Covariância.	24
2.5 - Filtros.	27
2.6 - O Espectro-Cruzado	34
2.7 - Distribuição Amostral dos Estimadores. .	36
CAPÍTULO III - ELIMINAÇÃO DA TENDÊNCIA.	40
3.1 - Considerações Gerais	40
3.2 - Testes para verificar a existência de tendência na série	41
3.3 - Métodos de Eliminação da tendência. . .	45
CAPÍTULO IV - ELIMINAÇÃO DO COMPONENTE SAZONAL. . .	59
4.1 - Introdução	59
4.2 - Método 1 (Kendall)	62
4.3 - Método 2 (Jorgenson)	63
4.4 - Método 3 (Kendall)	71
4.5 - Método 4 (Bureau do Censo)	73
4.6 - Método 5 (BLS)	74
4.7 - Método 6 (Residual).	75
CAPÍTULO V - ANÁLISE DOS EFEITOS DOS PROCESSOS DE AJUSTAMENTO SAZONAL.	76
CAPÍTULO VI - APLICAÇÕES.	87
APÊNDICE 1 - ALGUMAS SÉRIES TEMPORAIS E SEUS GRÁFI- COS.	94

APÊNDICE 2 - TABELAS E GRÁFICOS REFERENTES AO CAPÍTULO VI.	101
BIBLIOGRAFIA.	109

CAPÍTULO I

INTRODUÇÃO

Na natureza ocorrem muitos fenômenos que variam com o tempo e, desde épocas remotas, o homem tem procurado analisá-los. Qualquer coleção de dados que represente algum aspecto destes fenômenos pode ser considerada uma série temporal, mas o estudo sistemático de séries temporais é relativamente recente.

Segundo Kendall (1973), deve datar de aproximadamente 1.000 anos o registro de eventos de acordo com um eixo horizontal, onde intervalos iguais correspondem a iguais espaços de tempo. Possivelmente, um dos primeiros trabalhos onde aparecem gráficos de tempo na forma moderna, tenha sido feito por William Playfair em 1821, sobre a situação crítica da agricultura nos Estados Unidos.

No século XIX, a Estatística teórica ainda não era unificada e os trabalhos referentes às ciências físicas eram independentes das das ciências humanas e biológicas. Nesta época, as idéias dos físicos eram totalmente determinísticas;

isto é, um fenômeno que ocorria através do tempo era imaginado como se se comportasse segundo leis determinísticas.

Nos fins do século XIX, tentou-se aplicar às ciências humanas e biológicas os mesmos métodos das ciências físicas; e a moderna teoria estatística começou a se desenvolver baseada na idéia de que, embora um indivíduo isoladamente não se comportasse de modo determinístico, os agregados de indivíduos eram sujeitos a leis que poderiam ser apresentadas em termos matemáticos simples. E então começou a ser usado todo o aparato matemático que foi desenvolvido mais tarde.

Até 1925, aproximadamente, as irregularidades que ocorriam em uma série temporal eram consideradas "erros", análogos aos erros de observação. Elas aconteciam com freqüência, mas tinham a mesma conotação que tem, para os engenheiros eletrônicos, o "ruído": distúrbios fortuitos que ocorrem de maneira sistemática. Em particular, fenômenos flutuantes tais como os "ciclos de comércio" do século XIX, eram submetidos à Análise de Fourier como se consistissem num conjunto de termos harmônicos gerados por oscilações do tipo clássico.

Foi em 1927 que Udny Yule apresentou nova forma de estudo das séries temporais, ao observar o número de explosões solares. Ele verificou que eram irregulares as amplitudes da série e também as distâncias entre os picos e as depressões. A partir de então, começou a se desenvolver, com técnicas próprias, a análise de séries temporais.

DEFINIÇÃO 1.1 - Seja $\{Y(t, \omega), t \in T, \omega \in \Omega\}$ um Processo Estocástico (ou função aleatória), definido num espaço de probabilidade $S(\Omega, A, p)$.

Para cada t fixado, $Y(t, \omega)$ é uma variável aleatória.

Para ω fixo, $Y(t, \omega)$ é uma série temporal; isto é, a série temporal é definida como uma *realização* particular de um processo estocástico.

Vamos denotar uma série temporal por $\{Y(t), t \in T\}$.

O parâmetro t em geral é tempo, mas pode representar superfície, volume, etc.

Por exemplo, $Y(\underline{t})$ pode representar a temperatura do ar em um ponto fixo, onde $\underline{t} = (\text{latitude}, \text{longitude})$.

T em geral, é tomado como $\mathbb{Z}$ ou $\mathbb{R}$.

Em particular, seja $T = \mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$.

Temos:

$$(1.1) \quad \underline{Y}(t) = \begin{bmatrix} Y_1(t) \\ \vdots \\ Y_r(t) \end{bmatrix}$$

é um processo vetorial ou multivariado;

$$(1.2) \quad Y(\underline{t}) \text{ onde } \underline{t} = (t_1, \dots, t_n)$$

é um processo multidimensional;

$$(1.3) \quad \underline{Y}(\underline{t}) = \begin{bmatrix} Y_1(\underline{t}) \\ \vdots \\ Y_r(\underline{t}) \end{bmatrix} \text{ onde } \underline{t} = (t_1, \dots, t_n)$$

é um processo multivariado e multidimensional.

A função aleatória $\{Y(t, \omega)\}$ fica especificada se conhecermos todas as distribuições finitas, isto é:

$$(1.4) \quad F_t(y) = P[Y(t) \leq y] \quad \text{de } Y(t),$$

$$(1.5) \quad F_{t_1, t_2}(y_1, y_2) = P[Y(t_1) \leq y_1; Y(t_2) \leq y_2] \quad \text{de } Y(t_1, t_2)$$

$$(1.6) \quad F_{t_1, \dots, t_n}(y_1, \dots, y_n) = P[Y(t_1) \leq y_1, \dots, Y(t_n) \leq y_n] \quad \text{de } Y(t_1, \dots, t_n).$$

DEFINIÇÃO 1.2 - $\underline{Y}(t)$, de ordem $r \times 1$, é um processo *estritamente estacionário* se, para $t_1, \dots, t_n, k \in \mathbb{Z}$,

$$F_{t_1+k, \dots, t_n+k}(y_1, \dots, y_n) = F_{t_1, \dots, t_n}(y_1, \dots, y_n).$$

Isto é, se as variáveis aleatórias

$$(\underline{Y}(t_1), \dots, \underline{Y}(t_n)) \quad \text{e} \quad (\underline{Y}(t_1+k), \dots, \underline{Y}(t_n+k))$$

têm a mesma distribuição, para quaisquer $t_1, \dots, t_n, k \in \mathbb{Z}$.

DEFINIÇÃO 1.3 - $\underline{Y}(t)$ é um processo *estacionário em sentido amplo* ou *fracamente estacionário*, se

$$(1.7) \quad \left\{ \begin{array}{l} \text{a) } E|Y_j(t)|^2 < \infty \quad \forall j=1, \dots, r \\ \text{b) } E\{Y_j(t)\} = \text{constante independente de } t, j=1, \dots, r \\ \text{c) } c_{j\ell}(k) = \text{cov}\{Y_j(t+k), Y_\ell(t)\}, \quad j, \ell=1, \dots, r \end{array} \right.$$

Isto é, a covariância entre os valores das séries em dois tempos só depende da diferença dos tempos.

Dizemos que,

$Y(t)$, $t \in T$, é um processo real de parâmetro discreto se $T \subset \mathbb{Z}$ e $Y(t)$, $t \in T$, é um processo real de parâmetro contínuo se T é um intervalo da reta.

A série temporal pode apresentar:

- a) observações contínuas;
- b) observações discretas, a intervalos regulares;
- c) observações discretas, a intervalos irregulares;

Pouco trabalho teórico foi realizado sobre este último tipo de série. Podemos citar Granger (1963) que estudou os efeitos de intervalos diferentes do mensal em séries econômicas.

São exemplos de séries temporais:

- a) a quantidade diária de chuva em uma região da superfície terrestre;
- b) a produção anual de café em São Paulo;
- c) o registro eletrocardiográfico de um paciente, no intervalo de 60 segundos;
- d) o registro horário da temperatura, segundo o serviço meteorologia do Instituto Astronômico e Geofísico de São Paulo.

No Apêndice 1 deste trabalho, apresentamos alguns exemplos de séries e seus gráficos.

O levantamento de vários exemplos de séries tempo-

rais em economia sugere que podemos considerar a série geral $Y(t)$ como a mistura de três componentes:

$m(t)$ = tendência, ou movimento a longo prazo;

$s(t)$ = componente sazonal;

$\epsilon(t)$ = componente aleatório, também chamado irregular ou residual.

No que se segue, suporemos que $\epsilon(t)$ seja uma série estacionária em sentido amplo, e $t \in \mathbb{Z}$.

Podemos, pois, representar a série como:

$$(1.8) \left\{ \begin{array}{ll} Y(t) = m(t) + s(t) + \epsilon(t) & \text{(modelo aditivo)} \\ \text{ou} & \\ Y(t) = m(t) \cdot s(t) \cdot \epsilon(t) & \text{(modelo multiplicativo)} \\ \text{ou} & \\ Y(t) = m(t) \cdot s(t) + \epsilon(t) & \text{(modelo misto)} \end{array} \right.$$

O problema da existência do componente $c(t)$, isto é, ciclos ou flutuações sobre a tendência, quase não tem sido levado em conta e não será discutido neste trabalho. Mas podemos citar um teste, apresentado por Hannan (1961) para saber se a série possui ciclos.

Um dos objetivos da análise de uma série temporal é separar os componentes, para um estudo individual. Entretanto, devemos lembrar que, ao fazer isso, estamos impondo um modelo à situação.

É razoável supor que:

- a) tendências são devidas a forças permanentes que operam com certa uniformidade, mais ou menos na mesma direção;
- b) flutuações a curto prazo, sobre esses movimentos longos, são devidas a causas variadas e incontrolláveis;
- c) existe em ambas, algum distúrbio atribuível a eventos aleatórios, que fornece o resíduo.

Mas isso é apenas suposição, e pressupõe um tipo de hipótese que devemos rejeitar, caso nosso modelo não seja adequado aos dados observados.

Possivelmente, os componentes mais fáceis de se compreender são aqueles devidos, com certeza, a geradores cíclicos tais como: variações diurnas de temperatura, movimentos de maré associados com o mês lunar e variações sazonais. Devemos tomar cuidado para não confundir estes efeitos com flutuações de tipo pseudo-cíclicas, como por exemplo os "ciclos" de comércio ou os "ciclos" de explosão solar, nos quais não existe nenhum elemento subjacente de tipo periódico.

Um fator importante no estudo de Séries Temporais é o tamanho da série, definido como "o espaço entre a primeira e a última observação" ou como "o número de pontos encontrados". A precisão das estimativas não depende apenas do tamanho da série, mas também de sua estrutura interna, porque os valores sucessivos não são independentes. Essa dependência entre valores que estão próximos no tempo é chamada "correlação

serial" ou auto-correlação e é definida como segue:

A quantidade

$$(1.9) \quad \rho(k) = \frac{E\{Y(t)Y(t+k)\}}{E\{Y(t)\}^2}, \quad \rho(-k) = \rho(k), \quad t, k \in \mathbb{Z},$$

é a auto-correlação de ordem k , onde o numerador

$$\gamma(k) = E\{Y(t)Y(t+k)\}, \quad \gamma(-k) = \gamma(k), \quad t, k \in \mathbb{Z},$$

é chamado auto-covariância de ordem k .

Vamos considerar n observações $Y(1), \dots, Y(n)$ de uma Série Temporal $\{Y(t), t \in \mathbb{Z}\}$.

Os $n-1$ pares $[Y(1), Y(2)], [Y(2), Y(3)], \dots, [Y(n-1), Y(n)]$ constituem um conjunto de valores bivariados que têm um coeficiente de correlação de tipo padrão. O mesmo ocorre com os $n-2$ pares $[Y(1), Y(3)], \dots, [Y(n-2), Y(n)]$.

Nós chamamos de correlação serial de ordem k a correlação entre os termos $Y(t)$ e $Y(t+k)$, $t, k \in \mathbb{Z}$. Isto é:

$$(1.11) \quad r(k) = \frac{\frac{1}{n-k} \sum_{i=1}^{n-k} \left\{ \left[Y(i) - \frac{1}{n-k} \sum_{i=1}^{n-k} Y(i) \right] \left[Y(i+k) - \frac{1}{n-k} \sum_{i=1}^{n-k} Y(i+k) \right] \right\}}{\left\{ \frac{1}{n-k} \sum_{i=1}^{n-k} \left[Y(i) - \frac{1}{n-k} \sum_{i=1}^{n-k} Y(i) \right]^2 \cdot \frac{1}{n-k} \sum_{i=1}^{n-k} \left[Y(i+k) - \frac{1}{n-k} \sum_{i=1}^{n-k} Y(i+k) \right]^2 \right\}^{\frac{1}{2}}}$$

Esta expressão parece complicada porque o coeficiente de ordem k depende dos valores $Y(1), \dots, Y(n-k)$ para um membro do par, e dos valores $Y(1+k), \dots, Y(n)$ para o outro membro. Nós podemos modificar essa igualdade, medindo todas as variáveis em relação ao tamanho total da série. Temos então

$$(1.12) \quad r(k) = \frac{\frac{1}{n-k} \sum_{i=1}^{n-k} [Y(i) - \bar{Y}][Y(i+k) - \bar{Y}]}{\frac{1}{n} \sum_{i=1}^n [Y(i) - \bar{Y}]^2} \quad \text{onde} \quad \begin{cases} \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y(i) \\ r(-k) = r(k) \end{cases}$$

A correlação serial não impede que o método tradicional de mínimos quadrados forneça estimativas por ponto, razoavelmente boas, dos parâmetros de uma série temporal. Entretanto, ela prejudica as estimativas usuais de erros-padrão e, portanto, interfere na construção de intervalos de confiança e nos testes de hipóteses. Além disso, muitas séries, especialmente as de fenômenos sociais, se movem de modo muito complicado e por isso não é prático ajustar-lhes curvas matemáticas de tal forma que até mesmo estimativas por ponto não possam ser obtidas.

Os passos para análise de uma série temporal podem ser os seguintes (Granger, 1964):

- 1º) Investigação: através da inspeção visual, pode-se formar estatísticas sumárias úteis e fazer comentários do tipo "A série é sempre positiva".
- 2º) Ajustamento de um modelo, composto de 3 estágios:
 - 2a) limitação da escolha de um número finito de tipos de modelos possíveis;
 - 2b) seleção de um destes tipos;
 - 2c) seleção de um modelo, dentro do tipo escolhido; isto é, seleção numérica dos parâmetros.

Houve época em que o ajustamento de modelos foi con

siderado essencial, mas agora isso não é mais enfatizado; principalmente quando se percebe que as relações entre séries são tão complicadas que dificilmente elas poderão ser ajustadas a um modelo compreensível.

- 3º) Aplicação: tentativa de prever resultados futuros e de comparar o modelo ajustado com uma teoria particular.

Kendall, 1973, sugere os seguintes tipos de estudo de Séries Temporais:

- 1º) Ao nível mais superficial, tomamos uma Série Temporal particular e construímos um sistema simples, usualmente de tipo matemático, que "descreve" seu comportamento de modo conciso.
- 2º) Podemos tentar "explicar" o comportamento da série em termos de outras variáveis, e relacionar as observações com algumas regras estruturais de comportamento, através de um modelo.
- 3º) Podemos usar a análise resultante para tentar "prever" o comportamento futuro da série; conhecer o mecanismo gerador e fazer projeções futuras com confiança.
- 4º) Por meio do 2º item, podemos desejar "controlar" o sistema, tentando detectar sinais de eventos futuros ou examinar o que poderá ocorrer se alterarmos algum parâmetro.

59) De modo mais geral, podemos considerar o progresso conjunto de algumas variáveis através do tempo, isto é, considerar a variável como um vetor de observações. Neste caso, do ponto de vista estatístico, estaremos nos aproximando da construção de um modelo matemático. Muitas vezes é difícil distinguir onde termina a teoria de Série Temporal Múltipla e começa a de construção de modelos.

Neste trabalho, nós nos propusemos a abordar o problema do ajustamento sazonal de séries temporais. Mas, antes de introduzir o assunto, vamos tomar contacto com um método de análise de séries que tem sido bastante desenvolvido nas duas últimas décadas: a Análise Espectral.

CAPÍTULO II

TEORIA ESPECTRAL

2.1 - INTRODUÇÃO

Vamos abordar aqui alguns conceitos sobre análise espectral, necessários à compreensão de tópicos desenvolvidos nos capítulos V e VI, a saber:

- O espectro e sua relação com a auto-covariância.
- O espectro cruzado.
- Operação de filtragem. Propriedades.
- Estimação do espectro e distribuição amostral dos estimadores.

No capítulo I, vimos que, se $\underline{Y}(t)$ é um processo estacionário vetorial, podemos definir a função de covariância cruzada de ordem k como

$$\gamma_{j\ell}(k) = E\{Y_j(t) \underline{Y}_\ell(t+k)\}, \quad j, \ell=1, \dots, v \text{ e } k \in \mathbb{Z}.$$

Agora, se $\sum_{k=-\infty}^{\infty} |\gamma_{j\ell}(k)| < \infty$, então podemos definir a função de densidade espectral cruzada entre a série $Y_j(t)$ e a série $Y_\ell(t)$ como

$$(2.1) \quad f_{j\ell}(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_{j\ell}(k) e^{-i\lambda k}, \quad -\infty < \lambda < \infty,$$

que é a transformada de Fourier da auto-covariância. A relação inversa será

$$(2.2) \quad \gamma_{j\ell}(k) = \int_{-\pi}^{\pi} f_{j\ell}(\lambda) \cdot e^{i\lambda k} d\lambda, \quad k \in \mathbb{Z},$$

λ = frequência angular por unidade de tempo.

Se $j = \ell \Rightarrow f_{jj}(\lambda)$ é o espectro de $Y_j(t)$, e $f_{jj}(\lambda)$ é limitada, uniformemente contínua e tem período 2π , isto é $f_{jj}(\lambda + 2\pi) = f_{jj}(\lambda)$.

Além disso, como $\gamma_{jj}(-k) = \gamma_{jj}(k)$, então

$$\begin{aligned} f_{jj}(-\lambda) &= \frac{1}{2\pi} \sum_k \gamma_{j\ell}(k) e^{i\lambda k} \\ &= \frac{1}{2\pi} \sum_k \gamma_{j\ell}(-k) e^{-i\lambda k} \\ &= \frac{1}{2\pi} \sum_k \gamma_{j\ell}(k) e^{i\lambda k} = f_{jj}(\lambda) \end{aligned}$$

2.2 - DESCRIÇÃO DA SÉRIE TEMPORAL NO DOMÍNIO DO TEMPO

Como já definimos no capítulo I, a Série Temporal é uma "realização" de um processo estocástico. Isto é, a série é uma amostra única de um processo gerador particular.

Por "análise" da Série Temporal, entende-se a estimação e reconstrução das propriedades do processo gerador subjacente.

São exemplos de processos geradores:

EXEMPLO 2.1 - Processo de Média Movel

$$Y(t) = \sum_{k=0}^m b_k \epsilon(t-k),$$

onde b_k são constantes complexas e as $\epsilon(t)$ são variáveis aleatórias independentes.

EXEMPLO 2.2 - Processo Auto-Regressivo

$$Y(t) + \sum_{k=1}^m a_k Y(t-k) = \epsilon(t),$$

onde a_k são constantes complexas e as $\epsilon(t)$ são variáveis aleatórias independentes.

O valor atual $Y(t)$ é formado com a soma linear de valores passados e com um termo independente.

EXEMPLO 2.3 - Processo Misto: Médias Móveis - Auto Regressivo

$$Y(t) + \sum_{k=1}^m a_k Y(t-k) = \sum_{k=0}^m b_k \epsilon(t-k).$$

Se o processo gerador for normal, então é usual tentar explicar suas propriedades através de seus momentos de 1ª e 2ª ordem; isto é,

$$(2.3) \begin{cases} m(t) = E\{Y(t)\} & \text{(Média)} \\ \sigma^2(t) = E\{[Y(t)-m(t)]^2\} & \text{(Variância)} \\ \mu(t,s) = E\{[Y(t)-m(t)][Y(s)-m(s)]\} & \text{(Auto-covariância)} \end{cases}$$

determinam o processo $Y(t)$.

Em geral, estes momentos são funções do tempo t .

Mas, lembrando as definições do Capítulo I, devemos dizer da importância de séries geradas por processos cujos momentos de 1a. e 2a. ordem não são funções de t . Isto é:

$$(2.4) \quad \begin{cases} E\{Y(t)\} = m \\ E\{[Y(t)-m]^2\} = \sigma^2 \\ E\{[Y(t)-m][Y(s)-m]\} = \mu(j) \text{ onde } j=t-s \end{cases}$$

Estas séries são chamadas estacionárias de 2a. ordem, geradas por processos estacionários de 2a. ordem.

Em geral consideramos estacionária a série livre do componente de tendência, desde que o resíduo $\epsilon(t)$ seja estacionário e $s(t)$ seja suposta periódica.

Pela teoria geral de estimação, nós deveríamos dispor de M amostras do processo,

$$Y^{(k)}(t), \quad k=1, \dots, M,$$

para estimar $m(t)$ através de

$$\bar{Y}(t) = \frac{1}{M} \sum_{k=1}^M Y^{(k)}(t)$$

chamada "média de conjunto" e que goza da propriedade

$$\lim_{M \rightarrow \infty} \bar{Y}(t) = m(t)$$

Mas nós temos apenas *uma* amostra do processo. Por is

so, vamos nos valer do teorema Ergódico, segundo o qual podemos estimar os momentos através de médias sobre o tempo.

TEOREMA 2.1 - Teorema Ergódico - Seja $Y(t)$ um processo estacionário. Sob determinadas condições de regularidade, valem as seguintes relações:

$$(2.6) \left\{ \begin{array}{l} m = E\{Y(t)\} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N Y(t) \\ \mu(j) = E\{Y(t)Y(t+j)\} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N Y(t)Y(t+j) \end{array} \right\} \text{ se } t \text{ é discreto}$$

$$(2.7) \left\{ \begin{array}{l} m = E\{Y(t)\} = \lim_{T \rightarrow \infty} \int_0^T Y(t) dt \\ \mu(j) = E\{Y(t)Y(t+j)\} = \lim_{T \rightarrow \infty} \int_0^T Y(t)Y(t+j) dt \end{array} \right\} \text{ se } t \text{ é contínuo}$$

(Os limites são tomados em média quadrática).

Então, no caso discreto, as estimativas serão:

$$(2.8) \left\{ \begin{array}{l} \bar{Y} = \frac{1}{N} \sum_{t=1}^N Y(t) \\ s^2 = \frac{1}{N} \sum_{t=1}^N [Y(t) - \bar{Y}]^2 \\ c(j) = \frac{1}{2N} \sum_{t=-N+M}^{N-M} \{ [Y(t-j)Y(t)] + [Y(t)Y(t+j)] \} \quad j=0,1,\dots,M \end{array} \right.$$

2.3 - DESCRIÇÃO DA SÉRIE TEMPORAL NO DOMÍNIO DE FREQUÊNCIA

A vantagem de descrever a série no domínio de fre-

qüência reside no fato de se poder eliminar o problema da correlação serial, que foi definida no capítulo I. Na análise espectral, os componentes são ortogonais.

Vamos inicialmente considerar uma série normal.

DEFINIÇÃO 2.1 - Uma série $Y(t)$ é chamada normal se,

$$\forall n \text{ e } t_1, t_2, \dots, t_n \in \mathbb{Z},$$

então $Y(t_1), \dots, Y(t_n)$ têm distribuição normal n-variada.

Supondo todas as médias iguais a zero, a lei de probabilidade que governa o processo fica especificada pelas auto-covariâncias

$$E\{Y(t_i)Y(t_k)\} \text{ , } i,k=1,2,\dots,n$$

E, se o processo é estacionário, as auto-covariâncias só dependem de $j=t_i-t_k$.

EXEMPLO 2.4 - Exemplo de um Processo Aleatório Normal de Mé-dia Zero.

Sejam $0 \leq \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N \leq \pi$ constantes e a_i, b_i , $i=1, \dots, N$, variáveis aleatórias normais, independentes, com

$$(2.9) \quad \begin{cases} E\{a_i\} = E\{b_i\} = 0 & \forall i=1,\dots,N \\ E\{a_i b_k\} = 0 & \forall i,k \\ E\{a_i a_k\} = E\{b_i b_k\} = \begin{cases} \sigma_i^2 & \text{se } i=k \\ 0 & \text{se } i \neq k \end{cases} \end{cases}$$

A Série Temporal

$$(2.10) \quad Y(t) = \sum_{i=1}^N \{a_i \cos \lambda_i t + b_i \sin \lambda_i t\}$$

é uma série normal, de média zero.

Esta série é estacionária. De fato:

$$\begin{aligned} E\{Y(t)Y(t+j)\} &= E\left\{\sum_{i=1}^N \sum_{k=1}^N [a_i a_k \cos \lambda_i t \cos \lambda_k(t+j) + \right. \\ &\quad + a_i b_k \cos \lambda_i t \sin \lambda_k(t+j) + \\ &\quad + a_k b_i \cos \lambda_k(t+j) \sin \lambda_i t + \\ &\quad \left. + b_i b_k \sin \lambda_i t \sin \lambda_k(t+k)]\right\} \\ (2.9) \quad &= \sum_{i=1}^N \sigma_i^2 [\cos \lambda_i t \cos \lambda_i(t+j) + \sin \lambda_i t \sin \lambda_i(t+j)] \\ &= \sum_{i=1}^N \sigma_i^2 \cos \lambda_i \cdot j \\ &= \mu(j). \end{aligned}$$

A variância σ_i^2 , como função da frequência λ_i , é chamada *espectro* de $Y(t)$.

σ_i^2 e $\mu(j)$ formam um par de Fourier, isto é,

$$\sigma_i^2 = \frac{1}{\pi} \int_{-\pi}^{\pi} \mu(j) \cos \lambda_i \cdot j \, dj.$$

De fato:

$$\begin{aligned}
 \frac{1}{\pi} \int_{-\pi}^{\pi} \mu(j) \cos \lambda_k j \, dj &= \frac{1}{\pi} \int_{-\pi}^{\pi} \left\{ \sum_{k=1}^N \sigma_i^2 \cos \lambda_i j \cos \lambda_k j \right\} dj \\
 &= \frac{1}{\pi} \sum_{k=1}^N \sigma_i^2 \int_{-\pi}^{\pi} \cos \lambda_i j \cos \lambda_k j \, dj \\
 &= \frac{1}{\pi} \sum_{k=1}^N \sigma_i^2 \cdot \begin{cases} \pi & \text{se } \lambda_i = \lambda_k \\ 0 & \text{se } \lambda_i \neq \lambda_k \end{cases} \\
 &= \sigma_i^2
 \end{aligned}$$

Podemos, então, considerar qualquer das duas funções para descrever o processo: a auto-covariância descreve no domínio do tempo, e o espectro no de frequência. A relação entre as duas funções tem o nome de "Relação de Wiener-Khinchine".

Como apontou Tukey (1961), a estimação do espectro da série é análoga à análise dos componentes da variância, da Estatística clássica.

$$j = 0 \Rightarrow E\{Y(t)\}^2 = \text{Var}\{Y(t)\} = \sum_{i=1}^N \sigma_i^2$$

∴ a variância de $Y(t)$ é devida à soma das variâncias dos diferentes componentes de frequência.

Se os a_i e b_i de (2.10) não são supostos normais, o processo $Y(t)$ não será normal; mas, assim mesmo, o espectro e a auto-covariância descreverão o processo.

A pergunta que surge é se qualquer processo aleatório estacionário pode ser igualmente bem descrito, como acima, nos domínios de frequência e de tempo. A resposta é afirmativa, desde que não nos restrinjamos a esquemas paramétricos finitos como em (2.10).

Pode-se mostrar que qualquer processo aleatório $Y(t)$, estacionário até 2a. ordem, pode ser expresso na forma espectral

$$(2.11) \quad Y(t) = \int_0^\infty \cos \lambda t \, dU(\lambda) + \int_0^\infty \sin \lambda t \, dV(\lambda)$$

Isto é, o processo é "decomposto" em uma superposição de ondas seno e cosseno de frequências, e com amplitudes aleatórias $dU(\lambda)$ e $dV(\lambda)$ que têm as seguintes propriedades:

$$(2.12) \quad \begin{cases} E\{dU(\lambda)dU(\lambda')\} = E\{dV(\lambda)dV(\lambda')\} = 0, & \lambda \neq \lambda' \\ E\{dU(\lambda)dV(\lambda')\} = 0 & \forall \lambda, \lambda' \\ E\{dU(\lambda)\}^2 = E\{dV(\lambda)\}^2 = dF(\lambda) \end{cases}$$

A função $F(\lambda)$ é chamada "função de distribuição espectral" de $Y(t)$, sendo $F(-\infty) = 0$, $F(\infty) = \text{Var}\{Y(t)\}$. Ela é monotonicamente não decrescente e pode ser escrita na forma

$$F(\lambda) = F_1(\lambda) + F_2(\lambda) + F_3(\lambda)$$

onde:

$F_1(\lambda)$ é absolutamente contínua. É o componente mais interessante, do ponto de vista estatístico.

Podemos colocar $dF_1(\lambda)/d\lambda = f(\lambda)$.

O problema fundamental da análise espectral é estimar $f(\lambda)$, a partir de observações de $Y(t)$ num intervalo de tempo.

$F_2(\lambda)$ é uma função em escada. No caso de séries definidas como em (2.10), os degraus são os σ_i^2 nas frequências λ_i .

$F_3(\lambda)$ é o componente singular de $F(\lambda)$ e, segundo Nerlove (1964), pode ser desprezado.

A igualdade (2.11) é uma representação de $Y(t)$ no domínio de frequência do processo.

A série $Y(t)$ é normal se e só se as variáveis aleatórias $dU(\lambda)$ e $dV(\lambda)$ são normais, independentes, univariadas.

Vemos, portanto, que a lei de Probabilidade que governa o processo (2.11) é completamente especificada pela $F(\lambda)$ ou pela auto-covariância $\mu(j)$, que formam um par de Fourier.

De fato:

$$\mu(j) = E\{Y(t)Y(t+j)\}$$

$$\begin{aligned} &= E\left\{\int_0^\infty \int_0^\infty \cos \lambda t \cos \lambda'(t+j) dU(\lambda)dU(\lambda') + \right. \\ &+ \int_0^\infty \int_0^\infty \cos \lambda t \sin \lambda'(t+j)dU(\lambda)dV(\lambda') + \\ &+ \int_0^\infty \int_0^\infty \cos \lambda'(t+j) \sin \lambda t dU(\lambda')dV(\lambda) + \\ &\left. + \int_0^\infty \int_0^\infty \sin \lambda t \sin \lambda'(t+j)dV(\lambda)dV(\lambda')\right\} \end{aligned}$$

$$\begin{aligned} &\stackrel{(2.12)}{=} \int_0^\infty [\cos \lambda t \cos \lambda(t+j) + \sin \lambda t \sin \lambda(t+j)]dF(\lambda) \\ &= \int_0^\infty \cos \lambda j dF(\lambda). \end{aligned}$$

Caso importante:

$$j = 0 \Rightarrow \mu(0) = \int_0^\infty \cos 0 dF(\lambda) = \int_0^\infty dF(\lambda),$$

onde

$$\mu(0) = \text{Var}\{Y(t)\}$$

Logo, a variância pode ser decomposta nas contribuições dos diferentes componentes de frequência da série.

A representação espectral de $Y(t)$ pode ser colocada na forma complexa, lembrando que:

$$\cos \lambda t = \frac{e^{i\lambda t} + e^{-i\lambda t}}{2} \text{ e } \operatorname{sen} \lambda t = \frac{e^{i\lambda t} - e^{-i\lambda t}}{2i}$$

Então, a (2.11) fica

$$\begin{aligned} Y(t) &= \int_0^\infty \frac{e^{i\lambda t} + e^{-i\lambda t}}{2} dU(\lambda) + \int_0^\infty \frac{e^{i\lambda t} - e^{-i\lambda t}}{2i} dV(\lambda) \\ &= \int_{-\infty}^\infty e^{i\lambda t} \frac{dU(\lambda)}{2} - \int_{-\infty}^\infty ie^{i\lambda t} \frac{dV(\lambda)}{2} \end{aligned}$$

ou

$$(2.13) \quad Y(t) = \int_{-\infty}^\infty e^{i\lambda t} dZ(\lambda)$$

onde

$$\begin{cases} dZ(\lambda) = \frac{1}{2}[dU(\lambda) - idV(\lambda)] \\ dU(-\lambda) = dU(\lambda) \\ dV(-\lambda) = -dV(\lambda) \end{cases}$$

A variância de uma variável aleatória complexa X é definida como o valor esperado do quadrado de seu módulo, isto é,

$$\operatorname{Var}(X) = E|X|^2 = E\{X\bar{X}\}$$

Logo, de (2.12), temos

$$E\{dZ(\lambda) \overline{dZ(\lambda)}\} = \frac{1}{2} dF(\lambda)$$

e

$$\begin{aligned}\mu(j) &= \int_0^{\infty} \cos \lambda j \, dF(\lambda) \\ &= \int_0^{\infty} \frac{e^{i\lambda j} + e^{-i\lambda j}}{2} \, dF(\lambda)\end{aligned}$$

ou

$$(2.14) \quad \mu(j) = \int_{-\infty}^{\infty} e^{i\lambda j} \frac{dF(\lambda)}{2}$$

No caso de $F(\lambda)$ ser absolutamente contínua, ou seja, $dF(\lambda) = f(\lambda) \, d\lambda$, temos

$$f(\lambda) = \frac{1}{\pi} \int_{-\infty}^{\infty} e^{-i\lambda j} \mu(j) \, dj, \quad -\infty < \lambda < \infty,$$

que é a função de densidade espectral de $Y(t)$.

E, novamente, o espectro e a auto-covariância são pares de Fourier.

2.4 - EXEMPLOS DE ESPECTRO E AUTO-COVARIANÇA

EXEMPLO 2.5 - O espectro de uma sequência $\varepsilon(t)$ aleatória independente é uma constante sobre o intervalo $(-\pi, \pi)$,

com

$$\mu(j) = E\{\varepsilon(t)\overline{\varepsilon(t+j)}\} = \begin{cases} 0 & \text{se } j \neq 0 \\ \sigma^2 & \text{se } j = 0 \end{cases}$$

Logo

$$\begin{aligned} f(\lambda) &= \frac{1}{2\pi} \sum_j e^{-i\lambda j} \mu(j) \\ &= \frac{1}{2\pi} \cdot \mu(0) \\ &= \sigma^2/2\pi \end{aligned}$$

Podemos escrever:

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

onde

$$f(\lambda) = E\{dZ(\lambda_1) \overline{dZ(\lambda_2)}\} = \begin{cases} 0 & \lambda_1 \neq \lambda_2 \\ \sigma^2/2\pi & \lambda_1 = \lambda_2 \end{cases}$$

A série $\epsilon(t)$ é chamada "ruído branco".

EXEMPLO 2.6 - Seja o processo de médias móveis do exemplo 2.1

$$Y(t) = \sum_{k=0}^m b_k \epsilon(t-k)$$

Usando o exemplo 2.5, temos

$$\begin{aligned} Y(t) &= \int_{-\pi}^{\pi} \sum_{k=0}^m b_k e^{i\lambda(t-k)} dZ(\lambda) \\ &= \int_{-\pi}^{\pi} e^{i\lambda t} \left\{ \sum_{k=0}^m b_k e^{-i\lambda k} \right\} dZ(\lambda). \end{aligned}$$

Então,

$$\mu(j) = E\{Y(t)\overline{Y(t+j)}\}$$

$$= \sigma^2 \cdot b(\lambda) \overline{b(\lambda)}$$

e

$$f(\lambda) = \frac{1}{2\pi} \sum_j e^{-i\lambda j} \mu(j)$$

$$= \frac{\sigma^2}{2\pi} b(\lambda) \overline{b(\lambda)}$$

onde

$$b(\lambda) = \sum_{k=0}^m b_k e^{-i\lambda k}$$

EXEMPLO 2.7 - Seja o Processo Auto-Regressivo do exemplo 2.2

$$Y(t) + \sum_{k=1}^m a_k Y(t-k) = \varepsilon(t)$$

Vamos supor

$$Y(t) = \int_{-\pi}^{\pi} e^{i\lambda t} dG(\lambda)$$

onde

$$\begin{cases} E|dG(\lambda)|^2 = f(\lambda)d\lambda \\ f(\lambda) = \text{densidade espectral de } Y(t). \end{cases}$$

Usando novamente o exemplo 2.5, temos:

$$\int_{-\pi}^{\pi} e^{i\lambda t} dG(\lambda) + \sum_{k=1}^m a_k \int_{-\pi}^{\pi} e^{i\lambda(t-k)} dG(\lambda) = \int_{-\pi}^{\pi} e^{i\lambda t} dZ(\lambda)$$

Então

$$\mu(j) = \sigma^2 \frac{1}{a(\lambda) \overline{a(\lambda)}}$$

e

$$\begin{aligned} f(\lambda) &= \frac{1}{2\pi} \sum_j \mu(j) e^{-i\lambda j} \\ &= \frac{\sigma^2}{2\pi} \cdot \frac{1}{a(\lambda) \overline{a(\lambda)}} \end{aligned}$$

onde

$$a(\lambda) = 1 + \sum_{k=1}^m a_k e^{-i\lambda k}$$

2.5 - FILTROS

Muitas séries temporais apresentam tendências, isto é, mudanças das médias através do tempo. Por isso, o espectro de uma série típica apresentará potências muito altas concentradas em frequências próximas de zero, diminuindo nas frequências mais altas. (ver Apêndice 2).

A existência de picos em determinadas frequências sugere o uso de uma operação de "filtragem" da série.

Ela é usada quando se deseja alterar a forma do es-

pectro, sendo possíveis os seguintes casos:

- 1º) Queremos estudar somente uma particular faixa de frequência $(\lambda_i, \lambda_{i+1})$;
- 2º) Queremos remover inteiramente uma particular frequência λ_0 , que dá grande contribuição à variância (em λ_0 , o espectro tem um pico alto); isso permite estudar melhor outras frequências;
- 3º) Queremos "regularizar" a série a fim de estudar melhor seus movimentos de longa duração.

DEFINIÇÃO 2.2 - O filtro é um conjunto de operações matemáticas, usadas para transformar os dados originais da série $Y(t)$ em outra série $X(t)$, que será analisada. Segundo a teoria espectral, nós filtramos a série $Y(t)$ e depois comparamos os espectros de $Y(t)$ e de $X(t)$.

Seja uma operação A cujo domínio consiste de séries $Y(t)$, e cujo contradomínio consiste de séries $X(t)$, $t \in \mathbb{Z}$.

Indicamos:

$$X(t) = A[Y(t)].$$

DEFINIÇÃO 2.3 - A é linear se, para séries $Y_1(t)$ e $Y_2(t)$ no domínio de A , e com α_1 e α_2 constantes, temos

$$A[\alpha_1 Y_1(t) + \alpha_2 Y_2(t)] = \alpha_1 A[Y_1(t)] + \alpha_2 A[Y_2(t)]$$

DEFINIÇÃO 2.4 - A é invariante no tempo, se

$$A\{T^u[Y(t)]\} = A[Y(t+u)] = X(t+u)$$

onde T^u é o operador translação, isto é,

$$T^u[Y(t)] = Y(t+u)$$

Um exemplo de sistema linear e invariante no tempo é

$$(2.15) \quad X(t) = \int_{-\infty}^{\infty} a(u) Y(t-u) du$$

onde

$$\begin{cases} Y(t) \text{ é a "entrada" do sistema} \\ X(t) \text{ é a "saída" do sistema} \\ a(u) \text{ é a função de impulso ou núcleo} \\ \text{do filtro.} \end{cases}$$

O problema fundamental é estimar $a(u)$ ou sua transformada de Fourier

$$(2.16) \quad \begin{aligned} A(\lambda) &= \int_{-\infty}^{\infty} a(u) e^{-i\lambda u} du \\ &= u(\lambda) + iv(\lambda) \end{aligned}$$

chamada "função de transferência". Estimada $A(\lambda)$, dizemos que o sistema está identificado.

A importância do filtro linear em Análise Espectral se deve ao fato de que os espectros da saída e da entrada são relacionados de forma simples, como veremos agora.

De (2.16), temos:

$$\begin{aligned} A(\lambda) \cdot e^{i\lambda t} &= [u(\lambda) + iv(\lambda)][\cos \lambda t + i \sin \lambda t] \\ &= [u(\lambda) \cos \lambda t - v(\lambda) \sin \lambda t] + i[u(\lambda) \sin \lambda t + v(\lambda) \cos \lambda t], \end{aligned}$$

mas

$$(2.17) \quad A(\lambda) e^{i\lambda t} = A(\lambda) \cos \lambda t + iA(\lambda) \sin \lambda t.$$

Então, vemos que

- a) A parte real é a saída de um sistema cuja entrada é u uma onda coseno de frequência igual a λ e amplitude igual a 1.
- b) A parte imaginária é a saída de um sistema cuja entrada é uma onda seno de frequência igual a λ e amplitude igual a 1.

Seja o ângulo $\phi(\lambda)$ definido como

$$\phi(\lambda) = \arg[u(\lambda) + iv(\lambda)]$$

$$= \arctan \left[\frac{v(\lambda)}{u(\lambda)} \right]$$

tal que

$$\begin{cases} \cos \phi(\lambda) = \frac{u(\lambda)}{[u^2(\lambda) + v^2(\lambda)]^{1/2}} \\ \sin \phi(\lambda) = \frac{v(\lambda)}{[u^2(\lambda) + v^2(\lambda)]^{1/2}} \end{cases}$$

Aplicando em (2.17), temos:

$$A(\lambda)\cos \lambda t = [u^2(\lambda)+v^2(\lambda)]^{1/2}\cos[\lambda t+\phi(\lambda)]$$

$$A(\lambda)\sin \lambda t = [u^2(\lambda)+v^2(\lambda)]^{1/2}\sin[\lambda t+\phi(\lambda)] \text{ .}$$

Logo,

a') A parte real da saída do sistema é uma onda cosseno com amplitude igual

$$[u^2(\lambda)+v^2(\lambda)]^{1/2}$$

e fase igual a

$$\phi(\lambda) \neq 0 \text{ .}$$

b') A parte imaginária da saída do sistema é uma onda seno também com amplitude igual

$$[u^2(\lambda)+v^2(\lambda)]^{1/2}$$

e fase igual a

$$\phi(\lambda) \neq 0 \text{ .}$$

Portanto, o filtro modifica a amplitude da onda por um fator $[u^2(\lambda)+v^2(\lambda)]^{1/2}$ e desloca sua origem, no tempo, por um fator $\phi(\lambda)/\lambda$.

Os efeitos do filtro se tornam mais claros se colocarmos $A(\lambda)$ na forma polar

$$A(\lambda) = G(\lambda) \cdot e^{i\phi(\lambda)} \text{ ,}$$

onde $G(\lambda) = [u^2(\lambda)+v^2(\lambda)]^{1/2}$ é o "ganho" do filtro e $\phi(\lambda)$ é o "ângulo de fase" do filtro.

Agora podemos relacionar os espectros das séries $Y(t)$ e $X(t)$.

Seja a série dada por (2.13)

$$Y(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dZ_Y(\lambda)$$

e seja o sistema (2.15)

$$X(t) = \int_{-\infty}^{\infty} a(u) Y(t-u) du$$

$$(2.13) \quad = \int_{-\infty}^{\infty} a(u) \left[\int_{-\infty}^{\infty} e^{i\lambda(t-u)} dZ_Y(\lambda) \right] du$$

$$= \int_{-\infty}^{\infty} e^{i\lambda t} \left[\int_{-\infty}^{\infty} a(u) e^{-i\lambda u} du \right] dZ_Y(\lambda)$$

$$(2.16) \quad = \int_{-\infty}^{\infty} e^{i\lambda t} \cdot A(\lambda) \cdot dZ_Y(\lambda) \quad (I)$$

mas, de (2.13) também vale

$$X(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dZ_X(\lambda) \quad (II)$$

Comparando (I) e (II), temos:

$$(2.18) \quad dZ_X(\lambda) = A(\lambda) \cdot dZ_Y(\lambda)$$

De (2.12), vem

$$E\{dZ_X(\lambda) \overline{dZ_X(\lambda)}\} = \frac{1}{2} dF_{XX}(\lambda)$$

$$= \frac{1}{2} f_{XX}(\lambda) d\lambda \quad \text{se } F \text{ é absolutamente contínua.}$$

De (2.18), vem

$$E\{A(\lambda) dZ_Y(\lambda) \overline{A(\lambda) dZ_Y(\lambda)}\} = \frac{1}{2} f_{XX}(\lambda) d\lambda ,$$

ou

$$A(\lambda) \overline{A(\lambda)} E\{dZ_Y(\lambda) \overline{dZ_Y(\lambda)}\} = \frac{1}{2} f_{XX}(\lambda) d\lambda ,$$

ou

$$|A(\lambda)|^2 \cdot \frac{1}{2} f_{YY}(\lambda) d\lambda = \frac{1}{2} f_{XX}(\lambda) d\lambda .$$

Logo,

$$(2.19) \quad f_{XX}(\lambda) = |A(\lambda)|^2 f_{YY}(\lambda) .$$

EXEMPLO 2.8 - Seja o "Ruído Branco" do exemplo 2.5

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

com

$$f_{YY}(\lambda) = \sigma^2 / 2\pi .$$

E seja

$$X(t) = \sum_{k=0}^m b_k \epsilon(t-k)$$

o filtro de médias móveis do exemplo 2.6, aplicado a $\epsilon(t)$: Te-
mos

$$\begin{aligned} f_{XX}(\lambda) &= \frac{\sigma^2}{2\pi} A(\lambda) \overline{A(\lambda)} \\ &= f_{YY}(\lambda) \cdot |A(\lambda)|^2 \end{aligned}$$

onde

$$A(\lambda) = \sum_{k=0}^m b_k e^{-i\lambda k}$$

é a função de transferência do filtro.

2.6 - O ESPECTRO-CRUZADO

Lembrando, de (2.14), que

$$\mu_{YY}(j) = E\{Y(t)Y(t+j)\} = \int_{-\infty}^{\infty} e^{i\lambda j} f_{YY}(\lambda) d\lambda$$

podemos colocar de modo análogo, se $E\{X(t)\}=0$ e $E\{Y(t)\}=0$

$$(2.20) \quad \mu_{XY}(j) = E\{Y(t)X(t+j)\} = \int_{-\infty}^{\infty} e^{i\lambda j} f_{XY}(\lambda) d\lambda$$

onde $\mu_{XY}(j)$ é função de covariância cruzada entre $X(t)$ e $Y(t)$ e

$$(2.21) \quad f_{XY}(\lambda) = c(\lambda) - i q(\lambda)$$

DEFINIÇÃO 2.5 - $f_{XY}(\lambda) = c(\lambda) - i q(\lambda)$ é o espectro-cruzado entre $Y(t)$ e $X(t)$, onde

$c(\lambda)$ = co-espectro

$q(\lambda)$ = espectro quadratura.

Na representação complexa de processos, temos de (2.13)

$$X(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dZ_X(\lambda)$$

$$Y(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dZ_Y(\lambda)$$

onde

$$E\{dZ_X(\lambda_1)\overline{dZ_Y(\lambda_2)}\} = \begin{cases} 0 & \text{se } \lambda_1 \neq \lambda_2 \\ \frac{1}{2} f_{XY}(\lambda) & \text{se } \lambda_1 = \lambda_2 \end{cases}$$

Só nos interessa considerar as relações entre um particular componente de freqüência de $Y(t)$ e o mesmo componente de $X(t)$. Isto é, até que ponto o componente de freqüência λ de $Y(t)$ está relacionado com o de $X(t)$; e, ainda, "quanto" os dois componentes estão fora de fase.

DEFINIÇÃO 2.6 - O dobro da densidade co-espectral mede a covariância entre os componentes que estão "em fase", isto é:

$$\begin{aligned} (2.22) \quad 2c(\lambda)d\lambda &= E\{dU_X(\lambda)dU_Y(\lambda)\} \\ &= E\{dV_X(\lambda)dV_Y(\lambda)\} \end{aligned}$$

DEFINIÇÃO 2.7 - O dobro da densidade do espectro-quadratura mede a covariância entre os componentes que estão "em quadratura" ou $\Pi/2$ fora de fase, isto é:

$$\begin{aligned} (2.23) \quad 2q(\lambda)d\lambda &= E\{dU_X(\lambda)dV_Y(\lambda)\} \\ &= E\{dU_Y(\lambda)dV_X(\lambda)\} \end{aligned}$$

As medidas das relações entre os componentes de freqüência dos dois processos são as seguintes:

$$\text{1a. coerência em } \lambda: K(\lambda) = \frac{c^2(\lambda) + q^2(\lambda)}{f_{XX}(\lambda)f_{YY}(\lambda)} \quad 0 \leq K(\lambda) \leq 1$$

2a. fase: $\phi(\lambda) = \arctan [q(\lambda)/c(\lambda)]$

3a. ganho: $f_{YY}(\lambda) \cdot R_{XY}^2(\lambda) = f_{XX}(\lambda) \cdot K(\lambda)$

onde $R_{XY}(\lambda)$ é o coeficiente de regressão de $X(t)$ sobre $Y(t)$ na frequência λ .

2.7 - DISTRIBUIÇÃO AMOSTRAL DOS ESTIMADORES

Vamos estimar $f(\lambda)$ baseados em T observações

$$Y(0), Y(1), \dots, Y(T-1)$$

Seja

$$(2.24) \quad d^{(T)}(\lambda) = \sum_{t=0}^{T-1} Y(t) e^{-i\lambda t}, \quad -\infty < \lambda < \infty,$$

de período 2π , a transformada de Fourier finita de $Y(t)$, na frequência λ .

Os seguintes resultados podem ser demonstrados:

TEOREMA 2.3 (do Limite Central) - Se

$$\lambda_s \not\equiv 0 \pmod{2\pi} \quad \text{e} \quad \lambda_s \pm \lambda_q \not\equiv 0 \pmod{2\pi},$$

$s, q=1, \dots, m$, então:

$$d^{(T)}(\lambda_1), \dots, d^{(T)}(\lambda_m)$$

são assintoticamente independentes, e

$$(2.25) \quad d^{(T)}(\lambda_i) \rightarrow N^C(0; 2\pi T f(\lambda_i)), \quad i=1, \dots, m.$$

Então, um estimador de $f(\lambda)$ é

$$(2.26) \quad I^{(T)}(\lambda) = \frac{1}{2\pi T} \left| d^{(T)}(\lambda) \right|^2$$

chamado periodograma dos valores $Y(0), Y(1), \dots, Y(T-1)$.

Para demonstração e definição de variável normal complexa, ver Brillinger (1974), cap.4.

O teorema 2.3 também vale se

$$\lambda_s = \frac{2\pi s}{T} \text{ onde } \begin{cases} s \in \mathbb{Z} \\ \lambda_s \rightarrow \lambda \text{ quando } T \rightarrow \infty \end{cases}$$

Propriedades de $I(\lambda)$ (as mesmas do espectro):

$$P_1: I(\lambda + 2\pi) = I(\lambda)$$

$$P_2: I(\lambda) \geq 0$$

$$P_3: I(\lambda) \text{ é simétrico em relação a } \lambda = \pi$$

$$\text{se } T \rightarrow \infty \Rightarrow E\{I(\lambda)\} \rightarrow f(\lambda).$$

Logo, $I(\lambda)$ é assintoticamente não viciado.

TEOREMA 2.4 -

$$(2.27) \quad E\{I(\lambda)\} \approx \int_{-\pi}^{\pi} D(\lambda - \alpha) f(\alpha) d\alpha$$

onde

$$(2.28) \quad D(\lambda - \alpha) = \frac{1}{2\pi} \frac{\sin[T(\lambda - \alpha)/2]}{\sin[(\lambda - \alpha)/2]}$$

Isto é, o valor esperado de $I(\lambda)$ é uma média ponderada do espectro. O peso $D(\lambda - \alpha)$ é concentrado em torno de λ , para T grande.

TEOREMA 2.5 - Sejam

$$\begin{cases} s \in \mathbb{Z} \\ \lambda_s = \frac{2\pi s}{T} \rightarrow \lambda \text{ quando } T \rightarrow \infty \\ \lambda_s \pm \lambda_q \neq 0 \pmod{2\pi}, \quad 1 \leq s < q \leq m \end{cases}$$

Então os $I(\lambda_s)$ são assintoticamente independentes, cada um com distribuição limite

$$(2.29) \quad f(\lambda) \cdot \chi^2_2/2$$

Sejam

$$s \in \mathbb{Z} \quad \text{e} \quad \lambda_s = \frac{2\pi s}{T} \sim \lambda$$

Calculamos $2m+1$ ordenadas $I^{(T)}(\lambda_s)$ independentes, isto é:

$$(2.30) \quad I\left(\frac{2\pi(s+q)}{T}\right), \quad q=0, \pm 1, \pm 2, \dots, \pm m$$

e calculamos a média

$$(2.31) \quad f^{(T)}(\lambda_s) = \frac{1}{2m+1} \sum_{q=-m}^m I^{(T)}(2\pi(s+q)/T), \quad \lambda_s \neq 0 \pmod{2\pi}$$

que é chamado periodograma suavizado.

TEOREMA 2.6 - Sejam

$$\begin{cases} s \in \mathbb{Z} \\ \lambda_s^{(T)} = \frac{2\pi s}{T} \rightarrow \lambda_s \text{ quando } T \rightarrow \infty \\ \lambda_s \pm \lambda_q \neq 0 \pmod{2\pi} \end{cases}$$

Então os $f^{(T)}(\lambda_s)$ são assintoticamente independentes, com distribuição limite

$$(2.32) \quad f(\lambda) \cdot \chi^2_{2(2m+1)} / 2m+1, \quad \lambda \not\equiv 0 \pmod{\Pi}.$$

Dado que a distribuição de χ^2 é muito mais estável para um número maior de graus de liberdade, torna-se conveniente estimar o espectro da série através do periodograma suavizado.

+ □ +

-

CAPÍTULO III

ELIMINAÇÃO DA TENDÊNCIA

3.1 - CONSIDERAÇÕES GERAIS

Como já dissemos no Capítulo I, supomos a série temporal formada por três componentes: $m(t)$, $s(t)$ e $\epsilon(t)$.

Embora conceitualmente diferentes, parece que $m(t)$ e $s(t)$ são muito relacionados e, por isso, não podemos isolar um deles sem tentar isolar o outro. Se pudermos determinar $m(t)$ e $s(t)$ e removê-los da série, ficaremos com uma série que consiste num movimento oscilatório regular e suave, que suporemos ser estacionário.

Existem várias razões para se tentar examinar os efeitos sazonais:

- 1º comparar os valores de uma variável em diferentes épocas do ano;
- 2º remover os efeitos sazonais da série a fim de estudar os elementos não contaminados pela sazonalidade;

- 3º "corrigir" uma situação causada por efeitos sazonais;
- 4º prever valores futuros da série; esse caso só se aplica a séries estacionárias. Pouco trabalho foi realizado até hoje sobre previsão em séries não estacionárias.

Estes objetivos têm características diferentes e, por isso, um único método de determinação sazonal pode não ser conveniente para atingi-los. Na prática, diferentes técnicas têm sido empregadas.

Antes de apresentarmos os métodos de ajustamento sazonal (no Capítulo IV), vamos apresentar, os testes de tendência e os métodos para eliminá-la.

3.2 - TESTES PARA VERIFICAR A EXISTÊNCIA DE TENDÊNCIA NA SÉRIE

Wallis e Roberts (1956) apresentaram os testes através de um exemplo que aparece no Apêndice 1 deste trabalho: nível do Lago Huron-Michigan. São testes não paramétricos, cujas deduções podem ser encontradas em Draper-Smith (1966), Hajek (1969), Lindgren (1972) e Mood-Graybill (1963).

TESTE 1 - Teste de Postos

Sejam N observações de uma série temporal $Y(t), t=1, 2, \dots, N$.

Fixa-se o valor m_0 , mediana dos valores observados.

Atribui-se a cada valor da série o símbolo:

$$\begin{cases} A \text{ se ele } \hat{f}or \geq m_0 \\ B \text{ se ele } \hat{f}or < m_0 \end{cases}$$

Então, $N = n_1$ pontos A + n_2 pontos B.

Seja r_0 igual ao número observado de iterações ou subseqüências, isto é, grupos de símbolos iguais.

Para N grande, a variável aleatória r tem distribuiçã

$$r \sim N(\mu_r; \sigma_r)$$

onde

$$\begin{cases} \mu_r = \frac{2n_1n_2}{N} + 1 \\ \sigma_r = \sqrt{\frac{2n_1n_2(2n_1n_2 - N)}{N^2(N-1)}} \end{cases}$$

Podemos considerar não correlacionadas as observações próximas, a menos que N seja pequeno e $p=P(A)$ seja grande, isto é, $\frac{N-p}{N}$ seja muito pequeno.

E então o teste utiliza a mesma aproximação à normal da distribuição binomial, isto é:

$$z = \frac{r + \frac{1}{2} - \mu_r}{\sigma_r} : N(0;1)$$

Hipótese Nula: $H_0: r_0 = \mu_r$ (não existe tendência na série)

Hipótese alternativa: $H_a: r_0 < \mu_r$ (existe tendência).

Rejeitamos H_0 se $z < z_c$ onde z_c é tal que

$$P(z < z_c / H_0) = \alpha,$$

onde α é o nível de significância do teste.

O teste é unicaudal do lado esquerdo, porque apenas um número pequeno de iterações indica tendência na série.

EXEMPLO 3.1 - Na série do nível do lago Michigan-Huron foram encontrados os seguintes valores:

$$N = 96 \begin{cases} n_1 = 49 \text{ símbolos A} \\ n_2 = 47 \text{ símbolos B} \end{cases}$$

$$p = P(A) = 49/96 = 0,51$$

$$\frac{N-p}{N} = \frac{96-0,51}{96} \sim 1 \Rightarrow \text{dados não correlacionados.}$$

$$\left. \begin{array}{l} r_0 = 15 \text{ iterações} \\ \mu_r = 48,98 \\ \sigma_r = 4,87 \end{array} \right\} \Rightarrow z = \frac{15 + \frac{1}{2} - 48,98}{4,87} = -6,87$$

$$\alpha = 5\% \Rightarrow z_c = -1,64$$

conclusão: $z < z_c \Rightarrow$ rejeita-se H_0 . A série tem tendên-

cia.

O número de iterações é menor do que seria esperado, se as observações fossem independentes.

TESTE 2 - Teste do Sinal

Sejam N observações de uma série.

Em cada observação, a partir da segunda, coloca-se o sinal:

$$\begin{cases} + & \text{se o valor é } \geq \text{o anterior} \\ - & \text{se o valor é } < \text{o anterior} \end{cases}$$

Seja R_0 igual ao número observado de iterações, isto é, grupos de mesmo sinal.

Para N grande, a variável aleatória R tem distribuição aproximadamente

$$N(\mu_R; \sigma_R)$$

onde

$$\mu_R = \frac{2N-1}{3}$$

$$\sigma_R = \sqrt{\frac{16N-29}{90}}$$

$$\text{Portanto, } z = \frac{R + \frac{1}{2} - \mu_R}{\sigma_R} : N(0;1) \text{ aproximadamente.}$$

$H_0: R_0 = \mu_R \Rightarrow$ não existe tendência na série

$H_a: R_0 < \mu_R \Rightarrow$ existe tendência na série

Rejeitamos H_0 se $z < z_c$ onde z_c é tal que

$$P(Z < Z_c / H_0) = \alpha ,$$

α igual ao nível de significância do teste.

EXEMPLO 3.2 - Na série do nível do lago Michigan-Huron, foram encontrados os seguintes valores:

$$N = 96$$

$$\left. \begin{array}{l} R_0 = 49 \\ \mu_R = 63,6 \\ \sigma_R = 4,1 \end{array} \right\} \Rightarrow z = \frac{49 + \frac{1}{2} - 63,6}{4,1} = -3,44$$

$$\alpha = 5\% \Rightarrow z_c = -1,64$$

Conclusão: $z < z_c \Rightarrow$ rejeita-se H_0 .

O número de iterações é menor que o esperado, no caso de independência.

3.3 - MÉTODOS DE ELIMINAÇÃO DA TENDÊNCIA

No que se segue, vamos considerar a série observada $Y(t)$ satisfazendo o modelo aditivo

$$(3.1) \quad Y(t) = m(t) + s(t) + \epsilon(t) \quad , \quad t=1,2,\dots,T$$

onde:

$m(t)$ = tendência, suposta uma função determinística regular do tempo.

$s(t)$ = componente sazonal, suposto estritamente periódico, com período igual a 1 ano.

$\epsilon(t)$ = distúrbio aleatório, que supomos gerado por um processo estacionário.

Para simplificar, vamos considerar o intervalo de um mês entre as observações, já que isso é o que ocorre geralmente na prática.

MÉTODO 1 - Média Móvel Simples Centrada (método da diferença)

Seja a série $Y(t)$ observada mensalmente durante $p+1$ anos.

O modelo (3.1) torna-se

$$(3.2) \quad Y(t) = m(t) + \alpha(i) + \epsilon(t) \text{ onde } \begin{cases} t=12j+i \\ j=0,1,\dots,p \\ i=1,2,\dots,12 \end{cases}$$

Os $\alpha(i)$ são constantes mensais e indicam desvios das médias mensais em relação à média total.

Isto é,

$$\alpha(i) = \bar{s}(i) - \bar{s} ,$$

onde

$$\bar{s}(i) = \frac{1}{p+1} \sum_{j=0}^p s(12j+i)$$

$$\begin{aligned}
 \bar{s} &= \frac{1}{12} \sum_{i=1}^{12} \bar{s}(i) \\
 &= \frac{1}{12} \sum_{i=1}^{12} \frac{1}{p+1} \sum_{j=0}^p s(12j+i) \\
 &= \frac{1}{12(p+1)} \sum_{t=1}^{12(p+1)} s(t) .
 \end{aligned}$$

Então:

$$\begin{aligned}
 \sum_{i=1}^{12} \alpha(i) &= \sum_{i=1}^{12} [\bar{s}(i) - \bar{s}] \\
 &= \sum_{i=1}^{12} \bar{s}(i) - \sum_{i=1}^{12} \bar{s} \\
 &= 12\bar{s} - 12\bar{s} .
 \end{aligned}$$

Logo,

$$(3.3) \quad \sum_{i=1}^{12} \alpha(i) = 0 .$$

Um estimador $\hat{m}(t)$ de média móvel simples centrada é obtido tomando-se as médias aritméticas de 12 observações sucessivas. Mas como 12 é número par, considera-se cada observação como a média de 2 observações sucessivas.

EXEMPLO 3.3

$$\begin{aligned}
 \hat{m}(7) &= \frac{1}{12} \left\{ \frac{Y(1)+Y(2)}{2} + \frac{Y(2)+Y(3)}{2} + \dots + \frac{Y(11)+Y(12)}{2} + \frac{Y(12)+Y(13)}{2} \right\} \\
 &= \frac{1}{24} \{ Y(1) + 2Y(2) + \dots + 2Y(12) + Y(13) \} .
 \end{aligned}$$

Logo, temos:

$$(3.4) \quad \hat{m}(t) = \frac{1}{24} \{Y(t-6) + 2Y(t-5) + \dots + 2Y(t+5) + Y(t+6)\}$$

$$t=7, 8, \dots, 12p+6$$

Essa estimativa, devido a (3.3), é livre de elementos sazonais.

Seja $X(t) = Y(t) - \hat{m}(t)$ o desvio a partir da tendência e sejam $\bar{X}(1), \bar{X}(2), \dots, \bar{X}(12)$ as médias mensais dos $X(t)$, isto é:

$$(3.5) \quad \bar{X}(i) = \begin{cases} \frac{1}{P} \sum_{j=1}^P X(12j+i) & i=1, \dots, 6 \\ \frac{1}{P} \sum_{j=0}^{P-1} X(12j+i) & i=7, \dots, 12. \end{cases}$$

Em virtude de (3.3), as constantes mensais $\alpha(i)$ são estimadas por

$$(3.6) \quad \alpha(i) = \bar{X}(i) - \bar{X} \text{ onde } \bar{X} = \frac{1}{12} \sum_{i=1}^{12} \bar{X}(i)$$

$$= \frac{1}{12} \sum_{t=7}^{12p+6} X(t)$$

Desenvolvendo (3.6) em função dos $Y(t)$,
temos:

$$(3.7) \left\{ \begin{array}{l} \bar{X}(1) = \bar{Y}(1) - \bar{Y} + \frac{1}{24p} \{Y(7) - Y(12p+7)\} \\ \bar{X}(i) = \bar{Y}(i) - \bar{Y} + \frac{1}{24p} \left\{ 2 \sum_{r=7}^{i+5} [Y(r) - Y(12p+r)] + Y(i+6) - Y(12p+i+6) \right\} \\ \bar{X}(i) = \bar{Y}(i) - \bar{Y} + \frac{1}{24p} \left\{ -Y(i-6) + Y(12p+i-6) + 2 \sum_{r=i-5}^6 [-Y(r) + Y(12p+r)] \right\} \\ \bar{X}(12) = \bar{Y}(12) - \bar{Y} + \frac{1}{24p} \left\{ -Y(6) + Y(12p+6) \right\} \end{array} \right. \quad \begin{array}{l} \\ i=2, \dots, 6 \\ i=7, \dots, 11 \\ \end{array}$$

onde

$$\bar{Y} = \frac{1}{12p} \sum_{t=7}^{12p+6} Y(t)$$

$$= \frac{1}{12} \sum_{i=1}^{12} \bar{Y}(i)$$

$$\bar{Y}(i) = \begin{cases} \frac{1}{p} \sum_{j=1}^p Y(12j+i) & i=1, \dots, 6 \\ \frac{1}{p} \sum_{j=0}^{p-1} Y(12j+i) & i=7, \dots, 12. \end{cases}$$

MÉTODO 2 - Média Móvel Simples Centrada (método da razão)

É análogo ao método 1, mas os resíduos $X(t)$ aqui são calculados através da razão e não da diferença.

O modelo suposto é: $Y(t) = m(t) \cdot s(t) \cdot \varepsilon(t)$.

Temos

$$Y(t), \quad t=1, \dots, 12(p+1)$$

$$\hat{m}(t) = \frac{1}{24} \left\{ Y(t-6) + 2Y(t-5) + \dots + 2Y(t+5) + Y(t+6) \right\}$$

Então

$$(3.8) \quad X(t) = Y(t)/\hat{m}(t)$$

$$\bar{X}(i) = \frac{1}{p} \sum_j X(12j+i)$$

$$\bar{X} = \frac{1}{12} \sum_{i=1}^{12} \bar{X}(i)$$

E as constantes mensais são estimadas por

$$\hat{\alpha}(i) = \bar{X}(i)/\bar{X}$$

MÉTODO 3 - Constantes Sazonais com Tendência no Tempo

Se o padrão sazonal muda lentamente com o tempo, muitas vezes se usa um método de análise que ajusta aos desvios uma regressão linear sobre o tempo, a partir da tendência das observações em cada mês.

Para isso, necessita-se do par de estatísticas:

$$\bar{X}(i)$$

e

$$\frac{1}{p} \sum_j jX(12j+i)$$

Já nos referimos à primeira estatística no Método 1; agora, vamos tratar da segunda.

Sejam

$$(3.9) \quad \begin{cases} b(i) = \begin{cases} \frac{1}{P} \sum_{j=1}^P jX(12j+i) & i=1, \dots, 6 \\ \frac{1}{P} \sum_{j=0}^{P-1} jX(12j+i) & i=7, \dots, 12 \end{cases} \\ s(i) = \begin{cases} \frac{1}{P} \sum_{j=1}^P jY(12j+i) & i=1, \dots, 6 \\ \frac{1}{P} \sum_{j=0}^{P-1} jY(12j+i) & i=7, \dots, 12 \end{cases} \end{cases}$$

Aplicando $X(t) = Y(t) - \hat{m}(t)$ como no Método 1, e calculando $b(i)$ através de $s(i)$, temos:

$$(3.10) \quad \begin{cases} b(1) = s(1) - \bar{s} - \frac{1}{24} \{ \bar{Y}(7) + 2\bar{Y}(8) + \dots + 2\bar{Y}(12) + Y(12p+7) \} \\ b(i) = s(i) - \bar{s} - \frac{1}{24} \{ \bar{Y}(i+6) + 2\bar{Y}(i+7) + \dots + 2\bar{Y}(12) + 2Y(12p+7) + \dots + 2Y(12p+i+6) \} \quad i=2, \dots, 5 \\ b(6) = s(6) - \bar{s} - \frac{1}{24} \{ \bar{Y}(12) + 2Y(12p+7) + \dots + 2Y(12p+11) + Y(12p+12) \} \\ b(7) = s(7) - \bar{s} + \frac{1}{24} \{ Y(12p+1) - 2Y(12p+12) - \dots - 2Y(12p+6) + \bar{Y}(1) \} \\ b(i) = s(i) - \bar{s} + \frac{1}{24} \{ 2Y(12p+i-6) + \dots + 2Y(12p+6) + 2\bar{Y}(1) + \dots + 2\bar{Y}(i-7) + \bar{Y}(i-6) \} \quad i=8, \dots, 11 \\ b(12) = s(12) - \bar{s} + \frac{1}{24} \{ Y(12p+6) + 2\bar{Y}(1) + \dots + 2\bar{Y}(5) + \bar{Y}(6) \} \end{cases}$$

E finalmente, calculamos

$$(3.11) \quad \hat{\alpha}(i) = b(i) - \bar{b}$$

onde

$$\bar{b} = \frac{1}{12} \sum_{i=1}^{12} b(i)$$

MÉTODO 4 - Hannan (1960)

O estimador da tendência $m(t)$ é

$$(3.12) \quad \hat{m}(t) = \sum_{s=-m}^m \gamma(s) Y(t+s)$$

onde $\gamma(-s) = \gamma(s)$, independente de t .

O resíduo $X(t) = Y(t) - \hat{m}(t)$ fica:

$$(3.13) \quad X(t) = \sum_{s=-m}^m \delta(s) Y(t+s)$$

onde

$$\delta(-s) = \delta(s) = \begin{cases} 1 - \gamma(0) & \text{se } s=0 \\ -\gamma(s) & \text{se } s=1, 2, \dots, m \end{cases}$$

Vamos supor, para simplificar, que $Y(t)$ é definida para valores $t = -m+1, \dots, 12p+m$ e $X(t)$ para valores $t=1, \dots, \dots, 12p$.

Sejam

$$\begin{cases} \alpha(1), \dots, \alpha(12) \text{ as constantes mensais de } Y(t) \\ \beta(1), \dots, \beta(12) \text{ as constantes mensais de } X(t). \end{cases}$$

Então, de (3.13), temos:

$$(3.14) \quad \beta(q) = \sum_{s=-m}^m \delta(s)\alpha(q+s) \quad q=1, \dots, 12$$

onde

$$(3.15) \quad \alpha(r+12k) = \alpha(r) \quad r=1, \dots, 12; k \in \mathbb{Z}$$

Devemos estimar $\alpha(r)$ através de (3.14).

Vamos ilustrar essas considerações através de um exemplo, para $m=1$. Os resultados se estendem para m qualquer.

EXEMPLO 3.3 -

$$(3.16) \quad \begin{aligned} \beta(q) &= \sum_{s=-1}^1 \delta(s)\alpha(q+s) \quad q=1, \dots, 12 \\ &= \delta(-1)\alpha(q-1) + \delta(0)\alpha(q) + \delta(1)\alpha(q+1) \end{aligned}$$

Logo:

$$(3.17) \quad \begin{cases} \beta(1) = \delta(-1)\alpha(0) + \delta(0)\alpha(1) + \delta(1)\alpha(2) \\ \beta(2) = \delta(-1)\alpha(1) + \delta(0)\alpha(2) + \delta(1)\alpha(3) \\ \vdots \\ \beta(11) = \delta(-1)\alpha(10) + \delta(0)\alpha(11) + \delta(1)\alpha(12) \\ \beta(12) = \delta(-1)\alpha(11) + \delta(0)\alpha(12) + \delta(1)\alpha(13) \end{cases}$$

Usando (3.15), temos:

$$\begin{cases} \alpha(0) = \alpha(12+12(0)) = \alpha(12) \\ \alpha(13) = \alpha(1+12(1)) = \alpha(1). \end{cases}$$

Colocando (3.16) em forma matricial, temos:

$$\underline{\beta} = \underline{\delta} \cdot \underline{\alpha}$$

ou

$$\begin{pmatrix} \beta(1) \\ \beta(2) \\ \vdots \\ \beta(11) \\ \beta(12) \end{pmatrix} = \begin{pmatrix} \delta(0) & \delta(1) & 0 & 0 & \dots & 0 & 0 & 0 & \delta(-1) \\ \delta(-1) & \delta(0) & \delta(1) & 0 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & \delta(-1) & \delta(0) & \delta(1) \\ \delta(1) & 0 & 0 & 0 & \dots & 0 & 0 & \delta(-1) & \delta(0) \end{pmatrix} \cdot \begin{pmatrix} \alpha(1) \\ \alpha(2) \\ \vdots \\ \alpha(11) \\ \alpha(12) \end{pmatrix}$$

onde $\delta(-1) = \delta(1) \Rightarrow \underline{\delta}$ é simétrica.

Segundo Brillinger (1974), a matriz $\underline{\delta}$ assim formada chama-se "matriz circulante" e sua inversa é calculada imediatamente.

Para qualquer m, podemos calcular

$$\underline{\alpha} = \underline{\delta}^{-1} \underline{\beta}$$

através de

$$(3.18) \quad \hat{\alpha}(r) = \frac{1}{12} \sum_{q=1}^{12} \hat{\beta}(q) \sum_{s=1}^{11} \frac{\exp\left[\frac{11is(r-q)}{6}\right]}{\delta\left[\exp\left(\frac{11is}{6}\right)\right]} \quad r=1, \dots, 12$$

onde

$$\delta[z] = \sum_{s=-m}^m \delta(s) z^s$$

MÉTODO 5 - Hannan (1960)

É análogo ao método 4, mas usa somente os valores extremos da série.

Seja a série $Y(t)$, $t=-m+1, \dots, 12p+m$.

E seja $Y'(t)$ outra série tal que

$$(3.19) \quad Y'(t) = \begin{cases} Y(t) & t=1, \dots, 12p \\ Y(t+12p) & t=-m+1, \dots, 0 \\ Y(t-12p) & t=12p+1, \dots, 12p+m \end{cases}$$

Vamos calcular $Y''(t) = Y(t) - Y'(t)$, tal que $Y''(t) = 0$ para $t=1, 2, \dots, 12p$.

Aplicando (3.13) em $Y'(t)$ e $Y''(t)$, obtemos respectivamente os resíduos $X'(t)$ e $X''(t)$.

Sejam

$\alpha'(r)$ as constantes mensais de $Y'(t)$,

$\alpha''(r)$ as constantes mensais de $Y''(t)$,

$\beta'(r)$ as constantes mensais de $X'(t)$,

$\beta''(r)$ as constantes mensais de $X''(t)$.

Sejam ainda

$$\bar{Y}(r) = \frac{1}{p} \sum_{j=0}^{p-1} Y(12j+r),$$

$$\bar{Y} = \frac{1}{12p} \sum_{t=1}^{12p} Y(t),$$

com definições análogas para $X(t)$, $X'(t)$, $X''(t)$.

Queremos estimar $\alpha(r)$ através de $\hat{\alpha}'(r)$ e $\hat{\alpha}''(r)$.

Dada a circularidade de $Y'(t)$, temos de (3.13):

$$(3.20) \quad \bar{X}'(q) = \sum_{s=-m}^m \delta(s) \bar{Y}(q+s) \quad q=1, \dots, 12,$$

onde $\bar{Y}(r+12k) = \bar{Y}(r)$, $r=1, \dots, 12$; $k \in \mathbb{Z}$.

Fazendo $\hat{\beta}'(r) = \bar{X}'(r) - \bar{X}'$ e comparando (3.20) com (3.16), vemos que

$$\hat{\alpha}'(r) = \bar{Y}(r) - \bar{Y}.$$

Fazendo $\hat{\beta}''(q) = \bar{X}''(q) - \bar{X}''$ e usando (3.18), temos

$$(3.21) \quad \hat{\alpha}''(r) = A''(r) - \frac{1}{12} \sum_{s=1}^{12} A''(s),$$

onde

$$(3.22) \quad A''(r) = \frac{1}{12} \sum_{q=1}^{12} \bar{X}''(q) \sum_{s=1}^{11} \frac{\exp\left[\frac{\pi i s(r-q)}{6}\right]}{\delta\left[\exp\left(\frac{\pi i s}{6}\right)\right]}, \quad r=1, \dots, 12.$$

E o estimador de $\alpha(r)$ é:

$$(3.23) \quad \hat{\alpha}(r) = A(r) - \frac{1}{12} \sum_{s=1}^{12} A(s),$$

onde

$$(3.24) \quad A(r) = \bar{Y}(r) - A''(r).$$

MÉTODO 6 - Diferença de Variáveis

A função

$$(3.25) \quad \Delta Y(t) = Y(t+1) - Y(t)$$

é chamada "diferença de primeira ordem" das variáveis $Y(t)$, $t=1, \dots, T$.

A "diferença de segunda ordem" é definida como segue:

$$\begin{aligned} \Delta^2 Y(t) &= \Delta[\Delta Y(t)] \\ &= \Delta[Y(t+1) - Y(t)] \\ &= \Delta Y(t+1) - \Delta Y(t) \\ (3.26) \quad &= Y(t+2) - 2Y(t+1) + Y(t) \end{aligned}$$

Podemos, por recorrência, calcular a diferença de p -ésima ordem:

$$(3.27) \quad \Delta^p Y(t) = \sum_{k=0}^p \binom{p}{k} (-1)^k Y(t+p-k)$$

Se fizermos como nos métodos anteriores,

$$(3.28) \quad X(t) = Y(t) - \Delta^p Y(t)$$

vemos que:

$$\begin{cases} \Delta^p Y(t) \text{ é um caso particular de (3.12)} & e \\ X(t) \text{ é um caso particular de (3.13).} \end{cases}$$

MÉTODO 7 - Rhodes (1921) e Quenouille (1949)

Separa-se a série em segmentos e ajusta-se um polinômio de cada um.

São necessários muitos cálculos para assegurar que cada polinômio tenha ligação regular com o seguinte.

Este método tem sido muito pouco aplicado.

CAPÍTULO IV

ELIMINAÇÃO DO COMPONENTE SAZONAL

4.1 - INTRODUÇÃO

Quando examinamos séries temporais observadas em intervalos mensais, trimestrais ou semanais, notamos movimentos intra-anuais que chamamos de movimentos sazonais. (Ver no apêndice, as séries 2, 3, 4, 5 e 6)

DEFINIÇÃO 4.1: Sazonalidade, no domínio do tempo, segundo Thomas e Wallis (1971).

"Por variação sazonal, nós entendemos os movimentos intra-anuais sistemáticos, mas não necessariamente regulares, da série temporal econômica, que são causados muitas vezes por fenômenos não econômicos, tais como mudanças de clima, festas religiosas, etc."

DEFINIÇÃO 4.2 - Sazonalidade no domínio de frequência, segundo Nerlove (1964).

"Sazonalidade é aquela característica da série tem-

poral que provoca picos no espectro, nas frequências sazonais."

Vamos esclarecer o conceito através de um exemplo simples e irreai.

EXEMPLO 4.1 - Suponhamos a série $Y(t)$ perfeitamente periódica com período igual a 1 ano.

Ela pode ser representada por uma soma finita de senos e cosenos, com amplitudes aleatórias.

Tomando observações mensais, isto é, $1/12$ do período fundamental, são suficientes 11 termos. Temos então:

$$(4.1) \quad Y(t) = \sum_{i=1}^6 \left\{ \alpha_i \cos 2\pi \lambda_i t + \beta_i \sin 2\pi \lambda_i t \right\}$$

onde $\lambda_i = i/12$ são as frequências sazonais e $\sin[2\pi 6/12]t = 0 \quad \forall t$.

Sejam $\alpha_1, \dots, \alpha_6$ e $\beta_1, \dots, \beta_6$ variáveis aleatórias independentes, com

$$E\{\alpha_i\} = E\{\beta_i\} = 0$$

$$E\{\alpha_i^2\} = E\{\beta_i^2\} = \sigma_i^2$$

Então, como vimos no Capítulo II, o espectro de $Y(t)$ é composto de 6 linhas, nas frequências sazonais.

E a função de distribuição espectral consiste de uma função em escada, com "saltos" nos λ_i de tamanhos iguais a σ_i^2 .

Figura 1

Figura 2

Em séries reais, o espectro não consiste de linhas espectrais, mas de picos mais ou menos largos. Quanto mais estreito o pico, mais pronunciada e regular será a variação sazonal. (Ver Apêndice 2).

DEFINIÇÃO 4.3 - O ajustamento sazonal consiste em filtrar a potência alta das frequências sazonais.

O problema que surge é encontrar um filtro que remova os picos espectrais, mas que altere, o mínimo possível, as características resultantes da série, já que estas muitas vezes fornecem uma definição "a posteriori" do componente sazonal, através das relações entre a série original e a que foi sazonalmente ajustada.

A escolha do método de estimação e eliminação do componente sazonal depende da finalidade da análise. Vamos apresentar, aqui, os métodos mais usados.

4.2 - MÉTODO 1: REGRESSÃO, APRESENTADO POR KENDALL (1973)

Este é um método pouco usado.

Sejam Y_{ij} os valores de uma série para o ano i e o mês j . E sejam X_{ij} os valores da mesma série, livre de tendência.

Calculamos as retas de regressão

$$(4.2) \quad Y_{ij} = a_j + b_j X_{ij} \quad j=1, \dots, 12$$

Se $a_j = 0$, temos o modelo multiplicativo, e se $b_j = 1$, temos o modelo aditivo.

Ajustamos as constantes $\underline{a}$ e $\underline{b}$ para

$$(4.3) \quad \begin{cases} a_j^! = a_j - \bar{a}, \text{ onde } \bar{a} = \frac{1}{12} \sum_{j=1}^{12} a_j \\ b_j^! = b_j - \bar{b}, \text{ onde } \bar{b} = \frac{1}{12} \sum_{j=1}^{12} b_j \end{cases}$$

E os valores sazonalmente ajustados são dados por

$$(4.4) \quad y_{ij}^* = \frac{y_{ij} - a_j}{b_j} .$$

As desvantagens desse método são as seguintes:

- 1a.) são necessárias as estimativas de 24 constantes e, a menos que a série seja longa, as estimativas não são fidedignas.
- 2a.) as constantes a' e b' são estimadas para a série inteira e não levam em conta, portanto, as mudanças que podem ocorrer no padrão sazonal.

4.3 - MÉTODO 2: REGRESSÃO POR MÍNIMOS QUADRADOS, APRESENTADO POR JORGENSON (1964)

A suposição inicial é que o componente aleatório $\epsilon(t)$ seja estacionário e livre da correlação serial, que foi definida no Capítulo I.

O modelo da série deve satisfazer as hipóteses de um modelo estatístico linear geral e, portanto, as propriedades dos estimadores de $m(t)$ e $s(t)$ devem ser obtidas da teoria estatística existente para este modelo.

Além disso, os métodos de estimação devem ser obtidos diretamente da teoria geral de Análise de Regressão para séries estacionárias, como foi desenvolvido por Parzen (1961).

O problema que existe é que, na prática, não se tem

usado critérios baseados na inferência estatística, e a proposição deste método é anular a lacuna que separa a teoria da prática.

Seja o modelo aditivo

$$(4.5) \quad Y_i = m_i + s_i + \epsilon_i \quad i=1, \dots, N$$

Colocando em forma matricial, temos

$$\underline{Y} = \underline{m} + \underline{s} + \underline{\epsilon}$$

isto é:

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_N \end{pmatrix} = \begin{pmatrix} m_1 \\ \vdots \\ m_N \end{pmatrix} + \begin{pmatrix} s_1 \\ \vdots \\ s_N \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_N \end{pmatrix}$$

Vamos especificar cada um dos componentes, fazendo suposições sobre suas distribuições.

$$a) \quad \underline{\epsilon} \text{ é Normal, com } \begin{cases} E\{\underline{\epsilon}\} = 0 \\ E\{\epsilon_i \epsilon_j\} = \begin{cases} \omega I & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases} \end{cases}$$

onde ω é parâmetro constante, que representa a variância de cada elemento de $\underline{\epsilon}$.

b) $\underline{s} = \underline{S} \underline{g}$ onde $\underline{S}$ é a matriz cujas colunas representam as variáveis sazonais, e $\underline{g}$ é o vetor dos parâmetros.

$$(i) \quad \underline{\hat{\sigma}} \text{ é Normal, com } \begin{cases} E\{\underline{\hat{\sigma}}\} = \underline{\sigma} \\ V\{\underline{\hat{\sigma}}\} = \omega\{S'[I-M(M'M)^{-1}M']S\}^{-1} \end{cases}$$

(ii) $\underline{s}$ é determinístico e, portanto, as colunas de $\underline{s}$ são fixas, e não aleatórias.

(iii) $\underline{S}$ é de posto completo, e seja K_S igual ao posto de $\underline{S}$.

$\sigma_j, j=1, \dots, K_S$, são os efeitos sazonais, cada um tratado como parâmetro desconhecido; associado a cada um, existe uma variável sazonal fixada.

$s^j, j=1, \dots, K_S$, é o vetor de observações da j^{a} variável sazonal.

$\{s_{ij}\}$, elementos s^j , são unitários quando $i=j+r \cdot K_S^*$, onde $r \in \mathbb{Z}$ e K_S^* é o número de estações de K_S .

Cada estação tem efeito constante através do período amostral.

EXEMPLO 4.1 - Para o caso de quatro estações, temos:

$$s^1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix} ; s^2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix} ; s^3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix} ; s^4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ \vdots \\ \vdots \\ 1 \end{pmatrix}$$

logo:

$$\underline{s} = \underline{S}\underline{\sigma} = \sum_{j=1}^{K_S} \sigma_j \underline{s}^j$$

c) $\underline{m} = \underline{M}\underline{\mu}$

(i) $\hat{\underline{\mu}}$ é Normal, com
$$\begin{cases} E(\hat{\underline{\mu}}) = \underline{\mu} \\ V(\hat{\underline{\mu}}) = \omega \{ \underline{M}' [\underline{I} - \underline{S}(\underline{S}'\underline{S})^{-1} \underline{S}'] \underline{M} \}^{-1} \end{cases}$$

(ii) $\underline{M}$ é de posto completo, e seja K_m igual ao posto de $\underline{M}$.

$$\underline{m} = \underline{M}\underline{\mu} = \sum_{j=1}^{K_m} \mu_j \underline{m}^j$$

Dadas as suposições acima, o modelo fica

(4.6)
$$\underline{Y} = \underline{M}\underline{\mu} + \underline{S}\underline{\sigma} + \underline{\varepsilon}$$

Seja a matriz $\begin{bmatrix} \underline{M} \\ \underline{S} \end{bmatrix}$, que deve ser especificada, para que se possa identificar os parâmetros desconhecidos $\{\underline{\mu}; \underline{\sigma}\}$. $\underline{M}$ e $\underline{S}$ são de posto completo, portando $\begin{bmatrix} \underline{M} \\ \underline{S} \end{bmatrix}$ também é; se não fôr, deve-se transformá-la numa matriz de posto-completo, eliminando-se um dos parâmetros desconhecidos.

Existem muitas formas de transformação (Ver Jorgenson, 1964, pp. 693).

$\underline{M}$ e $\underline{S}$ devem ser não aleatórias, e $\begin{bmatrix} \underline{M} \\ \underline{S} \end{bmatrix}$ também.

Para estimar os parâmetros, lembremos que, se

$$\underline{Y} = \underline{X}\underline{\beta} + \underline{\varepsilon},$$

então o estimador de mínimos quadrados de $\underline{\beta}$ é

$$\hat{\underline{\beta}} = (\underline{X}'\underline{X})^{-1} \underline{X}'\underline{Y}$$

Então, temos:

$$(4.7) \quad \begin{bmatrix} \mu \\ \sigma \end{bmatrix} = \begin{bmatrix} M'M & M'S \\ S'M & S'S \end{bmatrix}^{-1} \cdot \begin{bmatrix} M'y \\ S'y \end{bmatrix}$$

ou

$$(4.8) \quad \hat{\mu} = (M'M)^{-1}M'(I-S\{S'[I-M(M'M)^{-1}M']S\}^{-1}S'[I-M(M'M)^{-1}M'])y$$

$$(4.9) \quad \hat{\sigma} = \{S'[I-M(M'M)^{-1}M']S\}^{-1}S'[I-M(M'M)^{-1}M']y$$

que são estimadores de mínimos quadrados ordinários, com matriz de covariância

$$(4.10) \quad \omega \cdot \begin{bmatrix} M'M & M'S \\ S'M & S'S \end{bmatrix}^{-1}.$$

O resíduo da regressão ajustada é

$$(4.11) \quad \hat{\varepsilon} = y - M\hat{\mu} - S\hat{\sigma}.$$

E a estimativa de ω é

$$(4.12) \quad \hat{\omega} = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{N-(K_m+K_s)}.$$

Vamos agora tratar do ajustamento sazonal, e vamos nos restringir ao ajustamento linear, através da transformação linear B, tal que

$$(4.13) \quad \underline{Y}_S^B = B \cdot \underline{Y}$$

No modelo (4.6), aplicaríamos a transformação e ob-

teríamos

$$\underline{Y}_S = \underline{Y} - S\underline{\sigma} = \underline{M}\underline{\mu} + \underline{\varepsilon} \quad ,$$

se os parâmetros fossem conhecidos.

Como são desconhecidos, usamos $\hat{\underline{\sigma}}$ como em (4.9), que é o único estimador linear, não viciado, de variância mínima.

$$\begin{aligned} \hat{\underline{Y}}_S &= \underline{Y} - S\hat{\underline{\sigma}} \\ &= B \cdot \underline{Y} \\ (4.14) \quad &= [I - S\{S'T_M S\}^{-1} S'T_M] \cdot \underline{Y} \quad , \end{aligned}$$

onde T_M é a matriz idempotente

$$I - M(M'M)^{-1}M' \quad .$$

Algumas propriedades dos métodos lineares de ajustamento sazonal são:

P_1 : A série sazonalmente ajustada é aditiva, isto é

$$\begin{aligned} \underline{Y}_S^B &= B\underline{Y} \\ &= B(\underline{M}\underline{\mu} + S\underline{\sigma} + \underline{\varepsilon}) \\ (4.15) \quad &= B\underline{M}\underline{\mu} + BS\underline{\sigma} + B\underline{\varepsilon} \quad . \end{aligned}$$

P_2 : Qualquer método deve ser caracterizado pela distribuição da diferença entre a série ajustada por ele e a série que seria obtida se os parâmetros do componente sazonal fossem conhecidos, isto

to é,

$$\begin{aligned} \underline{Y}_S^B - \underline{Y}_S &= [B\underline{M}_\mu + B\underline{S}\underline{\sigma} + B\underline{\epsilon}] - [\underline{M}_\mu + \underline{\epsilon}] \\ (4.16) \quad &= [B-I]\underline{M}_\mu + B\underline{S}\underline{\sigma} + [B-I]\underline{\epsilon} \end{aligned}$$

P_3 : Um método linear é não viciado se o valor esperado de (4.16) é igual a 0, isto é

$$(4.17) \quad E\{\underline{Y}_S^B - \underline{Y}_S\} = [B-I]\underline{M}_\mu + B\underline{S}\underline{\sigma} = 0 \text{ tal que } \begin{cases} [B-I]\underline{M} = 0 \\ B\underline{S} = 0 \end{cases}$$

P_4 : Um método linear não viciado deixa inalterado o componente de tendência e anula o componente sazonal.

No caso (4.14), o método é não viciado se

$$(4.18) \quad \begin{cases} -\underline{S}[\underline{S}'\underline{T}_M\underline{S}]^{-1}\underline{S}'\underline{T}_M\underline{M} = 0 \\ [\underline{I} - \underline{S}(\underline{S}'\underline{T}_M\underline{S})^{-1}\underline{S}'\underline{T}_M]\underline{S} = 0 \end{cases}$$

P_5 : Um método linear não viciado é de variância mínima, se fôr mínima a variância de uma combinação linear arbitrária da diferença de (4.16).

Seja $\underline{a}'$ um vetor arbitrário de coeficientes fixados. A combinação linear é

$$\underline{a}'(\underline{Y}_S^B - \underline{Y}_S),$$

cujas variância é

$$\begin{aligned}
 V[\underline{a}'(\underline{Y}_S^B - \underline{Y}_S)] &= V[\underline{a}'(B-I)\underline{\epsilon}] \\
 (4.19) \qquad \qquad \qquad &= \omega \underline{a}'(B-I)(B-I)'\underline{a}
 \end{aligned}$$

Para ajustamento sazonal, (4.19) deve ser escrita na forma

$$\begin{aligned}
 &\omega \underline{a}'[S\{S'T_M S\}^{-1}S' + (I - S\{S'T_M S\}^{-1}S'T_M - B)(\dots)']\underline{a} \\
 (4.20) \qquad &= \omega \underline{a}'[S\{S'T_M S\}^{-1}S' + S\{S'T_M S\}^{-1}S' + (B-I)T_M S\{S'T_M S\}^{-1}S' + \\
 &+ S\{S'T_M S\}^{-1}S'T_M(B-I)' + (B-I)(B-I)']\underline{a},
 \end{aligned}$$

onde

$$(4.21) \qquad (B-I)T_M S\{S'T_M S\}^{-1}S' = -S\{S'T_M S\}^{-1}S',$$

desde que

$$(4.22) \qquad \begin{cases} (B-I)T_M = (B-I)[I - M(M'M)^{-1}M'] = B-I \\ (B-I)S = -S \end{cases}$$

A expressão (4.20) é a soma de duas formas quadráticas semi-positivas definidas, e a primeira não depende da transformação B.

O mínimo de (4.20) é atingido quando se toma B como em (4.14).

∀ $\underline{a}'$, essa é a única B possível, para que a variância seja mínima.

4.4 - MÉTODO 3 - ANÁLISE HARMÔNICA, apresentado por Kendall
(1973)

O tratamento dos efeitos sazonais por análise harmônica ilustra o uso da função de transferência para medir o efeito da média móvel sobre os diferentes componentes da série.

Para uma média móvel de $2m+1$ termos, que podemos representar convenientemente com pesos simétricos

$$a(-k) = a(k),$$

a função de transferência se reduz à forma real

$$(4.23) \quad A(\lambda) = a(0) + \sum_{k=1}^m a(k) \cos \lambda k \quad 0 \leq \lambda \leq \pi$$

Usando essa média, estamos multiplicando as ordenadas do espectro pelo módulo de $A(\lambda)$.

Dados os pesos $a(k)$, podemos calcular $A(\lambda)$ para um conjunto de valores de 0 a π e, então, determinar o efeito da série filtrada.

O efeito geral consiste em remover, com a tendência, alguns ou todos os ciclos longos (de frequência baixa); com um eliminador conveniente de tendência, podemos determinar o componente sazonal, seguros de que as frequências mais altas ficam afetadas fracamente.

Se a sazonalidade for estritamente periódica, pode-

mos representá-la por uma soma de harmônicas, com frequências regulares

$$\lambda_s = 2\pi s/12$$

e comprimentos de onda de 12; 6; 4; 3; 2,4 e 2 meses, no caso de séries mensais. Ver o quadro abaixo.

FREQÜÊNCIA (CICLOS/MÊS)	PERÍODO (MESES)	NÚMERO DE VEZES EM QUE SE COMPLETA UM CICLO
1/12 = 0,0833	12	1
2/12 = 0,1667	6	2
3/12 = 0,2500	4	3
4/12 = 0,3333	3	4
5/12 = 0,4167	2,4	5
6/12 = 0,5000	2	6

Se $X(t)$ é o desvio, a partir da tendência, para o t^o mês em um conjunto de $12p$ meses, temos:

$$(4.24) \left\{ \begin{array}{l} a(k) = \frac{1}{12p} \sum_{t=1}^{12p} X(t) \cos \lambda(k)t \\ b(k) = \frac{1}{12p} \sum_{t=1}^{12p} X(t) \sin \lambda(k)t \end{array} \right\} \quad k=1, \dots, 5$$

$$a(6) = \frac{1}{12p} \sum_{t=1}^{12p} X(t) \sin \lambda(6)t$$

O movimento sazonal será representado por 11 constantes.

Agora podemos achar o espectro da série ajustada sazonalmente, multiplicando-a por $1/A(\lambda)$.

4.5 - MÉTODO 4: DO BUREAU DO CENSO DOS ESTADOS UNIDOS

É conhecido como Census Mark II e foi idealizado por Shiskin (1967).

Seu procedimento básico, para dados mensais, é o seguinte, supondo o modelo multiplicativo

$$(4.25) \quad Y(t) = m(t) \cdot s(t) \cdot \epsilon(t) \cdot$$

1º) Calcula-se uma média móvel de 12 termos para achar $\hat{m}(t)$.

2º) Calcula-se a razão $Y(t)/\hat{m}(t)$ para se obter

$$Y'(t) = s(t) \cdot \epsilon(t) \cdot$$

3º) A cada mês de $Y'(t)$, aplica-se uma média móvel de 5 termos para achar $\hat{s}(t)$.

4º) Calcula-se a razão $Y'(t)/\hat{s}(t)$ para se obter $\hat{\epsilon}(t)$.

5º) Para cada ano, calcula-se σ , desvio padrão móvel de 5 anos.

6º) Corrige-se cada valor de $\hat{\epsilon}(t)$, do seguinte modo: atribui-se

peso igual a 1, se $\hat{\epsilon}(t) < 1,5\sigma$

peso igual a 0, se $\hat{\epsilon}(t) > 2,5\sigma$

peso entre 0 e 1, se $1,5\sigma \leq \hat{\epsilon}(t) \leq 2,5\sigma$.

7º) substitui-se, em $Y'(t)$, o valor $\varepsilon(t)$ por $\hat{\varepsilon}(t)$ corrigido, isto é:

$$(4.27) \quad Y''(t) = \hat{s}(t) \cdot \hat{\varepsilon}(t) .$$

8º) A cada mês de $Y''(t)$, aplica-se uma média de 5 termos, para se obter novos fatores sazonais $\hat{s}'(t)$.

9º) Calcula-se a razão $Y(t)/\hat{s}'(t)$ e se obtém, finalmente, a série livre de sazonalidade.

Neste método, se o modelo for aditivo, substituímos as razões por diferenças.

4.6 - MÉTODO 5: DO BUREAU OF LABOR STATISTICS DOS ESTADOS UNIDOS (BLS), (1960)

É empregado para ajustar séries mensais de emprego e desemprego, expedidas pelo próprio Bureau.

O procedimento básico é semelhante ao do Bureau do Censo.

Emprego, desemprego e força de trabalho, estimados por levantamento amostral, consistem de grande número de componentes.

São feitas separações com base em idade, sexo, indústria agrícola ou não e, mais recentemente, cor.

No caso de desemprego, se faz uma classificação com base em sua duração.

No caso de séries de emprego total, existem quatro

categorias sexo-idade:

- Homens com idade entre 14 e 19 anos;
- Homens com idade a partir de 20 anos;
- Mulheres com idade entre 14 e 19 anos;
- Mulheres com idade a partir de 20 anos.

Cada categoria sofre ajustamento sazonal, e a soma das quatro vai fornecer a série total ajustada.

4.7 - MÉTODO 6: MÉTODO RESIDUAL

Aplica o Método 5, por subtração, isto é:

- 1º) Séries de emprego total e de força de trabalho são ajustadas pelo BLS.
- 2º) Subtraindo da série original o resultado acima, obtém-se a série ajustada de desemprego.

CAPÍTULO V

ANÁLISE DOS EFEITOS DOS PROCESSOS DE

AJUSTAMENTO SAZONAL

Essa análise foi apresentada por Nerlove (1964) e pretende esclarecer os métodos e vantagens do uso da análise espectral e do espectro cruzado, quando se faz ajustamento sazonal de séries temporais.

São recentes as tentativas de remoção dos efeitos sazonais em séries temporais, e nenhuma provou até agora ser estatisticamente adequada.

Além disso, todos os métodos utilizam filtros que são não lineares e, portanto, é difícil especificar, a priori, seus efeitos sobre a série.

Efeitos possíveis incluem

- a) remoção de mais elementos além dos que são estritamente sazonais;
- b) introdução de movimentos não sazonais;
- c) distorção das relações entre a série ajustada sazonal

mente e a série original.

Para constatar os efeitos acima, comparam-se os espectros entre a série original e a ajustada e observam-se o ganho e a fase do filtro aplicado.

As afirmações acima se baseiam simplesmente na observação visual dos espectros das duas séries, mas, para uma análise mais profunda, deve-se usar a técnica conhecida por análise de espectro-cruzado do qual já falamos no capítulo II.

Já vimos o que é filtro, função de transferência, ganho e ângulo de fase, no capítulo II, e vamos aplicar esses conceitos na análise de uma série que será ajustada por um filtro linear, invariante no tempo.

Seja a representação (2.13) de uma série temporal na forma complexa

$$Y(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dZ_Y(\lambda),$$

onde

$$dZ(\lambda) = \frac{1}{2}\{dU(\lambda) - idV(\lambda)\}.$$

Vamos aplicar em $Y(t)$ o filtro linear (2.15).

$$X(t) = L \cdot Y(t) = \int_{-\infty}^{\infty} a(u)Y(t-u)du,$$

onde o peso $a(u)$ é a função de impulso do filtro L , que opera sobre $Y(t)$ para produzir $X(t)$.

Ainda por (2.13), temos:

$$Y(t-u) = \int_{-\infty}^{\infty} e^{i\lambda(t-u)} dZ_Y(\lambda) .$$

Então $X(t)$ fica

$$\begin{aligned} X(t) &= \int_{-\infty}^{\infty} a(u) \left[\int_{-\infty}^{\infty} e^{i\lambda(t-u)} dZ_Y(\lambda) \right] du \\ &= \int_{-\infty}^{\infty} e^{i\lambda t} \left[\int_{-\infty}^{\infty} a(u) e^{-i\lambda u} du \right] dZ_Y(\lambda) . \end{aligned}$$

De (2.16), temos

$$A(\lambda) = \int_{-\infty}^{\infty} a(u) e^{-i\lambda u} du ,$$

que é a função resposta de frequência, ou função de transferência do filtro. Logo:

$$(5.1) \quad X(t) = \int_{-\infty}^{\infty} e^{i\lambda t} A(\lambda) dZ_Y(\lambda)$$

mas, de (2.13), vem

$$(5.2) \quad X(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dZ_X(\lambda)$$

Comparando (5.1) e (5.2), temos, como em (2.18):

$$(5.3) \quad dZ_X(\lambda) = A(\lambda) dZ_Y(\lambda)$$

Multiplicando os dois termos de (5.3) pelo complexo conjugado de $dZ_Y(\lambda)$, e tomando valores esperados, temos:

$$(5.4) \quad A(\lambda) = \frac{E\{dZ_X(\lambda) \overline{dZ_Y(\lambda)}\}}{E\{dZ_Y(\lambda) \overline{dZ_Y(\lambda)}\}}$$

Logo $A(\lambda)$ representa a inclinação da reta de regressão entre as variáveis aleatórias complexas $dZ_X(\lambda)$ e $dZ_Y(\lambda)$.

A interpretação dos valores da função de transferência como um conjunto de inclinações de regressão, uma para cada frequência na representação das séries de entrada e saída de um filtro, sugere uma aplicação interessante.

Em trabalho estatístico comum sobre séries temporais, nós frequentemente relacionamos duas variáveis através de uma reta de regressão de mínimos quadrados, mesmo quando sabemos que a relação entre elas não é verdadeiramente linear; a idéia é obter alguma medida da "relação média" entre as variáveis.

Aqui, de modo análogo, mesmo sabendo que o filtro não é linear nem invariante no tempo, podemos relacionar as séries de entrada e saída do sistema através de (5.4).

Então, a inclinação de regressão, obtida em cada frequência λ , representa a função resposta de frequência "média".

Em outras palavras, tomando uma série amostral, aplicando "qualquer" filtro, obtendo a série ajustada e fazendo uma estimativa da quantidade (5.4), estaremos estimando a função resposta de frequência de um filtro linear invariante no tempo, que é uma aproximação ao filtro verdadeiro, válida sobre a extensão das amplitudes aleatórias que caracterizam a

série original da forma como ela ocorre na prática.

Logo, os resultados fornecem meios de caracterizar qualquer filtro em termos do ganho e do ângulo de fase, nas frequências para as quais as estimativas são feitas.

Mas o problema que surge é: como formar estimativas das quantidades de (5.4)?

Quanto ao denominador, temos de (2.12):

$$E\{dU(\lambda)\}^2 = dF(\lambda) = E\{dV(\lambda)\}^2$$

e, de (2.13), temos

$$dZ(\lambda) = \frac{1}{2}\{dU(\lambda) - i dV(\lambda)\} .$$

Logo,

$$E\{dZ(\lambda) \overline{dZ(\lambda)}\} = \frac{1}{2} dF(\lambda) .$$

Portanto, o denominador de (5.4) é

$$\frac{1}{2} dF_{YY}(\lambda) .$$

Quanto ao numerador, vimos em (2.21) que

$$dF_{XY}(\lambda) = dC_{XY}(\lambda) - i dQ_{XY}(\lambda) ,$$

onde

$$\begin{cases} dC_{XY}(\lambda) = \text{Co-espectro} \\ dQ_{XY}(\lambda) = \text{espectro-quadratura} \end{cases}$$

ou

$$dF_{XY}(\lambda) = 2E\{dZ_X(\lambda)\overline{dZ_Y(\lambda)}\} \quad .$$

Logo

$$E\{dZ_X(\lambda)\overline{dZ_Y(\lambda)}\} = \frac{1}{2} dF_{XY}(\lambda) \quad .$$

PROPRIEDADE: O espectro cruzado e a covariância cruzada são pares de Fourier.

De fato:

$$\begin{aligned} E\{X(t)Y(t-u)\} &= E\{X(t)\overline{Y(t-u)}\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(\lambda-\lambda')t} e^{i\lambda'u} E\{dZ_X(\lambda)\overline{dZ_Y(\lambda')}\} \\ &= \int_{-\infty}^{\infty} e^{i\lambda u} \frac{dF_{XY}(\lambda)}{2} \end{aligned}$$

porque, como já vimos,

$$E\{dZ_X(\lambda)\overline{dZ_Y(\lambda')}\} = \begin{cases} 0 & \text{se } \lambda \neq \lambda' \\ \frac{1}{2} dF_{XY} & \text{se } \lambda = \lambda' \end{cases}$$

E mais,

$$(5.5) \quad dF_{XY}(\lambda) = \frac{1}{\pi} \int_{-\infty}^{\infty} e^{i\lambda u} [E\{X(t)Y(t-u)\}] du \quad .$$

Esta fórmula nos habilita a obter uma forma de es-

timar as partes real e complexa do espectro-cruzado.

A função F_{XY} , pode ser decomposta, na forma clássica, em termos de três funções:

F_1 = absolutamente contínua

F_2 = escada

F_3 = singular

Se considerarmos que F_2 e F_3 são irrelevantes, no caso de séries temporais, podemos escrever

$$(5.6) \quad \begin{aligned} dF_{XY}(\lambda) &= f_{XY}(\lambda) d\lambda \\ &= c_{XY}(\lambda) d\lambda - i q_{XY}(\lambda) d\lambda \end{aligned}$$

A fórmula discreta, análoga à (5.5), para um número M finito de pontos ("lags") é:

$$(5.7) \quad \begin{aligned} f_{XY}(\lambda) &= c_{XY}(\lambda) - i q_{XY}(\lambda) \\ &= \frac{1}{\pi} \sum_{u=-M}^M e^{-i\lambda u} E\{X(t)Y(t-u)\} \\ &= \frac{1}{\pi} \sum_{u=-M}^M (\cos \lambda u - i \sin \lambda u) E\{X(t)Y(t-u)\} \\ &= \frac{1}{\pi} \left[E\{X(t)Y(t)\} + \sum_{u=1}^M \cos \lambda u E\{X(t)Y(t-u) + X(t)Y(t+u)\} - \right. \\ &\quad \left. - i \sum_{u=1}^M \sin \lambda u E\{X(t)Y(t-u) - X(t)Y(t+u)\} \right] \end{aligned}$$

Sejam $R_{XY}(u)$ e $S_{XY}(u)$ as estimativas amostrais dos

valores esperados de (5.7), isto é:

$$(5.8) \quad \begin{cases} R_{XY}(u) = \frac{1}{2T} \sum_{t=-T+M}^{T-M} \{X(t)Y(t-u) + X(t)Y(t+u)\} \\ S_{XY}(u) = \frac{1}{2T} \sum_{t=-T+M}^{T-M} \{X(t)Y(t-u) - X(t)Y(t+u)\} \end{cases}$$

Então, substituindo os valores esperados de (5.7) por (5.8) e igualando entre si as partes reais e as partes imaginárias, teremos as estimativas de $c_{XY}(\lambda)$ e $q_{XY}(\lambda)$, isto é

$$(5.9) \quad \begin{cases} \hat{c}_{XY}(\lambda) = \frac{1}{\Pi} \{R(0) + \sum_{u=1}^M \cos \lambda u R_{XY}(u)\} \\ \hat{q}_{XY}(\lambda) = \frac{1}{\Pi} \left\{ \sum_{u=1}^M \sin \lambda u S_{XY}(u) \right\} \end{cases}$$

O problema é que estas estimativas não são consistentes, mas isso pode ser resolvido se se usar a média ponderada das estimativas cruas $R_{xy}(u)$ e $S_{xy}(u)$.

Antes de estimar o espectro-cruzado, pode-se aplicar às duas séries um filtro de "quase-diferenças", que consiste em eliminar tendência e a potência alta das frequências baixas, repetindo diferenças da forma

$$(5.10) \quad \Delta(k)Y(t) = Y(t) - kY(t-1) .$$

Pode-se observar que, para $k=1$, temos o método de diferença de variáveis, que foi definido no Capítulo III.

A esse processo de filtragem, dá-se o nome de "pré-

branqueamento" das séries $X(t)$ e $Y(t)$, e deve-se usar o mesmo valor k para as duas séries.

E, depois de estimar o espectro-cruzado, pode-se "re-colorir" as duas séries, isto é, dividi-las, em cada frequência, pela função de transferência do filtro (5.10), que é:

$$(5.11) \quad |A(\lambda)|^2 = [1 - 2k \cos 2\pi\lambda + k^2]^p$$

Para detalhes, ver Nerlove (1964), pp. 257.

Finalmente, dadas as estimativas

$$\hat{c}_{XY}(\lambda), \hat{q}_{XY}(\lambda), \hat{f}_{XY}(\lambda) \text{ e } \hat{f}_{YY}(\lambda),$$

a função de transferência é estimada por

$$(5.12) \quad \hat{A}(\lambda) = \frac{\hat{c}_{XY}(\lambda)}{\hat{f}_{YY}(\lambda)} - i \frac{\hat{q}_{XY}(\lambda)}{\hat{f}_{YY}(\lambda)}$$

e o ganho e o ângulo de fase são estimados por:

$$(5.13) \quad \hat{G}(\lambda) = \frac{\sqrt{\hat{c}_{XY}^2(\lambda) + \hat{q}_{XY}^2(\lambda)}}{\hat{f}_{YY}(\lambda)}$$

$$(5.14) \quad \hat{\phi}(\lambda) = \arctan \left[\frac{\hat{q}_{XY}(\lambda)}{\hat{c}_{XY}(\lambda)} \right]$$

Antes, somente estávamos habilitados a notar a remoção excessiva de potência das frequências não sazonais.

Agora, já podemos medir a "extensão" da remoção da potência, tanto em frequências diferentes da mesma série, como em frequências iguais de séries diferentes, através dos ganhos estimados.

A comparação entre os espectros das séries original e ajustada sazonalmente mostra que os filtros de ajustamento sazonal produzem forte mudança na fase inicial do processo. A fim de que nenhuma mudança ocorra na fase, é necessário que

$$A(\lambda) = \int_{-\infty}^{\infty} a(u) e^{-i\lambda u} du,$$
$$= u(\lambda) - i v(\lambda)$$

seja real, isto é, que $v(\lambda)$ seja igual a 0. E isto só pode ocorrer quando $a(u)$ for simétrica ao redor de $u=0$.

Parece desejável, doravante, tentar planejar filtros que minimizem as mudanças de fase ou, ao menos, as confinem às frequências mais altas.

Como todos os métodos de ajustamento sazonal, existentes até agora, distorcem as relações entre a série original e a ajustada, existe a dúvida sobre a prática padrão de ajustamento sazonal como chave de indicadores de modelos, embora seja necessária sua aplicação aos dados observados.

Segundo Wallis (1974), está aberta uma importante

linha de pesquisa, que consiste em aproveitar a prática econômica e a teoria estatística para se chegar a resultados aplicáveis e de validade científica.

- □ -

CAPÍTULO VI

APLICAÇÕES

Vamos utilizar, aqui, alguns dos conceitos apresentados nos capítulos anteriores, a fim de analisar uma série temporal publicada pelo Bureau do Censo dos Estados Unidos: Total de Vendas a Varejo nos Estados Unidos.

Os dados mensais e o gráfico da série aparecem na tabela 7. (Apêndice 2).

6.1 - TESTE DE VERIFICAÇÃO DE TENDÊNCIA

Vamos aplicar o teste de Postos, apresentado no Capítulo III, seção 3.2.

$$H_0: r_0 = \mu_r$$

$$H_a: r_0 < \mu_r$$

Ao nível de significância $\alpha=5\%$ encontramos o valor crítico $z=-1,64$. A mediana da série é $m_0=17,1$

Foram encontrados $\begin{cases} n_A = 74 \text{ valores a partir de } 17,1 \\ n_B = 70 \text{ valores abaixo de } 17,1 \end{cases}$

$$\left. \begin{array}{l} r_0 = 26 \text{ iterações observadas} \\ \mu_r = 72,94 \\ \sigma_r = 5,97 \end{array} \right\} \Rightarrow z = -7,78$$

$$z < z_c \Rightarrow \text{rejeitamos } H_0.$$

Logo, a série possui o componente $m(t)$ de tendência.

Vamos aplicar agora dois métodos de ajustamento sazonal apresentados no trabalho, e comparar os resultados:

do Bureau do Censo dos Estados Unidos (ver Capítulo IV - método 4);

de média móvel centrada (ver Capítulo III - método 2).

6.2 - AJUSTAMENTO SAZONAL PELO MÉTODO DO BUREAU DO CENSO

Vamos supor, inicialmente, que a série obedece ao modelo multiplicativo

$$y(t) = m(t) \cdot s(t) \cdot \epsilon(t), \quad t=1, \dots, 144.$$

Aplicando o procedimento do capítulo IV, seção 4.5, obtemos as estimativas de $m(t)$,

$$\hat{m}(t) = \frac{1}{24} \{y(t-6) + 2y(t-5) + \dots + 2y(t+5) + y(t+6)\} \\ t=7, \dots, 138,$$

e calculamos $y'(t) = y(t)/\hat{m}(t) = s(t) \cdot \epsilon(t)$, obtendo a série livre de tendência.

A seguir estimamos $s(t)$ e $\epsilon(t)$ separadamente.

E finalmente calculamos $x(t) = y(t)/\hat{s}(t)$ obtendo a

série livre do componente sazonal.

Estes resultados e o gráfico da série $x(t)$ são apresentados nas tabelas 8, 9 e 10.

6.3 - AJUSTAMENTO SAZONAL PELO MÉTODO DE MÉDIA MÓVEL CENTRADA

(apresentado por Durbin, 1963)

Lembrando o procedimento apresentado no método 2, capítulo III, supomos inicialmente o modelo

$$y(t) = m(t) \cdot \alpha(i) \cdot \epsilon(t)$$

$$\text{onde } \begin{cases} t=12j+i \\ j=0, \dots, 11 \\ i=1, \dots, 12 \end{cases}$$

Vamos estimar $m(t)$ como na secção anterior, isto é:

$$\hat{m}(t) = \frac{1}{24} \{y(t-6) + 2y(t-5) + \dots + 2y(t+5) + y(t+6)\}$$

$t=7, \dots, 138,$

$$\begin{aligned} \text{e calculamos } x'(t) &= y(t)/\hat{m}(t) \\ &= \alpha(i) \cdot \epsilon(t). \end{aligned}$$

A seguir calculamos

$$\begin{aligned} \bar{x}'(t) &= \frac{1}{p} \sum_j x'(12j+i) \\ \bar{x}' &= \frac{1}{12} \sum_j \bar{x}'(i) \\ \hat{\alpha}(i) &= \bar{x}'(i)/\bar{x}' \end{aligned}$$

Finalmente encontramos a série livre do componente

sazonal

$$x(t) = y(t)/\hat{a}(i) \cdot$$

Os resultados destes cálculos e o gráfico da série ajustada encontram-se nas tabelas 11 e 12.

Pela observação dos dois gráficos da série ajustada sazonalmente, verificamos que os dois métodos apresentados levam a resultados com poucas diferenças.

Isso nos leva a crer que os dois métodos têm propriedades semelhantes como técnicas de ajustamento sazonal.

A seguir vamos aplicar a análise espectral aos resultados obtidos pelo método do Bureau do Censo.

6.4 - ANÁLISE ESPECTRAL DO MÉTODO DE AJUSTAMENTO SAZONAL DO BUREAU DO CENSO

Utilizando os dados das séries original e ajustada pelo método do Bureau do Censo, foram calculados os espectros

$f_y(\lambda)$ para a série original $y(t)$ e

$f_x(\lambda)$ para a série ajustada $x(t)$.

Foi utilizado o programa BMD03T para efetuar os calculos.

Como vimos, no capítulo V, são encontradas 6 frequências sazonais, a saber:

FREQÜÊNCIAS	CICLOS
$\lambda_1 = 1/12 \approx 0,083$	12 meses
$\lambda_2 = 2/12 \approx 0,167$	6 meses
$\lambda_3 = 3/12 \approx 0,250$	4 meses
$\lambda_4 = 4/12 \approx 0,333$	3 meses
$\lambda_5 = 5/12 \approx 0,417$	2,4 meses
$\lambda_6 = 6/12 \approx 0,500$	2 meses

O programa calculou ainda os estimadores do ganho $G(\lambda)$, do ângulo de fase $\phi(\lambda)$ e da coerência $K(\lambda)$, que mostram as relações entre as séries original e ajustada, sob a suposição de que

$$X(t) = \int_u a(u)Y(t-u)$$

isto é, que o filtro é linear.

Os resultados e gráficos obtidos aparecem na tabela 13.

Comparando os espectros $f_y(\lambda)$ e $f_x(\lambda)$ vemos que o filtro eliminou a potência das frequências altas e eliminou todos os picos das frequências sazonais.

É, portanto, um filtro conveniente de ajustamento sazonal, mas não é suficientemente satisfatório, como se percebe pela análise de $G(\lambda)$, $\phi(\lambda)$ e $K(\lambda)$.

Como vimos no capítulo III, se o filtro é linear, va

lem as relações:

$$K(\lambda) = f_{xy}(\lambda)/f_x(\lambda) \cdot f_y(\lambda)$$

$$f_x(\lambda) = G^2(\lambda) f_y(\lambda)$$

$$\phi(\lambda) = \text{arc tg}[q(\lambda)/c(\lambda)]$$

Isso não se verificou nos nossos resultados, o que mostra que o filtro do Bureau do Censo é não linear.

O valor $K(\lambda)$ tem significado análogo ao do coeficiente de correlação linear, isto é, $0 \leq K(\lambda) \leq 1$.

Para que o filtro seja considerado satisfatório, é necessário que $K(\lambda) \approx 1 \quad \forall \lambda$. No nosso caso, isso não ocorre; em muitas frequências encontramos $K(\lambda) < 0,5$.

O valor $\phi(\lambda)$ mede o deslocamento da série ajustada em relação à série original. Espera-se que, usando um filtro conveniente, o valor $\phi(\lambda)$ seja $\approx 0 \pmod{\pi}$, $\forall \lambda$; isto é, que $q(\lambda)$ seja muito pequeno para que tenha ocorrido pouca mudança de fase.

No nosso caso, em muitas frequências, encontramos um valor $\phi(\lambda)$ distante de $0 \pmod{\pi}$; isto é, houve deslocamento da série $x(t)$ em relação a $y(t)$.

Além disso, quando o filtro é o mais conveniente, espera-se que $G(\lambda)$ seja como segue:

$G(\lambda) \approx 0$ nas frequências sazonais, isto é, o filtro eliminou o componente sazonal;

$G(\lambda) \approx 1$ nas frequências não sazonais, isto é, o filtro não removeu mais do que os fatores estritamente sazonais.

Isso não se verificou no nosso caso. Parece que o filtro removeu mais do que o considerado puramente sazonal.

Em compensação não introduziu fatores não sazonais, o que teria ocorrido se $G(\lambda) > 1$ para algum λ .

6.5 - CONCLUSÕES

Verificamos que o filtro do Bureau do Censo merece aprimoramento. Ele provoca algumas distorções nas relações entre as séries original e ajustada e, portanto, não atinge completamente o objetivo para o qual foi construído: eliminar somente o componente sazonal da série temporal.

Isso tem ocorrido, com algumas variações, em todos os filtros de ajustamento sazonal até hoje apresentados.

Nosso objetivo, neste trabalho, foi apresentar os métodos de ajustamento sazonal e as restrições que apresentam. No entanto, como saliente Nerlove (1973), é necessário comparar todos os métodos até agora apresentados a fim de tentar construir outros, de validade científica e aplicabilidade imediata, sem os defeitos anteriores.

APÊNDICE 1

ALGUMAS SÉRIES TEMPORAIS E SEUS GRÁFICOS

TABELA 1

NÍVEL MÉDIO MENSAL DAS ÁGUAS DO LAGO MICHIGAN-HURON

Fonte: Departamento Norte-americano de Exploração de Lagos
Corpo de Engenheiros do Exército dos Estados Unidos

ANO	NÍVEL	ANO	NÍVEL	ANO	NÍVEL	ANO	NÍVEL	ANO	NÍVEL	ANO	NÍVEL
1860	583,3	1876	583,6	1892	581,0	1908	581,8	1924	579,6	1940	579,3
61	583,5	77	582,7	93	581,3	09	581,1	25	578,5	41	579,0
62	583,2	78	582,5	94	581,4	10	580,5	26	578,5	42	580,2
63	582,6	79	581,5	95	581,2	11	580,0	27	579,6	43	581,5
1864	582,2	1880	582,1	1896	580,0	1912	580,7	1928	580,6	1944	580,8
65	582,1	81	582,2	97	580,8	13	581,3	29	582,3	45	581,0
66	581,7	82	582,6	98	580,8	14	580,7	30	581,2	46	581,0
67	582,2	83	583,3	99	581,1	15	580,0	31	579,1	47	581,1
1868	581,6	1884	583,1	1900	580,7	1916	581,1	1932	578,6	1948	580,8
69	582,1	85	583,3	01	581,1	17	581,9	33	578,7	49	579,7
70	582,7	86	583,7	02	580,8	18	581,9	34	578,0	50	580,0
71	582,8	87	582,0	03	580,8	19	581,3	35	578,6	51	581,6
1872	581,5	1888	582,3	1904	581,5	1920	581,0	1936	578,7	1952	582,7
73	582,2	89	581,8	05	581,6	21	580,5	37	578,6	53	582,1
74	582,3	90	581,6	06	581,5	22	580,6	38	579,7	54	581,7
75	582,1	91	580,9	07	581,6	23	579,8	39	580,0	55	581,5

TABELA 2

MILHAS PERCORRIDAS PELOS AVIÕES DAS COMPANHIAS INGLESAS

Fonte: Kendall, M.G. (1973)

(1:1.000.000)

	1963	1964	1965	1966	1967	1968	1969	1970
Janeiro..	6,8	7,3	8,4	8,2	8,3	8,6	9,5	10,8
Fevereiro	6,2	6,8	7,8	7,4	7,9	8,8	8,9	10,4
Março....	7,1	7,8	8,8	8,5	10,0	10,9	11,6	13,6
Abril....	8,2	8,4	9,9	9,9	10,1	10,5	8,8	13,4
Maió.....	8,5	9,1	10,6	10,2	10,8	11,2	12,5	13,1
Junho....	9,6	10,2	11,2	12,3	10,9	10,6	14,8	14,9
Julho....	10,5	11,0	11,4	11,6	12,2	10,8	13,7	14,1
Agosto...	10,7	10,9	11,4	11,6	12,2	12,8	13,7	14,0
Setembro.	10,0	10,3	10,7	12,4	13,3	13,8	15,1	16,2
Outubro..	8,2	9,1	9,4	9,6	10,4	10,9	12,2	12,4
Novembro.	6,8	7,7	7,8	8,1	8,7	9,3	10,6	11,6
Dezembro.	7,0	7,6	7,9	9,3	9,6	10,9	12,2	12,8

TABELA 3

TEMPERATURA EM SÃO PAULO

Fonte: IAG - USP

-

Média de 24 leituras diárias

MESES ↓	ANOS →	1949	1950	1951	1952	1953	1954	1955	1956	1957
Janeiro		-	20,6	20,5	20,8	21,9	22,1	21,1	23,4	21,3
Fevereiro		-	20,9	20,7	20,1	20,7	21,5	21,9	21,2	20,4
Março		-	20,6	20,0	20,5	20,9	21,0	20,7	20,3	20,4
Abril		-	18,8	15,9	17,1	18,1	18,3	18,6	18,1	18,1
Maio		-	17,5	15,7	16,7	16,9	16,5	15,7	15,1	15,9
Junho		-	16,1	14,6	15,2	15,5	16,3	14,9	13,5	15,0
Julho		14,5	14,7	13,2	14,8	13,3	15,7	14,7	14,0	-
Agosto		15,8	17,0	14,0	16,8	15,8	16,2	15,5	13,6	-
Setembro		16,0	17,0	16,3	15,9	17,8	17,7	16,6	17,6	-
Outubro		17,7	16,9	17,5	18,0	18,4	17,7	16,4	17,0	-
Novembro		17,6	17,8	19,0	19,1	18,6	18,5	16,9	16,7	-
Dezembro		19,6	19,9	18,4	19,8	19,4	19,6	20,2	19,1	-

TABELA 4

NASCIDOS VIVOS NOS ESTADOS UNIDOS

Fonte: National Office of Vital Statistics (1:10.000)

MESES ↓	ANOS →	1949	1950	1951	1952	1953
Janeiro.		29,9	29,7	30,4	31,2	32,2
Fevereiro.		27,3	27,3	28,2	30,0	29,6
Março.		30,0	29,4	31,3	31,7	31,5
Abril.		27,1	25,9	29,6	29,2	28,7
Maio		28,2	27,6	31,3	30,0	30,7
Junho.		28,5	29,4	30,7	31,1	32,1
Julho.		31,8	31,6	32,8	34,5	35,4
Agosto		32,3	32,5	33,4	35,0	35,6
Setembro		31,3	31,5	32,9	34,4	34,8
Outubro.		31,1	31,2	32,9	33,6	33,4
Novembro		29,1	29,2	30,4	31,5	32,0
Dezembro		29,3	30,1	31,2	32,3	34,0

TABELA 5

VENDA MENSAL DE CHAMPAGNE

Fonte: Revue Française d'Automatique, Informatique et Recherche
Operationnelle; março de 1973, Vol. I (por 1.000 garrafas)

	1964	1965	1966	1967	1968	1969	1970
JAN.	-	5.375	3.633	4.016	2.639	3.934	4.348
FEV.	-	3.088	4.292	3.957	2.899	3.162	3.564
MAR.	-	3.713	4.154	4.510	3.370	4.286	4.577
ABR.	-	4.514	4.121	4.276	3.740	4.676	4.788
MAI.	-	4.520	4.647	4.968	2.927	5.010	4.618
JUN.	3.986	4.539	4.753	4.677	3.936	4.874	-
JUL.	3.260	3.663	3.965	3.523	4.217	4.633	-
AGO.	1.573	1.643	1.723	1.821	1.738	1.659	-
SET.	3.528	4.739	5.048	5.222	5.221	5.951	-
OUT.	5.211	5.428	6.922	6.873	6.424	6.981	-
NOV.	7.614	8.314	9.858	10.803	9.842	9.851	-
DEZ.	9.254	10.651	11.331	13.916	13.076	12.670	-

TABELA 6

SÉRIE DE CHUVAS EM SÃO PAULO

Fonte - IAG - Dept^o Meteorologia - USP (em milímetros)

	1965	1966	1967	1968	1969	1970	1971	1972
Janeiro. .	-	234,4	215,7	257,1	109,6	268,8	193,9	258,8
Fevereiro. .	-	232,5	247,9	117,6	147,0	362,8	347,3	206,3
Março. . .	-	261,2	234,4	249,6	94,6	107,8	177,2	85,1
Abril. . .	-	192,1	48,7	40,4	77,6	55,2	79,0	53,0
Maio . . .	-	36,7	22,4	71,9	42,9	96,9	51,1	43,3
Junho. . .	-	7,5	87,6	42,5	62,7	55,1	96,9	5,1
Julho. . .	81,5	24,5	39,6	19,3	6,3	21,5	42,6	-
Agosto . .	7,8	65,6	5,2	64,4	36,1	92,0	19,0	-
Setembro .	87,3	93,9	107,9	23,0	41,3	101,9	89,2	-
Outubro. .	189,5	155,8	222,5	154,9	189,5	118,7	103,3	-
Novembro .	132,6	64,9	136,7	79,3	258,3	60,3	122,5	-
Dezembro .	201,3	261,4	123,7	229,2	109,5	134,2	160,5	-

APÊNDICE 2

TABELAS E GRÁFICOS REFERENTES AO CAPÍTULO VI

TABELA 7

TOTAL DE VENDAS A VAREJO NOS ESTADOS UNIDOS

Série original: $y(t) = m(t)e(t)$

(1:1.000.000.000)
em US\$

	JAN.	FEV.	MAR.	ABR.	MAI.	JUN.	JUL.	AGO.	SET.	OUT.	NOV.	DEZ.
1953	12,9	12,2	13,7	14,1	14,5	14,4	14,2	14,0	13,9	14,8	13,8	16,3
1954	12,2	11,8	13,6	14,0	14,1	14,5	14,3	13,8	14,0	14,5	14,4	17,7
1955	13,1	12,6	14,6	15,5	15,3	15,6	15,3	15,5	15,8	15,7	15,7	19,1
1956	13,7	13,6	15,5	15,1	16,1	16,6	15,4	16,2	15,6	16,1	16,5	19,4
1957	14,7	14,1	15,9	16,3	17,2	17,1	16,9	17,5	16,4	16,9	17,1	19,8
1958	15,3	13,8	15,5	16,4	17,4	16,6	16,6	17,0	16,3	17,4	17,0	21,2
1959	16,2	15,0	17,0	17,8	18,6	18,7	18,3	18,1	17,6	19,1	17,6	21,5
1960	16,3	15,8	17,6	19,0	18,5	18,9	18,1	18,2	17,8	18,6	18,4	22,2
1961	15,8	15,1	17,7	17,6	18,5	18,9	17,9	18,3	18,2	18,8	19,2	22,9
1962	17,0	16,0	19,2	19,1	20,2	20,3	19,1	19,9	18,9	20,6	20,9	24,1
1963	18,3	17,1	19,7	20,5	21,2	20,7	20,5	21,0	19,3	21,5	21,5	25,1
1964	19,2	18,6	20,5	21,2	22,5	22,2	22,1	21,8	21,3	22,6	21,7	27,7

Fonte: Bureau of the Census - Technical Paper nº 15.

TABELA 8

TOTAL DE VENDAS A VAREJO NOS ESTADOS UNIDOS

tendência $\hat{m}(t)$

	JAN.	FEV.	MAR.	ABR.	MAI.	JUN.	JUL.	AGO.	SET.	OUT.	NOV.	DEZ.
1953	-	-	-	-	-	-	14,1	14,0	14,0	14,0	14,0	14,0
1954	14,0	13,9	13,9	13,9	13,9	14,0	14,1	14,2	14,3	14,4	14,5	14,6
1955	14,7	14,8	14,9	15,0	15,1	15,2	15,3	15,4	15,5	15,5	15,5	15,6
1956	15,6	15,7	15,7	15,7	15,8	15,8	15,9	15,9	16,0	16,0	16,1	16,2
1957	16,3	16,4	16,5	16,5	16,6	16,6	16,7	16,7	16,7	16,6	16,7	16,6
1958	16,6	16,6	16,6	16,6	16,6	16,6	16,7	16,8	16,8	17,0	17,1	17,3
1959	17,5	17,6	17,7	17,8	17,9	17,9	18,0	18,0	18,1	18,1	18,2	18,2
1960	18,2	18,2	18,2	18,2	18,2	18,3	18,3	18,2	18,2	18,1	18,1	18,1
1961	18,1	18,1	18,1	18,1	18,1	18,2	18,3	18,4	18,5	18,6	18,7	18,9
1962	19,0	19,1	19,2	19,3	19,4	19,6	19,7	19,8	19,8	19,9	20,0	20,1
1963	20,1	20,2	20,3	20,4	20,4	20,5	20,6	20,7	20,8	20,8	20,9	21,0
1964	21,2	21,3	21,4	21,5	21,6	21,7	-	-	-	-	-	-

TABELA 9

TOTAL DE VENDAS A VAREJO NOS ESTADOS UNIDOS

fatores sazonais $\hat{s}(t)$

	JAN.	FEV.	MAR.	ABR.	MAI.	JUN.	JUL.	AGO.	SET.	OUT.	NOV.	DEZ.
1953	0,89	0,86	0,98	1,00	1,02	1,03	1,01	1,00	0,99	1,01	1,00	1,22
1954	0,89	0,86	0,98	1,00	1,02	1,03	1,00	1,00	0,99	1,01	1,00	1,21
1955	0,89	0,86	0,98	1,00	1,02	1,03	1,00	1,01	0,98	1,01	1,01	1,21
1956	0,89	0,86	0,98	0,99	1,02	1,03	1,00	1,01	0,98	1,01	1,02	1,21
1957	0,90	0,85	0,97	0,99	1,03	1,03	1,00	1,01	0,97	1,01	1,01	1,20
1958	0,91	0,85	0,97	0,99	1,03	1,03	1,00	1,01	0,97	1,02	1,01	1,21
1959	0,91	0,85	0,97	0,99	1,03	1,04	1,00	1,00	0,97	1,02	1,01	1,21
1960	0,90	0,84	0,97	0,99	1,03	1,04	0,99	1,00	0,97	1,03	1,02	1,21
1961	0,90	0,84	0,97	1,00	1,03	1,03	0,99	1,00	0,97	1,03	1,03	1,21
1962	0,90	0,84	0,97	1,00	1,03	1,03	0,98	1,01	0,97	1,03	1,03	1,21
1963	0,90	0,84	0,97	1,00	1,04	1,02	0,98	1,01	0,96	1,03	1,03	1,20
1964	0,90	0,84	0,97	0,99	1,04	1,02	0,98	1,01	0,96	1,03	1,03	1,20

TABELA 10

SÉRIE SAZONALMENTE AJUSTADA $x(t)$ SEGUNDO O MÉTODO
DO BUREAU DO CENSO DOS EUA

	JAN.	FEV.	MAR.	ABR.	MAI.	JUN.	JUL.	AGO.	SET.	OUT.	NOV.	DEZ.
1953	14,6	14,2	14,0	14,1	14,3	14,0	14,1	14,0	14,1	14,6	13,9	13,4
1954	13,8	13,9	13,9	14,0	13,9	14,1	14,2	13,7	14,2	14,4	14,4	14,6
1955	14,8	14,7	14,9	15,5	15,1	15,1	15,2	15,4	16,0	15,5	15,6	15,8
1956	15,4	15,8	15,9	15,2	15,7	16,0	15,4	16,0	16,0	15,9	16,2	16,1
1957	16,4	16,4	16,4	16,5	16,7	16,6	16,8	17,3	16,8	16,7	16,9	16,5
1958	16,9	16,2	16,0	16,5	16,8	16,0	16,6	16,9	16,8	17,0	16,9	17,5
1959	17,9	17,6	17,6	17,9	18,0	18,1	18,3	18,0	18,0	18,6	17,4	17,7
1960	18,1	18,7	18,2	19,1	18,0	18,2	18,2	18,1	18,3	18,2	18,1	18,2
1961	17,5	17,9	18,2	17,7	18,0	18,3	18,2	18,3	18,7	18,2	18,7	18,9
1962	18,9	19,1	19,7	19,2	19,6	19,7	19,5	19,8	19,5	20,0	20,2	20,0
1963	20,2	20,3	20,3	20,6	20,4	20,2	20,8	20,8	20,0	20,8	20,8	20,9
1964	21,2	22,2	21,2	21,3	21,6	21,6	22,5	21,6	22,1	21,9	21,0	23,1

TABELA 11

$$x'(t) = y(t)/\hat{m}(t) = \alpha(i) \cdot \varepsilon(t)$$

	JAN.	FEV.	MAR.	ABR.	MAI.	JUN.	JUL.	AGO.	SET.	OUT.	NOV.	DEZ.	
1953	-	-	-	-	-	-	1,01	1,00	0,99	1,06	0,98	1,16	
1954	0,87	0,86	0,98	1,01	1,01	1,04	1,01	0,97	0,98	1,01	0,99	1,21	
1955	0,89	0,85	0,98	1,03	1,01	1,03	1,00	1,01	1,02	1,01	1,01	1,22	
1956	0,88	0,87	0,99	0,96	1,02	1,05	0,97	1,02	0,97	1,01	1,02	1,20	
1957	0,90	0,86	0,96	0,99	1,04	1,03	1,01	1,05	0,98	1,02	1,02	1,19	
1958	0,92	0,83	0,93	0,99	1,05	1,00	0,99	1,01	0,96	1,02	0,99	1,23	
1959	0,93	0,85	0,96	1,00	1,04	1,04	1,02	1,01	0,97	1,05	0,97	1,18	
1960	0,90	0,87	0,97	1,04	1,02	1,03	0,99	1,00	0,98	1,03	1,02	1,23	
1961	0,87	0,83	0,98	0,97	1,02	1,04	0,98	0,99	0,98	1,01	1,03	1,21	
1962	0,89	0,84	1,00	0,99	1,04	1,04	0,97	1,00	0,95	1,03	1,04	1,20	
1963	0,91	0,85	0,97	1,00	1,04	1,01	1,00	1,01	0,93	1,03	1,03	1,20	
1964	0,91	0,87	0,96	0,99	1,04	1,02	-	-	-	-	-	-	
$\bar{x}'(i)$	0,90	0,85	0,97	1,00	1,03	1,03	0,99	1,01	0,97	1,02	1,02	1,21	$\sum_1 \bar{x}'(i)=12$
$\hat{g}(i)=\bar{x}'(i)/\bar{x}'$	0,90	0,85	0,97	1,00	1,03	1,03	0,99	1,01	0,97	1,02	1,02	1,21	$\sum_1 \hat{g}(i)=12$

TABELA 12
SÉRIE SAZONALMENTE AJUSTADA $\times(t)$ SEGUNDO O MÉTODO
DE DURBIN

	JAN.	FEV.	MAR.	ABR.	MAI.	JUN.	JUL.	AGO.	SET.	OUT.	NOV.	DEZ.
1953	14,3	14,3	14,1	14,1	14,1	14,0	14,3	13,9	14,3	14,5	13,5	17,5
1954	13,6	14,0	14,0	14,0	13,7	14,1	14,4	13,7	14,4	14,2	14,1	14,6
1955	14,6	14,8	15,0	15,5	14,8	15,1	15,5	15,3	16,3	15,4	15,4	15,8
1956	15,2	16,0	16,0	15,1	15,6	16,1	15,5	16,0	16,1	15,8	16,2	16,0
1957	16,3	16,6	16,4	16,3	16,7	16,6	17,1	17,3	16,9	16,6	16,8	16,4
1958	17,0	16,2	16,0	16,4	16,9	16,1	16,8	16,8	16,8	17,1	16,7	17,5
1959	18,0	17,6	17,5	17,8	18,1	18,1	18,5	17,9	18,1	18,7	17,3	17,8
1960	18,1	18,6	18,1	19,0	18,0	18,3	18,3	18,0	18,3	18,2	18,0	18,3
1961	17,5	17,8	18,2	17,6	18,0	18,3	18,1	18,1	18,8	18,4	18,8	18,9
1962	18,9	18,8	19,8	19,1	19,6	19,7	19,3	19,7	19,5	20,2	20,5	19,9
1963	20,3	20,1	20,3	20,5	20,6	20,1	20,7	20,8	19,9	21,1	21,1	20,7
1964	21,3	21,9	21,1	21,2	21,8	21,6	22,3	21,6	21,9	22,2	21,3	22,9

TABELA 13

FREQUÊNCIAS	$\hat{f}_Y(\lambda)$	$\hat{f}_X(\lambda)$	$\hat{G}(\lambda)$	$\hat{\phi}(\lambda)$	$\hat{K}(\lambda)$
0,000	4,54	6,99	5,605	0,0	0,995
0,016	3,85	4,65	4,211	-0,7	0,995
0,031	2,10	2,32	2,197	1,0	0,996
0,047	0,83	0,96	0,876	2,7	0,986
0,063	0,21	0,23	0,168	10,0	0,763
0,078	4,15	0,08	0,133	-132,9	0,234
0,094	1,76	0,06	0,169	-85,0	0,515
0,109	0,16	0,04	0,021	-19,1	0,252
0,125	0,23	0,11	0,088	11,3	0,544
0,141	0,37	0,08	0,105	67,3	0,622
0,156	7,76	0,06	0,083	36,1	0,122
0,172	17,55	0,05	0,654	-162,4	0,702
0,188	0,42	0,04	0,034	173,7	0,269
0,203	0,10	0,05	0,055	111,1	0,794
0,219	0,21	0,11	0,082	-11,5	0,538
0,234	0,50	0,10	0,091	-40,8	0,410
0,250	10,13	0,10	0,061	80,4	0,060
0,266	0,44	0,05	0,081	-63,5	0,523
0,281	0,08	0,05	0,024	-63,9	0,379
0,297	0,13	0,12	0,085	-14,0	0,690
0,313	0,14	0,14	0,054	45,3	0,383
0,328	7,39	0,02	0,288	166,5	0,729
0,344	4,13	0,38	0,480	-5,0	0,385
0,359	0,16	0,19	0,150	38,2	0,839
0,375	0,11	0,01	0,012	45,9	0,312
0,391	0,10	0,08	0,017	-43,3	0,198
0,406	2,84	0,03	0,150	-107,8	0,500
0,422	5,83	0,05	0,194	-28,8	0,368
0,438	0,35	0,10	0,163	3,9	0,888
0,453	0,07	0,01	0,025	14,0	0,942
0,469	0,11	0,06	0,067	11,1	0,846
0,484	0,26	0,09	0,099	-15,2	0,643
0,500	3,31	0,03	0,173	0,0	0,582

B I B L I O G R A F I A

- [01] - ANDERSON, T.W. - The Statistical Analysis of Time Series, John Wiley & Sons, Inc., 1971.
- [02] - BORGMAN, L.E. - Statistical Time Series Analysis, University of California, Davis, 1966.
- [03] - BRILLINGER, D.R. - Frequency Analysis of Vector-Valued Time Series, Holt, 1974.
- [04] - DOOB, J.L. - Stochastic Processes, John Wiley & Sons, Inc., 1953.
- [05] - DURBIN, J. - Trend Elimination for the Purpose of Estimating Seasonal and Periodic Component of Time Series, Proceeding of the Symposium on Time Series Analysis, John Wiley & Sons, Inc., 1963, pp 3-16.
- [06] - " " - Trend Elimination by Moving-Average and Variate - Difference Filters - Bulletin de L'Institut International de Statistique, 33^e Session; 1961, pp. 1-12.
- [07] - DRAPER, N.R. e SMITH, H. - Applied Regression Analysis, John Wiley & Sons, Inc., 1966.
- [08] - GOODMAN, N.R. - Some Comments on Spectral Analysis of Time Series, Technometrics, Vol. 3, n° 2, Maio de 1961, pp. 221-228.
- [09] - GRANGER, C.W.J. e HATANAKA, M. - Spectral Analysis of Economic Time Series, Princeton, 1964.
- [10] - HAJEK, J. - Nonparametric Statistics, Holden Day, 1969.
- [11] - HANNAN, E.J. - The Estimation of Seasonal Variation, Australian Journal of Statistics, Vol. 2, 1960, pp. 1-15.

- [12] - _____ - Time Series Analysis, Methuen, London, 1960.
- [13] - _____ - Testing for a Jump in the Spectral Function, Journal of Royal Statistical Society, Vol.23, 1961, pp. 396-404.
- [14] - JENKINS, G.W. e WATTS, D.G. - Spectral Analysis, Holden Day, 1969.
- [15] - JORGENSON, D.W.- Minimum Variance, Linear Unbiased Seasonal Adjustment of Economic Time Series, Journal of the American Statistical Association, n° 15, - 1964, pp. 681-724.
- [16] - _____ - Seasonal Adjustment of Data for Econometric Analysis, Journal of the American Statistical Association, n° 38, 1967, pp. 137-140.
- [17] - KENDALL, M.G. - Time Series, Griffin, London, 1973.
- [18] - LINDGREEN, B.W. e MC ELRATH, G.W. - Introduction to Probability and Statistics, McGraw-Hill, 1972.
- [19] - MOOD, A.M. e GRAYBILL, F.A. - Introduction to the Theory of Statistics, McGraw-Hill, 1963.
- [20] - NERLOVE, M. - Spectral Analysis of Seasonal Adjustments Procedures, Econometrica, Vol. 32, 1964, pp. 241-286.
- [21] - PARZEN, E. - Mathematical Considerations in the Estimation of the Spectra, Technometrics, Vol. 3, 1961, pp. 167-190.
- [22] - SHISKIN, J.; YOUNG, A.H. e MUSGRAVE, J.C. - The X-11 - Variant of the Census Method II Seasonal Adjustment Program, Bureau of The Census, Technical Paper, n° 15, 1967.
- [23] - THOMAS, J.J. e WALLIS, K.F. - Seasonal Variation in Regression Analysis, Journal of the Royal Statistical Society, n° 134, 1971, pp. 57-72.

- [24] - TOLSTOV, G.P. - Fourier Series, Prentice Hall, 1962.
- [25] - TUKEY, J.W. - Discussion Emphasizing the Connection
Between Analysis of Variance and Spectrum Analysis,
Technometrics, Vol. 3, 1961, pp. 1-29.
- [26] - U.S. BUREAU OF LABOR STATISTICS - The B.L.S. Seasonal
Factor Method, U.S. Department of Labor, 1966.
- [27] - WALLIS, K.F. - Seasonal Adjustment and Relations
Between Variables, Journal of The American Statis-
tical Association, Vol. 69, 1974, pp. 18-31.
- [28] - WALLIS W.A. e ROBERTS, H.V. - Statistics: A New Appro-
ach, Glencoe III, Free Press, 1956, Cap. 18.
- [29] - YAGLOM, A.M. - Introduction to Stationary Random Func-
tions, Prentice Hall, 1962.