



6º Simpósio Nacional de Probabilidade e Estatística 1984

SÉRIES TEMPORAIS MULTIVARIADAS

BASILIO DE BRAGANÇA PEREIRA

INSTITUTO DE MATEMÁTICA

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

ERRATA

Pág.	linha	onde se le	leia-se
ii	13	Journal of Time Series	Journal of Forecasting
iv	6	ARMA(p,d,q)	ARIMA(p,d,q)
2	8	o valor exportações	o valor das exportações
2	18,19,21	... t ∈ T}	... t ∈ T}
3	última	$\sum_{v=0}^{\infty}$	$\sum_{u=0}^{\infty}$
19	última	Cov(Y(t ₁), Y(t ₂)) = C(τ).	Cov(Y(T ₁), Y(T ₂)) = C(τ), (τ = t ₂ - t ₁).
22	7	+ cos λτdC(w)]	+ cos wτdC(w)].
23	9	Como no	No
24	4	α _i sen w _i t	α _i sen w _i t
28	17	F _y e Y _{t+}	F _{y_t} , Y _{t+1}
29	2	a ₁ L ²	a ₂ L ²
31	última	λ _t	λ _j
32	8	X _t β	B(L)X _t
33	2	X _{t+}	X _{t+1}
33	penúltima	+ E ⁻¹	+ EΓ ⁻¹
36	7	(ψ ₃ - ψ ₁ ψ ₁)	(ψ ₃ - ψ ₁ ψ ₂)
37	5	ψ(L)Y _t	ψ(L)ε _t
40	4	Y =	Y _t =
41	7	- φY _{t-1}	- φ ₁ Y _{t-1}
41	14	+ φ ² ε _{t-1} + ...	+ φ ² ε _{t-2} + ...
42	4	K ≥ 0	k ≥ 0
43	6	β _k = φ ₁ ρ _{k-1} + φ ₂ Y _{k-2}	ρ _k = φ ₁ ρ _{k-1} + φ ₂ ρ _{k-2}
44	17	Y _t - φ ₁ Y _t - ...	Y _t - φ ₁ Y _{t-1} - ...
45	14	θ _q (L)ε _t	θ(L)ε _t
45	19	ε + θ ₁ ε _{t-1}	ε + θε _{t-1}

pág.	linha	onde se lê	leia-se
45	última	$\theta^2 L^2 + \dots$	$\theta^2 L^2 - \dots$
47	10	$\gamma_0 = (1 + \sigma_1^2 + \dots + \sigma_q^2) \sigma_\varepsilon^2$	$\gamma_0 = (1 + \theta_1^2 + \dots + \theta_q^2) \sigma_\varepsilon^2$
49	7 e 9	$\phi_p(L)$ e $\theta_q(L)$	$\phi(L)$ e $\theta(L)$
49	14	$= \phi_1 Y_{t-1} + \varepsilon_t + \phi \varepsilon_{t-1}$	$= \phi Y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$
49	17	$\frac{1 - \phi L}{1 - \theta L} Y_t$	$\frac{1 - \phi L}{1 + \theta L} Y_t$
49	18	$\frac{1 + \phi L}{1 - \phi L} \varepsilon_t$	$\frac{1 + \theta L}{1 - \phi L} \varepsilon_t$
50	1,2,3,4	ϕ_1^k e θ_1	ϕ e θ
50	11	$\gamma_k = \phi_1^{k-1} \gamma$	$\gamma_k = \phi^{k-1} \gamma_{k-1}, k > 1$
51	7	$E\{Y_t, \varepsilon_{t-l}\}$	$E\{Y_t, \varepsilon_{t+l}\}$
53	6,8	$\phi_p(L)$ e $\theta_q(L)$	$\phi(L)$ e $\theta(L)$
53	10	ARMA(p,d,q)	ARIMA(p,d,q)
54	10	$\lambda = 0.9, -0.8 \dots$	$\lambda = -0.9, -0.8 \dots$
55	Fig. II1(b)	ΔY_t	∇Y_t
58	Fig. II4	ΔY $\Delta^2 Y$	∇Y e $\nabla^2 Y$
59	5	$\theta = 0$	$\theta_0 = 0$
61	2,3	$\theta_1 n_{t-s}^- , \theta_1 L^S - , \theta(L) n_t$	$\theta_1 n_{t-s}^+ , \theta_1 L^S + , \theta(L^S) t$
61	10	$\phi_1 Y_{t-s}^- \dots - \phi_p Y_{t-ps}$	$\phi_1 Y_{t-s}^+ \dots + \phi_p Y_{t-ps}$
62	14	(L^{12})	(L^S)
62	20,22	$\theta_1 L^- , \theta_1 L^S -$	$\theta_1 L^+ , \theta_1 L^S +$
70	2	AR(1) e AR(2)	AR(1) e MA(2)
71	penúltima	$ P_k^+ $	$ P^* $
72	17	parcial é infinita, tendendo	parcial tende
72	20	ii)	iii)
73	7	era sujeita	é sujeita
74	8	$\hat{\rho}_k^*$	$\hat{\rho}_k^+$
75	1	i)	ii)
76	17	$\phi(L^S) \phi(L) - \nabla_S^D \dots$	$\phi(L^S) \phi(L) \nabla_S^D \dots$
76	última	Hosking (1980)	Hosking.(1980b)

pág.	linha	onde se lê	leia-se
77	8	$\phi_k + \phi_1 \phi_{k-1} + \dots$	$\phi_k + \phi_1 \phi_{k+1} + \dots$
78	5	$\dots + \theta_q Y_{t-p} + \varepsilon_t$	$\dots + \theta_q Y_{t-q} + \varepsilon_t$
78	7	$\dots + \theta_q \rho^q I_{k-p}$	$\dots + \theta_q \rho^q I_{k-q}$
78	11	$\dots + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \dots + \phi_p \varepsilon_{t-p}$	$\dots + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$
79	20	Hosking (1980) e Bhansali (1981)	Hosking (1980a) e Bhansali (1980)
80	12	FAI	FACI
82	14	(1968)	(1978)
83	17,21	$P(1,2), P(k,1)$	$P(1,2), P(k,1)$
85	penúltima	$P^+(i,k)$	$P^*(i,k)$
88	9	$\dots - U_d B^d$	$\dots - U_d L^d$
91	4	e $MA(q)$	e $MA(q_0)$
91	penúltima	estepwise	stepwise
92	entre 13 e 14	p \vdots $p+1$	p \vdots $p+1$
96	3	a forma	a ordem
101	1	ϕ_1	ϕ
101	5	$\varepsilon_2 = Y - \phi_1 Y_1 \dots$	$\varepsilon_2 = Y_2 - \phi Y_1 \dots$
102	12	1983	1982
104	8	Q	Q
105	8	$\rho_k(\hat{\varepsilon}_t)$	$\hat{\rho}_k(\hat{\varepsilon}_t)$
105	15	$\varepsilon_t = \phi^{-1}(L) \dots$	$\varepsilon_t = \phi^{-1}(L) \dots$
106	4	SARIMA(p,d,q)	SARIMA(p,d,q) x
107	3	Y_t, Y_{t+1}, \dots	Y_t, Y_{t-1}, \dots
107	5	$\psi(B)\varepsilon_t$	$\psi(L)\varepsilon_t$
107	11	a_t, a_{t-1}	$\varepsilon_t, \varepsilon_{t-1}$
110	12	$\phi_1 Y_t(p+j+1)$	$\phi_1 Y_t(p+j-1)$
110	15	$\theta \varepsilon_{t-1}$	$\theta \varepsilon_{t-1}$
111	6	$\phi_1 Y_t(p+j-1)$	$\phi_1 Y_t(p+j-1)$

pág.	Linha	onde se lê	leia-se
114	12	determinado	determinada
118	13	$\ell n \sigma^2 + \frac{2M}{T}$	$\ell n \hat{\sigma}^2 + \frac{2M}{T}$
119	10	$\{Y_t - Y_{t-\ell}\}^2$	$\{Y_t - \phi_\ell Y_{t-\ell}\}^2$
121	5	Hankel	Hankel truncada
138	17	$\beta_j x_{ij}$	$\beta_j x_{jt}$
138	21	$\beta_i (X_{it} - X_{i7})$	$\alpha_i (x_{it} - x_{7t})$
139	3	$(1-L)(1-L^{12})(X_{it} - X_{i7})$	$\alpha_i (1-L)(1-L^{12})(x_{it} - x_{7t})$
140	22	e $w(L) = 1 - \delta_1 L - \dots$	e $\delta(L) = 1 - \delta_1 L - \dots$
141	1	$w(L) = 0$ e $\delta(L) = 0$ estão	$w(L) = 0$ estão sobre ou fora de círculo unitário e $\delta(L) = 0$ estão fora do círculo unitário.
141	4	$w_0 - w_1 L - \dots - w_\delta L^\delta$	$(w_0 - w_1 L - \dots - w_\delta L^\delta) L^b$
142	5	$(1 - \delta_1 L - \dots - \delta_r L^r)$	$(1 - \delta_1 L - \dots - \delta_r L^r)$
143	13	$(X_t - X)(Y_{t+k} - Y)$	$(X_t - X)(Y_{t+k} - Y)$
149	18	HH	H'H
152	5	pode ser adequado	pode não ser adequado
157	6,7,8,9	B	L
159	1	$B_0 Y = S$	L
164	6,8	ϕ_0, \sum	ϕ_0, \sum
167	3	$(mp, (m-1)(p+q))$	$(mp, (m-1)p+q)$
167	13	modelo	modelo
169	8,9	$(1 - \theta_{11} L), (1 - \theta_{22} L)$	$(1 - \theta_{11} L), (1 - \theta_{22} L)$
174	4	$\Gamma(i-1), \phi_1$	$\Gamma(i-1), \phi_1$
180	18	$c_t(\ell)$	$e_t(\ell)$
194	19	u_t	u_t
195	13	$\text{Corr}(\alpha_t, n_{t+k})$	$\text{Corr}(\alpha_t, n_{t+k})$
197	10	$\Pi_{12j} = 0$	$\Pi_{12j} = 0$
210	5,7	Hosking 1980, Hosking 1980	Hosking 1980a, Hosking 1980b

SÉRIES TEMPORAIS MULTIVARIADAS

Basilio de Bragança Pereira

Instituto de Matemática e Coordenação dos
Programas de Pós-Graduação de Engenharia,
Universidade Federal do Rio de Janeiro.

6º SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA

Rio de Janeiro

1984

PREFÁCIO

A análise de séries temporais tem tido um grande desenvolvimento nos anos recentes e o objetivo deste livro é apresentar os principais resultados sobre modelos lineares para séries temporais.

A ênfase é na análise no domínio do tempo. A análise no domínio da frequência embora abordada no início do livro está sendo apresentada em maior detalhe em outra publicação. O presente trabalho é dividido em três capítulos.

O Capítulo I apresenta uma introdução geral sobre séries temporais indicando os fundamentos probabilísticos e os principais enfoques para análise (domínio do tempo x domínio da frequência) bem como suas justificativas e relacionamento. Paraphraseando Sir Maurice Kendall "esta era a *introdução* que eu gostaria que tivessem colocado em minhas mãos quando eu pela primeira vez me interessei por séries temporais".

O Capítulo II apresenta metodologias para modelagem de séries univariadas. A metodologia de Box-Jenkins, suas modificações e extensões são estudadas. Diversos métodos de identificação são abordados e metodologias alternativas recentes para modelagem são descritas.

O Capítulo III apresenta metodologias para modelagem de séries temporais multivariadas e de séries temporais com informações concomitantes. Neste capítulo apresenta-se a análise dos modelos de: regressão com erros ARMA, efeitos de calendário, análise de intervenção e observações perdidas, função de transferência, séries vetoriais (MARMA ou ARMAV); bem como o estudo de causalidade. Em particular, diferentes enfoques para identificar função de transferência são indicadas. Para séries vetoriais os resultados sobre identificação e as metodologias alternativas do Capítulo II são generalizadas neste capítulo.

Muitos contribuíram para a realização deste trabalho:
Meus agradecimentos

- à Comissão Organizadora do 6 SINAPE pelo convite para ministrar o curso;
- aos meus alunos, principalmente do IM-UFRJ e da COPPE-UFRJ que suportaram minhas experiências com partes do material deste texto em cursos de graduação, mestrado e doutorado;
- aos meus colegas do IM-UFRJ e da COPPE-UFRJ pelo estímulo constante;
- ao IM-UFRJ e COPPE-UFRJ pelo suporte financeiro e pelo ambiente de pesquisa e ensino;
- à Spyros Makridakis (International Forecasting Symposium e Journal of Time Series) e à Oliver Anderson (International Time Series Meetings e Journal of Time Series Analysis) que por suas iniciativas motivaram uma maior comunicação entre analistas de séries temporais e da qual muito tenho me beneficiado;
- às bibliotecárias do CT-UFRJ que gentilmente sempre colaboraram em minhas pesquisas bibliográficas;
- à Cilly Gabrielsen pela datilografia eficiente.

Finalmente, dedico este livro à minha querida Iracema pelo incentivo constante e ao Marcel pela alegria de nosso convívio.

Maio 1984

Basilio de Bragança Pereira

ÍNDICE

PREFÁCIO	i
----------------	---

ÍNDICE	iii
--------------	-----

I - INTRODUÇÃO A TEORIA DE SÉRIES TEMPORAIS

1. Considerações Gerais	1
1.1 - Classificação	2
1.2 - Objetivos	3
1.3 - Enfoques	5
1.4 - Tipos de Variação	6
2. Fundamentos Probabilísticos	9
2.1 - Introdução	9
2.2 - Processos Estocásticos	9
2.3 - Processos Estacionários	14
2.4 - Ergodicidade	17
3. Teoremas Gerais de Representação	20
3.1 - Introdução	20
3.2 - Decomposição Preditiva (Domínio do Tempo)	20
3.3 - Representação Espectral (Domínio da Frequência)	22
3.4 - Outras Considerações sobre os Teoremas de Representação	26
4. Alguns Modelos de Séries Temporais e Econometrias	28

II - MODELOS LINEARES UNIVARIADOS: BOX-JENKINS E EXTENSÕES

1. Introdução	35
2. Modelos Autorregressivos, Médias Móveis e Mistos	41
2.1 - Modelos Autorregressivos - AR(p)	41
2.2 - Estacionariedade e Estabilidade	44
2.3 - Modelos Média Móveis - MA(q)	45
2.4 - Inversibilidade e Identificabilidade	47

2.5 - Modelos Autorregressivos - Médias Móveis - ARMA(p,q)	49
3. Modelos com Média não Nula, Não Estacionários e Sazonais	53
3.1 - Modelos com Média Não Nula	53
3.2 - Modelos Não Estacionários - ARMA(p,d,q) .	53
3.3 - Modelos Sazonais - SARIMA (p,d,q)(P,D,Q) _s	60
4. Agregação	66
4.1 - Introdução	66
4.2 - Agregação de Diferentes Séries	67
4.3 - Agregação Temporal e Observações Perdidas	67
5. Identificação	69
5.1 - Introdução	69
5.2 - Função de Autocorrelação Parcial	70
5.3 - Inferência sobre as Autocorrelações e Au tocorrelações Parciais	73
5.4 - Autocorrelações Inversas	76
5.5 - Arranjos R* e S* e o Método "Corner"	80
5.6 - Função de Autocorrelação Parcial Generali zada e Função de Autocorrelação Estendida	85
5.7 - Critérios Automáticos	92
6. Estimação e Verificação	96
6.1 - Estimação dos Parâmetros	96
6.2 - Verificação da Adequabilidade	102
6.3 - Uso dos Resíduos para Alterar o Modelo ..	105
7. Previsão	106
7.1 - Representações e Notação	106
7.2 - Variância da Previsão e Intervalos de Con fiança	107
7.3 - Cálculo e Atualização das Previsões	108
7.4 - Robustez e Parcimônia	111
7.5 - Combinação de Previsões	112
7.6 - Relação com Outros Métodos	112

8. Outras Metodologias	114
8.1 - Modelagem Recursiva de Hannan-Rissanen ..	114
8.2 - Modelagem ARMA($n, n-1$) de Pandit-Wu	116
8.3 - Modelagem ARARMA(p, q) de Parzen	117
8.4 - Modelagem em Espaço de Estados de Akaike.	119
8.5 - Decomposição por Componentes Principais .	124
9. Apêndice	
A1 - Resumo de Identificação	127
A2 - Correlogramas	128

III - MODELOS LINEARES MULTIVARIADOS

1. Introdução	134
2. Regressão em Séries Temporais	135
2.1 - Regressão com Erros ARMA	135
2.2 - Efeitos de Calendário	140
3. Função de Transferência	140
3.1 - Introdução	140
3.2 - Função de Correlação Cruzada	142
3.3 - Identificação Empírica	144
3.4 - Identificação de Haugh-Box	145
3.5 - Identificação de Box-Jenkins	147
3.6 - Identificação por Mínimos Quadrados e Mé- todo "Corner"	148
3.7 - Estimação, Verificação e Previsão	151
4. Análise de Intervenção	155
4.1 - Modelos para intervenção	155
4.2 - Observações Perdidas	159
5. Modelos Vetoriais ARMA	161
5.1 - Introdução	161
5.2 - Algumas Propriedades	163
5.3 - Modelos com Médias não Nulas e Sazonais .	167
5.4 - Função de Transferência e Análise de In- venção	168
5.5 - Identificação	170

5.6 - Estimaco e Verificaco	178
5.7 - Previso	180
5.8 - Anlise de Autovalores e Autovetores	181
6. Outras Metodologias	185
6.1 - Modelage Recursiva de Hannan - Kavalieris	185
6.2 - Modelagem ARMAV(n,n-1) de Pandit-Wu	185
6.3 - Modelagem em Espaco de Estados de Akaike	186
6.4 - Previso por Componentes Principais	188
7. Causalidade	190
7.1 - Definies	190
7.2 - Caracterizaco	192
7.3 - Testes de Causalidade	197
IV - PROGRAMAS DE COMPUTADOR	200
REFERNCIAS	201

CAPÍTULO I

INTRODUÇÃO A TEORIA DE SÉRIES TEMPORAIS

1. Considerações Gerais

Uma série temporal é um conjunto de observações feitas seqüencialmente no tempo; entretanto, tempo pode ser substituído por qualquer variável como espaço, profundidade etc. As observações vizinhas são dependentes e o estudo de uma série temporal consiste em analisar e modelar esta dependência.

A maior parte dos procedimentos estatísticos foi desenvolvida para analisar observações independentes e, logo, a maioria dos resultados não é relevante para o estudo de séries temporais. Portanto novas técnicas e resultados têm-se desenvolvido para aplicação específica em séries temporais, cujo estudo se constitui uma importante área da estatística. Exemplos ocorrem em vários campos do conhecimento como economia (preços diários de ações, desemprego mensal, exportação mensal), sociologia (criminalidade mensal, greves anuais), medicina (eletrocardiograma, eletroencefalograma), demografia (população anual, nascimentos e mortes mensais), epidemiologia (ocorrências semanais de sarampo), mercadologia (vendas semanais, gastos semanais com propaganda), meteorologia (precipitação pluviométrica, temperatura diária, velocidade do vento), oceanografia (maré horária).

1.1 Classificação

A série temporal é representada por um conjunto de observações $\{Y(t), t \in T\}$ de uma variável Y , onde o conjunto T é um conjunto de índices (tempo, espaço etc.).

Dependendo da natureza de T e de Y a série temporal pode ser

- i) Discreta, quando T é um conjunto finito de pontos, $T = \{1, 2, \dots, T\}$. Por exemplo, o valor exportações mensais de 1970 a 1980. Notação Y_t .
- ii) Contínua, quando T é um intervalo finito, $T = \{t : 0 < t < T\}$. Por exemplo, medições durante dois minutos de um eletrocardiograma. Notação $Y(t)$.
- iii) Multivariada, (discreta ou contínua), $\{Y_1(t), \dots, Y_k(t), t \in T\}$ ou $\{Y(t), t \in T\}$. Por exemplo: vendas (Y_{1t}) e gastos com propaganda (Y_{2t}) semanais de um produto; pressão uterina da parturiente (Y_{1t}) e batidas do coração do feto (Y_{2t}) em trabalho de parto.
- iv) Multidimensional, quando temos $\{Y(t), t \in T\}$ e t é um vetor. Por exemplo, $\{Y(t, r, \ell), t \in T\}$ onde Y é a altura de um ponto do oceano, t o tempo, r a latitude e ℓ a longitude; $\{X(t, b), t \in T\}$ onde X é o número de casos de meningite semanais (t) por bairro (b).

A variável Y observada pode também ser discreta ou contínua; às vezes, a variável é discreta (ex.: população) mas

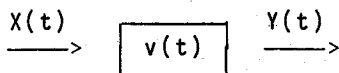
por sua magnitude pode ser considerada contínua.

Em resumo, a variável observada Y pode ser discreta ou contínua, univariada ou multivariada e o índice "tempo" pode ser discreto ou contínuo, unidimensional ou multidimensional. Da forma bastante geral uma série temporal pode ser $\{Y(t), t \in T\}$ onde $Y(t)$ é um vetor $k \times 1$ e t é um vetor $p \times 1$. Dizemos neste caso que a série é multivariada e multidimensional.

Em geral vamos supor t discreto, unidimensional e observado em intervalos equiespaçados.

1.2 Objetivos

Muitas situações em ciências físicas, biológicas e sociais e em engenharia envolvem o conceito de sistema dinâmico, caracterizado por uma série de entrada $X(t)$, uma série de saída $Y(t)$ e uma função de transferência $v(t)$. (Figura 1).



- Sistema Dinâmico

De particular importância são os sistemas lineares, onde a saída é relacionada com a entrada através de um funcional linear envolvendo $v(t)$. Em exemplo típico é

$$Y(t) = \sum_{v=0}^{\infty} v(t)X(t-u)$$

também chamado modelo de função de transferência.

Existem dois principais objetivos para usar séries temporais:

- 1) entender o mecanismo do sistema gerador da série temporal e
- 2) prever o comportamento futuro do sistema.

Para isso, problemas de interesse são:

- a) descrever o comportamento da série; neste caso ferramentas úteis são: a construção de gráficos da série, construção de histogramas e diagramas de dispersão, obtenção de estatísticas descritivas simples, verificação de tendências, ciclos e variações sazonais, pontos de mudanças, observações espúrias, etc.
- b) investigar o mecanismo gerador da série temporal, procurar periodicidades relevantes nos dados.
- c) fazer previsões de valores futuros da série a partir de valores passados.
- d) estimar a função de transferência $v(t)$, conhecendo-se as séries de entrada $X(t)$ e saída $Y(t)$.
- e) fazer previsões da série $Y(t)$, com o conhecimento de observações da série $X(t)$ e de $v(t)$.
- f) estudar o comportamento do sistema, simulando-se a série de entrada.

g) controlar a série de saída $Y(t)$, de modo a trazê-la o mais próximo possível de um valor desejado, ajustando-se convenientemente a série de entrada $X(t)$; este controle é necessário, devido às perturbações que normalmente afetam um sistema.

1.3 Enfoques

Hã dois aspectos a considerar no estudo de séries temporais: análise e modelagem.

A principal razão para modelar uma série temporal é tornar possível previsões de valores futuros. Como em qualquer análise de dados, procura-se obter um modelo matemático que descreva o sistema de maneira parcimoniosa para o objetivo desejado. Sua forma funcional deve ser simples e o número de parâmetros deve ser mínimo; além disso o modelo deve, se possível, ser parametrizado de tal forma que cada parâmetro possa ser interpretado facilmente e identificado com algum aspecto da realidade. Finalmente sua forma funcional deve permitir uma fácil manipulação matemática, necessária para inferência a seu respeito.

O objeto da análise da série temporal é sumarizar as propriedades da série e caracterizar seu comportamento identificando ou sugerindo um modelo adequado. Hã basicamente dois enfoques usados na análise de séries temporais.

No primeiro enfoque, a análise é feita no domínio do

tempo, isto é, estamos interessados na magnitude de eventos que ocorrem em determinado instante de tempo e na relação entre observações em diferentes instantes de tempo. A ferramenta utilizada é a função de autocorrelação (e certas funções da mesma) e a análise é baseada em um modelo paramétrico.

No segundo enfoque, a análise é feita no domínio da frequência, isto é, estamos interessados na frequência com que certos eventos ocorrem em um período de tempo (movimentos cíclicos por exemplo). A ferramenta utilizada é o espectro (que é uma transformada da função de autocorrelação) e a análise é baseada em modelos não paramétricos.

As duas formas de análise não são alternativas, mas sim complementares, mostrando cada uma diferentes aspectos da natureza da série temporal. Ambos os enfoques, como veremos, são justificados por teoremas de representação (Wold e Cramer, respectivamente) e podemos dizer que a análise no domínio da frequência é conveniente na procura de características determinísticas da série temporal enquanto a análise no domínio do tempo é usada na análise de processos não determinísticos.

1.4 Tipos de Variação

Métodos tradicionais de análise de séries temporais tratam principalmente da decomposição da série em: tendência, variação sazonal, ciclos e outras flutuações irregulares. O interesse por este enfoque tem se reativado recentemente.

Os diferentes tipos de variação que se procura analisar são:

- a) Efeito sazonal - muitas séries exibem uma variação periódica. (diária, semanal, anual, etc.). Este tipo de variação é fácil de entender e pode ser medida e extraída dos dados, fornecendo uma série dessazonalizada. Por exemplo, medidas de temperatura exibem uma variação anual periódica, a baixando durante o inverno e aumentando no verão.
- b) Outros efeitos cíclicos - além de efeitos sazonais, algumas séries exibem outras variações de período fixo. Por exemplo, as variações de temperatura ao longo do dia. Além disso, outras séries apresentam variações cujo período não é fixo, porém são possíveis de ser previstos. Por exemplo, dados econômicos podem ser afetados pelos chamados ciclos econômicos com períodos variando de 5 a 7 anos, segundo algumas teorias econômicas. Dados de vazões de rios parecem obedecer ciclos com períodos em torno de 50 anos e dados de ocorrência de surtos de sarampo obedecem a períodos de 2,5 a 3 anos. Basicamente variações cíclicas são variações periódicas com período sujeito a pequenas variações, em contraste com sazonalidade que tem período constante.
- c) Tendência - são os efeitos a longo prazo na média. A dificuldade é definir longo prazo. Por exemplo, variáveis climáticas às vezes exibem variações com períodos de 50 anos. Se tivermos 20 anos de dados esta variação de longo prazo será confundida com tendência. Porém se tivermos centenas de anos o ciclo será visível. Outro exemplo seria o de da

dos sobre desemprego mensal observado durante 2 anos, supondo a existência de ciclos de 7 anos na economia. Novamente confundiríamos a variação de longo prazo com tendência.

- d) Flutuação irregular - após remover a sazonalidade, ciclo e tendência a série fica reduzida a uma série de resíduos livre de qualquer variação regular. Esta série de resíduos é então modelada usando algum modelo probabilístico em que a noção de dependência é explorada.

2. Fundamentos Probabilísticos

2.1 Introdução

Os sistemas em que estamos interessados são sistemas que se desenvolvem no tempo (ou espaço) sujeitos a variações que podem ser descritas por leis probabilísticas e esse tipo de sistema é o objeto de estudo da teoria dos processos estocásticos.

Na teoria estatística de séries temporais o mecanismo gerador é considerado como um processo estocástico e a série temporal observada é uma das realizações possíveis. Logo os conceitos de processos estocásticos e série temporal são análogos respectivamente aos conceitos de população e amostra na Estatística Elementar. O que distingue a análise de séries temporais de outras análises estatísticas é o reconhecimento explícito da importância da ordem em que as observações são feitas, isto é, da dependência entre as observações.

2.2 Processos Estocásticos

Definimos um processo estocástico como uma família $\{Y(t), t \in T\}$ tal que para cada $t \in T$, $Y(t)$ é uma variável aleatória. O conjunto T é arbitrário. O conjunto dos valores $\{Y(t), t \in T\}$ é chamado espaço de estados do processo e os valores y_t são chamados estados.

Como vemos através da Figura 11 o conceito é uma

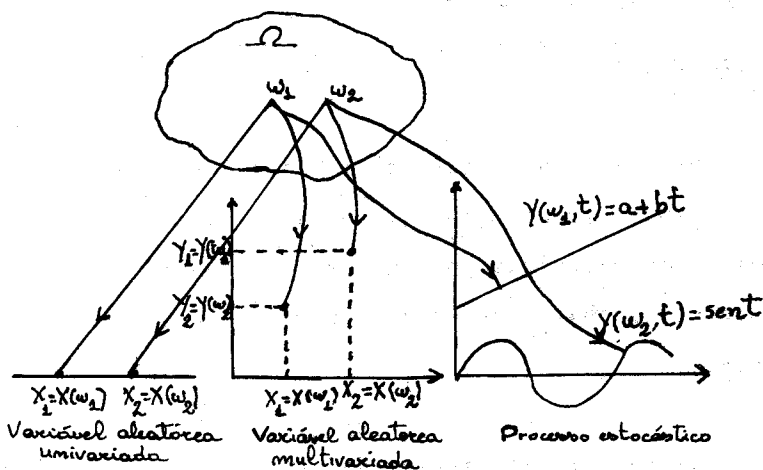


Fig. 11 - Extensão do conceito de variável aleatória para processos estocásticos. Ω - espaço amostral do experimento.

extensão do conceito de variável aleatória multivariada para dimensão infinita. Essa extensão foi possível devido a um teorema de Kolmogorov que é englobado na seguinte definição mais completa:

Uma família de variáveis aleatórias $\{Y(t), t \in T\}$ constitui um processo estocástico se

- a) as variáveis aleatórias $Y(t_1), \dots, Y(t_n)$ são bem definidas, isto é, têm funções de distribuição finito-dimensionais

$$F(y_1, \dots, y_n; t_1, \dots, t_n) = P\{Y(t_1) \leq y_1, \dots, Y(t_n) \leq y_n\}$$

conhecidas para todo $n \geq 1$. Isto significa que para $n = 1$

temos as distribuições unidimensionais da variável aleatória $Y(t_1)$, $t_1 \in T$; para $n=2$ temos as distribuições bidimensionais da variável aleatória $(Y(t_1), Y(t_2))$, $t_1, t_2 \in T$ e assim por diante.

b) estas funções de distribuição satisfazem as condições de

- simetria: tem o mesmo valor quando y_j e t_j são submetidos a mesma permutação

$$\text{(ex.: } F(y_1, y_2, y_3; t_1, t_2, t_3) = F(y_2, y_1, y_3; t_2, t_1, t_3)\text{)}.$$

- consistência:

$$\lim_{y_n \rightarrow \infty} F(y_1, \dots, y_n; t_1, \dots, t_n) = F(y_1, \dots, y_{n-1}; t_1, \dots, t_{n-1}).$$

Entretanto, esta definição não é operacional na prática pois o conhecimento das distribuições é muito difícil de ocorrer na prática, senão impossível. Felizmente para o tipo de processo que vai nos interessar bastará nos restringirmos aos momentos de 1ª e 2ª ordem destas distribuições, isto é,

i) a função média de $Y(t)$,

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

ii) a função de auto-covariância de $Y(t)$,

$$\begin{aligned} C(t_1, t_2) &= E\{[Y(t_1) - \mu(t_1)][Y(t_2) - \mu(t_2)]\} = \\ &= E\{Y(t_1)Y(t_2)\} - \mu(t_1)\mu(t_2). \end{aligned}$$

Em particular

$$C(t, t) = \sigma^2(t) = \text{Var}\{Y(t)\} = E\{Y^2(t)\} - \mu^2(t)$$

além disso como $C(t_1, t_2)$ depende da unidade de $Y(t)$, na

maioria das vezes utilizamos a

iii) função de auto-correlação de $Y(t)$

$$\rho(t_1, t_2) = \frac{C(t_1, t_2)}{\sigma(t_1)\sigma(t_2)}$$

que independe da unidade de medida.

Observe que se as distribuições finito-dimensionais de $\{Y(t), t \in T\}$ são normais, isto é, o processo é Gaussiano, o conhecimento de $\mu(t)$ e $C(t_1, t_2) \forall t, t_1, t_2$ caracteriza toda a distribuição do processo.

Como vemos para cada t , $Y(t)$ é uma variável aleatória e esta interpretação é ilustrada na Figura I2a

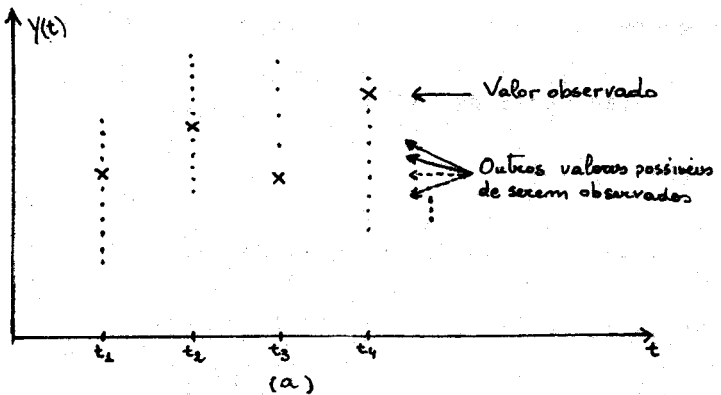


Fig. I2a - Valores observados (X) e valores possíveis de serem observados (·) do processo estocástico $Y(t)$.

ou mais genericamente na Fig. I2b

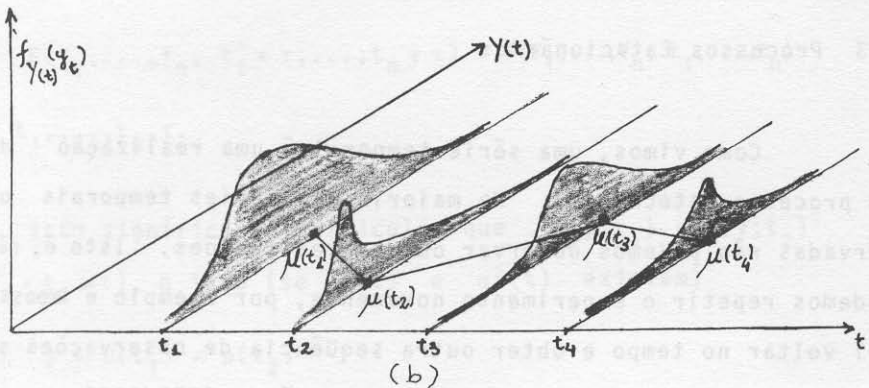


Fig. 12b - Família de variáveis aleatórias e suas densidades.

onde $f_{Y(t)}(y_t)$ é a função de densidade de $Y(t)$ para t fixado. É possível que as funções de densidade sejam diferentes para instantes de tempo distintos t_1 e t_2 .

Por outro lado podemos considerar o processo em função de t , ou seja, uma realização ou trajetória do processo estocástico. Designando por $Y^{(1)}(t), Y^{(2)}(t), \dots$ etc as realizações, o conjunto de todas as possíveis realizações denomina-se "ensemble". A Figura 13 ilustra esta interpretação.

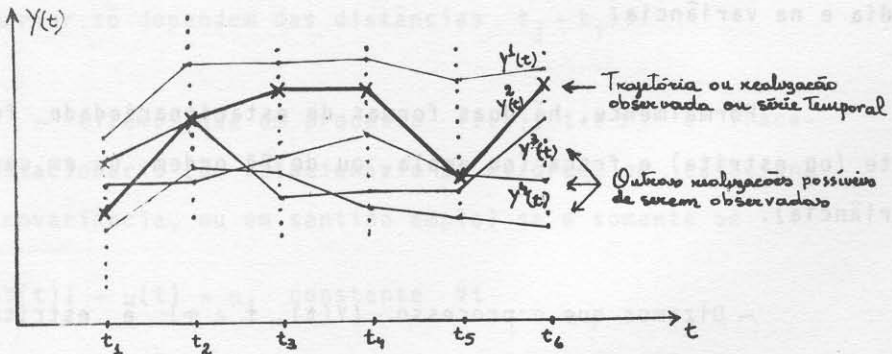


Fig. 13 - Ensemble de um processo estocástico.

2.3 Processos Estacionários

Como vimos, uma série temporal é uma realização de um processo estocástico. Na maioria das séries temporais observadas não podemos observar outras realizações. Isto é, não podemos repetir o experimento novamente, por exemplo é impossível voltar no tempo e obter outra seqüência de observações sobre exportações mensais do Brasil no período 1970/1983.

Portanto a cada instante de tempo t , em geral temos somente uma observação $y(t)$ de $Y(t)$. Se desejamos por exemplo saber alguma coisa sobre $\mu(t)$ ou $\sigma^2(t)$ fica impossível obter estas informações com apenas uma observação $y(t)$, a menos que se faça alguma suposição adicional sobre o tipo de processo analisado.

A primeira suposição normalmente feita é a estacionariedade. Intuitivamente um processo é estacionário se ele desenvolve-se no tempo de modo que a origem dos tempos não é importante ou ainda se não existe mudanças sistemáticas na média e na variância.

Formalmente, há duas formas de estacionariedade forte (ou estrita) e fraca (ou ampla, ou de 2ª ordem, ou em covariância).

— Dizemos que o processo $\{Y(t), t \in T\}$ é estritamente estacionário se suas funções de distribuições finito-dimensionais são invariantes por translações no tempo, isto é

$$F(Y_1, \dots, Y_n, t_1 + \tau, \dots, t_n + \tau) = F(Y_1, \dots, Y_n; t_1, \dots, t_n)$$

para $\forall t_1, \dots, t_n, \tau$.

Isto significa em particular que $F(y; t_1) = F(y; t_2)$

(pois $t_2 = t_1 + \tau$) e logo (se $\mu(t)$ e $\sigma^2(t)$ existem)

$$\mu = \mu(t_1) = \mu(t_2)$$

$$\sigma^2 = \sigma^2(t_1) = \sigma^2(t_2)$$

isto é, a média ($\mu(t) = \mu$) e variância ($\sigma^2(t) = \sigma^2$) do processo são constantes.

Também tem-se $F(y_1, y_2; t_1, t_2) = F(y_1, y_2, t_1 + \tau, t_2 + \tau)$

logo $C(t_1, t_2) = C(t_1 + \tau, t_2 + \tau)$ e fazendo $\tau = -t_1$ tem-se

$$C(t) = C(0, t_2 - t_1) \text{ para } t = t_2 - t_1$$

isto é, a função de autocovariância (autocorrelação) só depende da distância $t = t_2 - t_1$.

(Mais genericamente, mostra-se que os momentos de ordem superior são dependentes das distâncias $t_j - t_1$).

— Dizemos que um processo $\{Y(t), t \in T\}$ é fracamente estacionário (ou estacionário de 2ª ordem ou estacionário em covariância, ou em sentido amplo) se e somente se

i) $E\{Y(t)\} = \mu(t) = \mu$, constante $\forall t$

ii) $\text{Var}\{Y(t)\} = \sigma^2(t) = \sigma^2$ é finita (e constante) $\forall t$

$$\text{iii) } C(t_1, t_2) = C(t_2 - t_1) = C(\tau) \quad (\tau = t_2 - t_1)$$

Como mencionado anteriormente, estaremos interessados apenas nos momentos de 1ª e 2ª ordem e, portanto, a partir de agora, somente trataremos dos processos fracamente estacionários, que denominaremos somente de estacionários.

Note que se os momentos existem para um processo, esta estacionariedade forte implica estacionariedade fraca. Entretanto estacionariedade fraca não implica estacionariedade forte se o processo é Gaussiano, isto é, quando as distribuições finito-dimensionais são normais.

É claro que nem todas as séries temporais são realizações de processos estacionários, o que se faz é transformar a série em estacionária e utilizar a teoria dos processos estacionários para a série transformada.

Finalmente, sendo $C(\tau)$ a função de autocovariância de um processo estacionário $Y(t)$, as seguintes propriedades são satisfeitas:

$$\text{a) } C(0) \geq 0$$

$$\text{b) } C(-\tau) = C(\tau)$$

$$\text{c) } |C(\tau)| \leq C(0)$$

d) $C(\tau)$ é positiva definida no sentido que

$$\sum_{j=1}^n \sum_{k=1}^n a_j a_k C(\tau_j - \tau_k) \geq 0$$

quaisquer reais a_1, \dots, a_n e t_1, \dots, t_n .

A função de autocorrelação de $Y(t)$ é definida por

$$\rho(\tau) = \frac{C(\tau)}{C(0)}$$

e propriedades análogas a a) - d) valem para $\rho(\tau)$.

2.4 Ergodicidade

Usualmente, em problemas de estimação estatística, temos disponível diversas repetições (observações) de um mesmo fenômeno. Portanto, como estimativa da média (μ) de uma variável aleatória Y , tomamos a média amostral dos valores observados y^1, y^2, \dots, y^n . Segundo este procedimento, para estimarmos a média $\mu(t_0)$ e a função de autocovariância (ou autocorrelação), $C(\tau) = C(t_1, t_2)$ ($\rho(\tau) = \rho(t_1, t_2)$) teríamos que fazer diversas observações $y^1(t_0), y^2(t_0), \dots, y^n(t_0)$ e $(y^1(t_1), y^1(t_2)), (y^2(t_1), y^2(t_2)) \dots (y^n(t_1), y^n(t_2))$ e tomar a média amostral e a autocovariância amostral. A Figura I4 ilustra esse procedimento

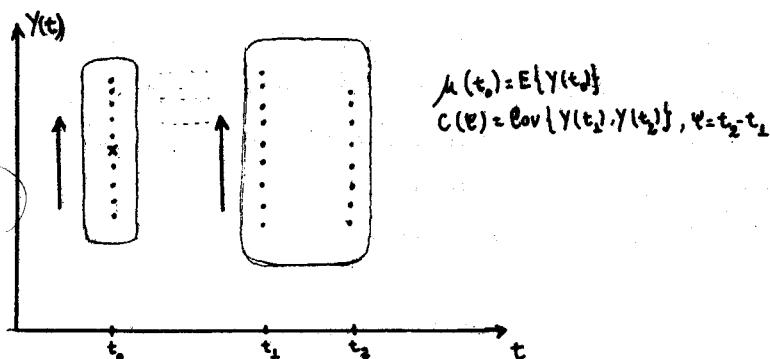


Fig. I4 - Médias estatísticas, tomadas na direção das setas. • Valores observados.

Entretanto, como já foi apontado, na prática geralmente só temos uma observação da série temporal e não podemos estimar os parâmetros da forma indicada, neste caso.

Porém, com a série $Y(1), Y(2) \dots Y(T)$, podemos calcular as seguintes médias no tempo (ou em fase)

$$m = \frac{1}{T} \sum_{t=1}^T Y(t) \quad \text{e} \quad c(\tau) = \frac{1}{T} \sum_{t=1}^T Y(t)Y(t+\tau)$$

ilustrados na Figura I5

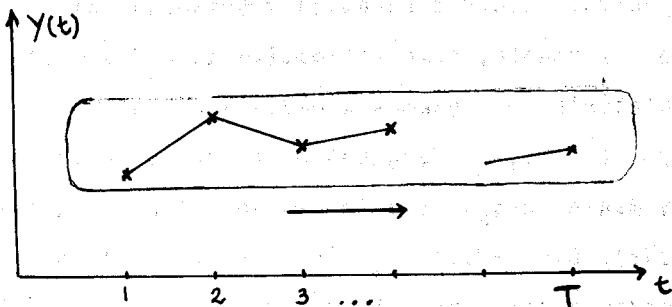


Fig. I5 - Médias em fase, tomadas na direção da seta.

De fato, a importância prática da análise de autocorrelações de processos estacionários é em grande parte devida ao fato de que, se $Y(t)$ é um processo estacionário, sua média μ e sua função de autocovariância $C(\tau)$ podem ser estimadas usando somente uma realização de $Y(t)$. Isto ocorre se $Y(t)$ é ergódico, isto é, satisfaz os chamados teoremas ergódicos.

Se temos um processo estacionário e só podemos observar uma realização, podemos fazer a suposição que o processo

é ergódico já que neste caso não há uma maneira operacional de verificar ergodicidade. Além disso a maioria dos processos estacionários observados na prática são ergódicos (Yaglom, 1962, pg. 16).

Na realidade o que os teoremas em geral requerem é que as variáveis $Y(t)$ e $Y(t+k)$ se tornem independentes quando k aumenta. Isto é uma suposição bastante razoável na prática; isto é as observações irem se tornando independentes quanto mais afastadas se encontrem. Neste caso

$$\mu = \mu(t) = m \text{ e } C(t) = \text{Cov}(Y(t_1), Y(t_2)) = c(\tau).$$

3. Teoremas Gerais de Representação

3.1 Introdução

Em trabalhos teóricos e aplicados existe um certo dualismo entre os dois enfoques usados na análise de séries temporais. Este dualismo é devido aos objetivos da análise enquanto do ponto de vista formal os dois enfoques são fortemente relacionados. Nesta seção apresentam-se informalmente sem demonstração e sem muito rigor alguns resultados que fornecem a justificativa de cada enfoque e seu relacionamento. Nos limitaremos ao caso de processos discretos no tempo.

3.2. Decomposição Preditiva (Domínio do Tempo)

- Teorema de Wold - Todo processo estacionário Y_t admite a representação

$$Y_t = D_t + Z_t$$

onde

- D_t e Z_t são não correlacionados
- D_t é determinístico, no sentido de que valores futuros de D_t são previstos com precisão baseado em função linear de valores passados D_{t-1}, D_{t-2}, \dots
- Z_t é um processo linear discreto geral (ou média móvel infinito, $MA(\infty)$)

$$Z_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots$$

onde $E\{\epsilon_t\} = 0$ e $E\{\epsilon_t \epsilon_{t+k}\} = 0 \forall t, k$, isto é ϵ_t é um ruído branco.

Observe que a decomposição não exclui a possibilidade de um mecanismo não linear melhor representar o processo. A decomposição significa que é razoável esperar que o verdadeiro mecanismo do processo possa ser adequadamente representado por um mecanismo que gera um tipo simples de processo determinístico mais um resíduo Z_t , que veremos como aproximar.

Esta decomposição é de grande importância prática, pois muitas séries temporais, particularmente séries econômicas, contêm componentes de tendência e sazonalidade. Um modelo clássico envolvendo estas componentes supõe que a variável econômica é representada por

$$Y_t = D_t + Z_t = T_t + S_t + Z_t$$

onde T_t é a componente determinística representando a tendência, S_t é a componente determinística representando a sazonalidade e Z_t é a componente estocástica.

A decomposição de Wold mostra a importância dos processos lineares discretos em geral e indica a razão da grande utilização desses modelos, popularizados e tornados acessíveis por Box e Jenkins (1976).

3.3 Representação Espectral (Domínio da Frequência)

A análise de uma série temporal no domínio da frequência é baseada no seguinte par de teoremas.

A) Teorema de Bochner-Khintchin - A função de autocorrelação $\rho(\tau)$ de um processo estacionário discreto no tempo admite a representação

$$\rho(\tau) = \int_{-\pi}^{\pi} e^{i\omega\tau} dF(\omega) = \int_0^{\pi} [\text{sen } \omega\tau dQ(\omega) + \text{cos } \omega\tau dC(\omega)]$$

onde $F(\omega)$ tem as propriedades de uma função de distribuição de probabilidade no intervalo $(-\pi, \pi)$, isto é, $F(-\pi) = 0$ e $F(\pi) = 1$. A notação $dF(\omega)$ significa integração ou soma conforme $F(\omega)$ seja contínuo ou uma função escada. A função $F(\omega)$ é denominada função de distribuição espectral.

Como $\rho(\tau)$ tem a propriedade de ser positiva-definida para demonstrar a representação basta usar um teorema da análise devido a Bochner que afirma que uma função $\rho(\tau)$ é positiva-definida se e somente se tem a representação acima.

Lembrando conceitos elementares de probabilidades vemos que a função de autocorrelação $\rho(\tau)$ funciona como a função característica da distribuição $F(\omega)$, e a função de distribuição espectral $F(\omega)$ portanto admite a decomposição de Lebesgue

$$F(\omega) \approx a_1 F_1(\omega) + a_2 F_2(\omega)$$

onde $F_1(\omega)$ e $F_2(\omega)$ são funções de distribuição do seguinte tipo:

- a) $F_2(\omega)$ é contínua
- b) $F_1(\omega)$ é uma função escada com saltos p_r nos pontos ω_r , $r = 1, 2, \dots$ e $\sum p_r = 1$.

A Figura 16 ilustra uma distribuição espectral na forma geral

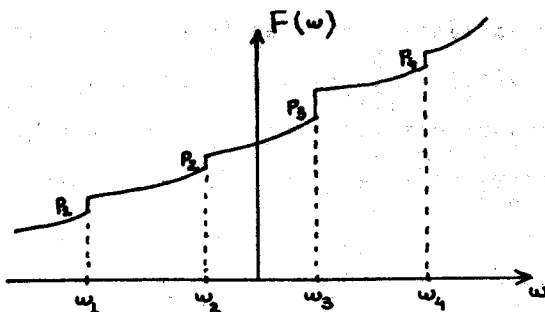


Fig. 16 - Forma geral da distribuição espectral.

Como no caso de distribuições de probabilidade a maioria dos problemas práticos envolve apenas uma das componentes, sendo somente contínua ou somente discreta conforme a variável aleatória era contínua ou discreta respectivamente. O mesmo ocorre para distribuições espectrais, entretanto neste caso ocorre mais freqüentemente o caso de ambas as componentes presentes.

Vejamos os seguintes casos:

CASO 1 - $a_1 = 1$ e $a_2 = 0$

Aqui $F(\omega) = F_1(\omega)$ e o espectro consiste apenas de componentes discretos e dizemos que temos um espectro de linha. O exemplo principal é o Processo Harmônico:

$$Y_t = \sum_{i=1}^k A_i \cos(\omega_i t + \phi_i) = \sum_{i=1}^k (\alpha_i \sin \omega_i t + \beta_i \cos \omega_i t)$$

onde k_i , A_i e ω_i são constantes e ϕ_i são variáveis aleatórias uniformes em $(-\pi, \pi)$; alternativamente α_i e β_i são variáveis aleatórias normais não correlacionadas entre si e com média 0 e variância σ_i^2 . Neste caso verifica-se que $E(Y_t) = 0$ e $V(Y_t) = \sum_{i=1}^k \sigma_i^2$. Portanto a variância de Y_t é decomposta na soma das variâncias de cada componente harmônico (de uma forma análoga à análise de variância clássica).

CASO 2 - $a_1 = 0$ e $a_2 = 1$

Aqui $F(\omega) = F_2(\omega)$ e o espectro consiste apenas do componente contínuo. Neste caso $dF(\omega) = f(\omega)d\omega$ chama-se função de densidade espectral. Os seguintes modelos que serão o principal objeto de estudo deste livro são exemplos de processos com espectro contínuo:

Processo Aleatório Puro ou Ruído Branco

Processo Linear Discreto

Processo Autorregressivo - AR

Processo Média Móvel - MA

Processo Misto Autorregressivo-Média Móvel - ARMA

CASO 3 - $a_1 > 0$ e $a_2 > 0$

Neste caso, $F(\omega)$ é uma mistura de ambas as componentes contínua e discreta. Este caso corresponde ao caso de uma variável aleatória com uma distribuição de probabilidade mista. Um exemplo de tal processo é obtido tomando

$$Y_t = D_t + Z_t$$

onde D_t e Z_t são não correlacionados, D_t tem espectro contínuo e Z_t tem espectro discreto.

B) Teorema de Cramer - Todo processo estacionário discreto no tempo com função de distribuição espectral $F(\omega)$ admite a representação

$$Y_t = \int_{-\pi}^{\pi} e^{i\omega t} dZ(\omega) = \int_0^{\pi} [\text{sen } \omega t \alpha(\omega) d\omega + \text{cos } \omega t \beta(\omega) d\omega]$$

onde $\alpha(\omega)$ e $\beta(\omega)$ são funções estocásticas com propriedades análogas a α_i e β_i do processo harmônico. Da mesma forma a variância de Y_t é decomposta na soma $\int_{-\pi}^{\pi} dF(\omega) = F(\pi)$. Esta representação é uma extensão do Processo Harmônico

$$Y_t = \sum_{i=1}^k [\alpha_i \text{sen } \omega_i \tau + \beta_i \text{cos } \omega_i \tau]$$

ao fazermos $k \rightarrow \infty$ e permitirmos a inclusão de número infinito não enumerável de termos trigonométricos no intervalo $[0, \pi]$. Esta extensão permite analisar funções estocásticas, isto é, processos não determinísticos. A análise espectral trata do estudo da estrutura da série identificando componentes (senões) importantes de Y_t através da análise e estimação de $F(\omega)$.

3.4 Outras Considerações Sobre os Teoremas de Representação

É interessante apontar explicitamente as relações e distinções que observamos nos dois tipos de decomposição de uma série temporal.

- i) Da expressão $\rho(\tau) = \int_{-\pi}^{\pi} e^{i\omega\tau} dF(\omega)$ observamos que há uma correspondência um a um entre $\rho(t)$ e $F(\omega)$; logo, toda a informação sobre o processo contida em $\rho(\tau)$ está contida em $F(\omega)$ e vice versa.
- ii) Na decomposição de Wold temos $Y_t = D_t + Z_t$ e na decomposição da distribuição espectral temos $F(\omega) \sim F_1(\omega) + F_2(\omega)$; a componente determinística D_t corresponde a função espectral $F_1(\omega)$ e a componente não determinística corresponde à função contínua $F_2(\omega)$.
- iii) Mostra-se que a parte determinística na decomposição de Wold é um processo harmônico, isto é,

$$D_t = \sum_{i=1}^K (\alpha_i \sin \omega_i t + \beta_i \cos \omega_i t)$$

correspondendo ao espectro de linha $F_1(\omega)$. Mostra-se, também, que dada uma função de distribuição espectral contínua $F_2(\omega)$ sempre podemos encontrar um processo média móvel com distribuição espectral $F_2(\omega)$.

- iv) Há uma distinção de tratamento da componente estocástica nos dois enfoques. Na decomposição de Wold a componente estocástica é representada parametricamente por meio da sequência ψ_1, ψ_2, \dots . Na representação espectral a com

ponente estocástica tem uma representação não paramétrica em termos da função $F_2(\omega)$.

- v) A decomposição de Wold foi estendida por Cramer em 1962 (ver Cramer e Leadbetter, 1966, pg.107) para o caso de processos não estacionários. A modificação sendo que os coeficientes ψ_1, ψ_2, \dots passam a depender do tempo, isto é,

$$Z_t = \varepsilon_t + \psi_1(t)\varepsilon_{t-1} + \psi_2(t)\varepsilon_{t-2} + \dots$$

A análise espectral de processos não estacionários foi estudada inicialmente por Priestley em 1966 (ver Priestley, 1981, Cap. 11).

- vi) As decomposições de Wold e de Cramer bem como a expansão de Karhunen-Loeve (ver Fukunaga, 1972, Cap. 8) podem ser vistas como casos particulares do método de expansão canônica devido a Pugachev, em 1957, e estudado elegantemente por Parzen, em 1961, usando conceitos de espaços de Hilbert. Neste método tenta-se expressar qualquer processo estocástico (função aleatória) em termos de combinações lineares de variáveis aleatórias não correlacionadas.

Nesta abordagem, uma série temporal é tratada como um vetor de um espaço infinito-dimensional e pode ser convenientemente estudada pela sua decomposição ortogonal ou canônica, expressa como soma ou integral. No domínio do tempo os eixos são variáveis aleatórias não correlacionadas enquanto os coeficientes formam uma função determinística; no domínio da frequência os eixos são funções determinísticas ortogonais tais como $e^{i\lambda t}$ e os coeficientes são variáveis aleatórias.

4. Alguns Modelos de Séries Temporais e Econométricos

Nesta seção listaremos alguns modelos especiais bem como apresentaremos algumas notações. Será útil escrever os modelos na forma geral

$$\text{dados} = \text{sinal} + \text{ruído}$$

i.e.,

$$Y_t = \delta_t + \varepsilon_t.$$

Antes de apresentarmos os modelos, introduziremos a notação de operadores de defasagem ou de lag, que facilitam bastante a manipulação dos modelos e que obedecem às regras algébricas:

— Operador de defasagem, retardo ou translação para o passado

$$Ly_t = y_{t-1}, L^m y_t = y_{t-m}$$

— Operador de diferenças

$$\Delta y_t = y_t - y_{t-1} = (1 - L)y_t, \Delta = (1 - L)$$

— Operador de translação para o futuro

$$Fy = L^{-1}y_t = y_{t+1}, L^{-m}y_t = y_{t+m}$$

— Operador soma

$$\begin{aligned} Sy_t &= \sum_{j=0}^{\infty} Y_{t-j} = y_t + y_{t-1} + \dots = (1 + L + L^2 + \dots)y_t = \\ &= (1 - L)^{-1}y_t = \Delta^{-1}y_t. \end{aligned}$$

Representaremos em geral por $a(L)$ o polinômio de

defasagem

$$a(L) = \sum_{j=0}^m a_j L^j = a_0 + a_1 L + a_1 L^2 + \dots + a_m L^m.$$

Os operadores de defasagem operam linearmente nas variáveis como por exemplo

$$\begin{aligned} \phi(L)Y_t &= (\phi_0 + \phi_1 L + \dots + \phi_p L^p)Y_t = \\ &= \phi_0 Y_t + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} = \sum_{j=0}^p \phi_j Y_{t-j} \end{aligned}$$

e são manipulados como polinômios na variável L para efeito de multiplicação, divisão e inversão.

Por exemplo

(multiplicação)

$$Y_t = (\phi_0 + \phi_1 L)X_t \quad \text{e} \quad X_t = (\theta_0 + \theta_1 L)\varepsilon_t$$

$$\text{então} \quad Y_t = (\phi_0 + \phi_1 L)(\theta_0 + \theta_1 L)\varepsilon_t =$$

$$= \phi_0 \theta_0 \varepsilon_t + (\phi_1 \theta_0 + \phi_0 \theta_1) \varepsilon_{t-1} + \phi_1 \theta_1 \varepsilon_{t-2}$$

(divisão)

$$\phi(L)Y_t = \varepsilon_t \quad \text{então} \quad Y_t = \phi^{-1}(L)\varepsilon_t$$

onde $\phi^{-1}(L)$ é o inverso de $\phi(L)$

(inversão)

Se $\phi(L)$ é um polinômio, $\psi(L)$ seu inverso é obtido de

$$\phi(L)\psi(L) = 1$$

isto é,

$$(\phi_0 + \phi_1 L + \phi_2 L^2 + \dots)(\psi_0 + \psi_1 L + \psi_2 L^2 + \dots) = 1 + 0L + 0L^2 + \dots$$

logo $\phi_0 \psi_0 = 1 \implies \psi_0 = \frac{1}{\phi_0}$

$$(\phi_0 \psi_1 + \phi_1 \psi_0)L = 0 \implies \psi_1 = -\frac{\phi_1}{\phi_0^2}$$

$$(\phi_0 \psi_2 + \phi_1 \psi_1 + \phi_2 \psi_0)L^2 = 0 \implies \psi_2 = \frac{\phi_1^2}{\phi_0^3} - \frac{\phi_2}{\phi_0^2}$$

⋮

i) Processo Aleatório Puro ou Ruído Branco

$$Y_t = \delta_t + \varepsilon_t$$

$$\delta_t = \mu$$

onde o conjunto dos ε_t é uma seqüência de variáveis aleatórias não correlacionadas. Para estacionariedade de Y_t temos $E\{Y_t\} = \mu$ e $E\{Y_t - \mu\}^2 = \sigma^2$. Em geral $\mu = 0$ e a seqüência dos ε_t é o ruído branco. Adotamos esta notação nos exemplos a seguir.

ii) Modelo Linear Discreto Geral - MA(∞)

$$Y_t = \delta_t + \varepsilon_t$$

$$\delta_t = \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots = \psi(L) \varepsilon_t$$

onde ε_t é um ruído branco.

iii) Modelo Autorregressivo de Ordem p - AR(p)

$$Y_t = \Delta_t + \varepsilon_t$$

$$\Delta_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} = \phi^*(L)Y_t$$

iv) Modelo de Médias Móveis de Ordem q - MA(q)

$$Y_t = \Delta_t + \varepsilon_t$$

$$\Delta_t = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} = \theta^*(L)\varepsilon_t$$

v) Modelo Autorregressivo - Média Móvel - ARMA(p,q)

$$Y_t = \Delta_t + \varepsilon_t$$

$$\Delta_t = \phi^*(L)Y_t + \theta^*(L)\varepsilon_t$$

vi) Modelos de Regressão ou de Variáveis Exógenas

$$Y_t = \Delta_t + \varepsilon_t$$

$$\Delta_t = X_t \beta$$

X_t é uma seqüência de tendências determinísticas ou de defasagem

ex.: constante $Y_t = \mu + \varepsilon_t$

linear $Y_t = \alpha + \beta t + \varepsilon_t$

polinomial $Y_t = \sum_{j=0}^m \beta_j t^j + \varepsilon_t$

regressão $Y_t = \alpha + \beta x_t + \varepsilon_t$

harmônico $Y_t = \sum_{j=1}^p \{\alpha_j \cos \lambda_j t + \beta_j \sin \lambda_j t\} + \varepsilon_t$

$$\lambda_t = \frac{2\pi j}{p} \quad \text{se o período é } p.$$

$$\text{defasagem } Y_t = \alpha + \beta_1 x_t + \beta_2 x_{t-1} + \varepsilon_t$$

$$Y_t = \beta_1 x_t + \beta_2 x_{t-1} + \dots + \varepsilon_t = B(L)x_t + \varepsilon_t$$

vii) Modelo Autorregressivo Exógeno - ARX

$$Y_t = \delta_t + \varepsilon_t$$

$$\delta_t = A^*(L)Y_t + X_t' \beta.$$

viii) Modelo ARMA + Exógeno - ARMAX

$$Y_t = \delta_t + \varepsilon_t$$

$$\delta_t = A^*(L)Y_t + \theta^*(L)\varepsilon_t + X_t' \beta.$$

A vantagem deste modelo comparado com os modelos de função de transferência é sua linearidade.

ix) Modelo de Função de Transferência Ruído Branco

$$Y_t = \delta_t + \varepsilon_t$$

$$\delta_t = \frac{1}{1 + A^*(L)} B^*(L)X_t$$

x) Modelo Função de Transferência Ruído ARMA

$$Y_t = \delta_t + v_t$$

$$\delta_t = \frac{1}{1 + A^*(L)} B^*(L)X_t$$

$$v_t = \frac{1 + \theta(L)}{1 + \phi^*(L)} \varepsilon_t$$

xi) Modelo de Espaço de Estados ou de Markov

$$Y_t = \delta_t + v_t$$

$$\Delta_t = X_t$$

$$X_{t+1} = FX_t + w_t$$

onde v_t e w_t são ruídos branco com

$$E\left\{\begin{pmatrix} w_t \\ v_t \end{pmatrix} \begin{pmatrix} w_t v_t \end{pmatrix}\right\} = \begin{bmatrix} Q & M \\ M & R \end{bmatrix}$$

Desde que $M \neq 0$ existe uma correspondência um a um entre modelos de Markov e modelos ARMA. Entretanto, enquanto um modelo ARMA pode ser escrito em forma de espaço de estados ou de Markov em infinitas maneiras, somente algumas formas têm utilidade prática.

xii) Modelos de Equações Simultâneas

$$Y_t' \Gamma + X_t' B = \varepsilon_t$$

onde

Y_t - vetor $l \times 1$ de variáveis endôgenas,

X_t - vetor $m \times 1$ de variáveis exôgenas,

ε_t - vetor $l \times 1$ de ruído branco com matriz de covariância Σ .

Sendo N o tamanho da amostra escrevendo $Y = (Y_1, \dots, Y_N)'$ etc, temos

$$Y\Gamma + XB = E.$$

Estas equações são chamadas equações estruturais.

Se Γ tem posto máximo, as equações

$$Y = XT + W = XB\Gamma^{-1} + E^{-1}$$

são chamadas de forma reduzida do modelo.

O modelo dinâmico de equações simultâneas é

$$\Gamma'(L)Y_t + B'(L)X_t = \epsilon_t$$

onde $\Gamma'(L)$, $B'(L)$ são polinômios matriciais em L .

Observe que todos os modelos anteriores podem ser entendidos para o caso multivariado, bastando considerar as variáveis Y_t como vetores e os polinômios em L como polinômios matriciais.

CAPÍTULO II

MODELOS LINEARES UNIVARIADOS: BOX-JENKINS E EXTENSÕES

Introdução

Motivado pelo Teorema de Wold, estudaremos neste capítulo a classe de modelos lineares discretos. Vamos supor, em geral, que a parte determinística foi convenientemente extraída por mínimos quadrados ou por diferenças (ver Priestley, 1981, pg. 587 e pg. 771). Os nomes de Box e Jenkins estão intimamente ligados aos modelos a serem estudados, não só por terem os mesmos sido os pioneiros na sua divulgação e aplicação, mas também por terem desenvolvido métodos estatísticos necessários para sua implementação. Aqui apresentaremos um resumo desta metodologia com as diversas modificações e aperfeiçoamentos sugeridos desde o trabalho original de Box e Jenkins.

i) Modelo Linear Discreto e Propriedades

O modelo linear discreto introduzido no capítulo anterior é escrito

$$\begin{aligned} Y_t &= \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots = \\ &= (1 + \psi_1 L + \psi_2 L^2 + \dots) \varepsilon_t = \psi(L) \varepsilon_t \end{aligned} \quad (1)$$

onde ε_t é ruído branco com média 0, variância σ_ε^2 .

A forma (1) expressa Y_t em termos dos ruídos pre-

sente e passados. Uma forma alternativa expressa Y_t em termos do ruído presente e de todos os Y_t 's passados e é obtida fazendo sucessivamente para $t, t-1, \dots$

$$\varepsilon_t = Y_t - \psi_1 \varepsilon_{t-1} - \psi_2 \varepsilon_{t-2} - \dots$$

$$\varepsilon_{t-1} = Y_{t-1} - \psi_1 \varepsilon_{t-2} - \psi_2 \varepsilon_{t-3} - \dots \quad \text{etc}$$

Eliminando ε_{t-1} , obtêm-se

$$Y_t = \psi_1 Y_{t-1} + \varepsilon_t + (\psi_2 - \psi_1^2) \varepsilon_{t-2} + (\psi_3 - \psi_1 \psi_2) \varepsilon_{t-3} + \dots$$

Continuando as substituições de $\varepsilon_{t-2}, \varepsilon_{t-3}, \dots$ etc, tem-se

$$Y_t = \pi_1 Y_{t-1} + \pi_2 Y_{t-2} + \dots + \varepsilon_t$$

$$\text{i.e.,} \quad Y_t - \pi_1 Y_{t-1} - \pi_2 Y_{t-2} - \dots = \varepsilon_t$$

$$(1 - \pi_1 L - \pi_2 L^2 - \dots) Y_t = \varepsilon_t$$

$$\pi(L) Y_t = \varepsilon_t \quad (2)$$

onde os π_i 's são funções dos ψ_i . Observe que $\pi(L) = \psi^{-1}(L)$ e da expressão (1) podemos considerar que Y_t é a saída de um sistema dinâmico linear com entrada ruído branco e função de transferência $\psi(L)$ (A obtenção de $\pi(L) = \psi^{-1}(L)$ é válida sob certas condições, ver Priestley, pg. 145).

Vemos, portanto, que em ambas as expressões

$$Y_t = \psi(L) \varepsilon_t \quad \text{e} \quad \pi(L) Y_t = \varepsilon_t$$

temos um número infinito de parâmetros. Na prática (com um

número finito de observações) contornaremos este fato utilizando o resultado conhecido de que um polinômio infinito (em L) pode ser representado por uma função racional, isto é, razão entre dois polinômios de graus finitos, isto é,

$$Y_t = \psi(L)Y_t = \frac{\theta(L)}{\phi(L)} \epsilon_t \quad (3)$$

ou

$$\pi(L)Y_t = \frac{\phi(L)}{\theta(L)} Y_t = \epsilon_t \quad (4)$$

onde: $\psi(L) = 1 + \psi_1 L + \psi_2 L^2 + \dots$

$$\pi(L) = 1 - \pi_1 L - \pi_2 L^2 - \dots$$

$$\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q$$

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p.$$

Os momentos de primeira e segunda ordem do modelo (1) são (supondo $\sum_{i=1}^{\infty} \psi_i < \infty$, $\sum_{i=1}^{\infty} \psi_i^2 < \infty$ e $\sum_{i=1}^{\infty} \psi_i \psi_{i+k} < \infty$ para que a ordem de integração e soma infinita possam ser trocadas).

$$E\{Y_t\} = \mu = E\{\epsilon_t + \psi_1 \epsilon_{t-1} + \dots\} = E\{\epsilon_t\} + \psi_1 E\{\epsilon_{t-1}\} + \dots = 0 \quad (5a)$$

$$\begin{aligned} V\{Y_t\} &= \sigma_Y^2 = \gamma_0 = V\{\epsilon_t + \psi_1 \epsilon_{t-1} + \dots\} = V\{\epsilon_t\} + \psi_1^2 V\{\epsilon_{t-1}\} + \dots = \\ &= (1 + \psi_1^2 + \psi_2^2 + \dots) \sigma_\epsilon^2 = \sigma_\epsilon^2 \sum_{i=1}^{\infty} \psi_i^2, \quad \psi_0 = 1 \end{aligned} \quad (5b)$$

$$\begin{aligned} C\{Y_t, Y_{t-k}\} &= \gamma_k = E\{(\epsilon_t + \psi_1 \epsilon_{t-1} + \dots)(\epsilon_{t-k} + \psi_1 \epsilon_{t-k-1} + \dots)\} = \\ &= \sigma_\epsilon^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+k}. \end{aligned} \quad (5c)$$

Observe que a média, variância e covariância não dependem de t , logo o processo é estacionário. É importante res

saltar que este resultado foi obtido supondo convergência das séries $\sum_{i=0}^{\infty} \psi_i$, $\sum_{i=0}^{\infty} \psi_i \psi_{i+k}$ e $\sum_{i=0}^{\infty} \psi_i^2$. Lembrando que $|C(Y_t; Y_{t-k})| < [V\{Y_t\}V\{Y_{t-k}\}]^{1/2} = \sigma^2 \sum_{i=0}^{\infty} \psi_i^2$ temos "a condição para que o processo linear geral seja estacionário: $\sum_{i=0}^{\infty} \psi_i^2 < \infty$ ".

ii) Metodologia de Box-Jenkins e extensões

Da mesma forma que em teoria de probabilidades, onde cada modelo tem suas características, (por exemplo, a distribuição de Poisson tem média e variância igual; a distribuição normal tem assimetria zero e curtose igual a três) os modelos em (3) (ou 4) têm também características próprias.

Um pesquisador menos cuidadoso poderia sempre que tivesse um conjunto de dados, ajustar, por exemplo, uma distribuição de Poisson independente das características dos mesmos. Isto é mais ou menos o que ocorre em análise de séries temporais se sempre utilizarmos, por exemplo, amortecimento exponencial independente das características da série.

Em contraste, Box-Jenkins sugerem que se "deixe os dados falarem por si mesmo", reforçando a tese também deles de que o estatístico ou analista "deve se apaixonar por seus dados e não por seus modelos".

A metodologia consiste na primeira etapa, denominada "identificação", em obter estacionariedade da série e reconhecer padrões que caracterizam algum modelo da classe geral (3)

4) através da inspeção da função de autocorrelação. - FAC estral e de transformações da mesma. (função de autocorrelação parcial - FACP; FAC inversa - FACI, FACP inversa - FACPI, parâmetros R, S e método corner, FACP generalizada, FAC estendida). Originalmente Box-Jenkins sugerem apenas inspecionar a função de autocorrelação e FACP.

Após a etapa de identificação preliminar do modelo, passa-se à etapa de "estimação" eficiente dos parâmetros e em seguida efetua-se a "verificação" que consiste em testar a adequabilidade dos modelos. Se não é considerado adequado, o modelo deve ser modificado, retornando-se ao ciclo: identificação, estimação e verificação. Após um modelo adequado ser selecionado passamos a sua utilização em previsão e/ou controle.

A parte mais difícil desta metodologia é que mais importante de sua aplicação é a etapa de identificação, pois requer experiência do analista. Na tentativa de minimizar e mesmo evitar esta dificuldade, diversos autores (Akaike, Parzen, Hannan, Schwartz etc) tem proposto métodos automáticos para escolha dos modelos.

Também na parte de estimação, novos algoritmos e métodos (Ansley, Harvey e Phillips, McLeod etc) têm sido desenvolvidos recentemente, bem como na parte de verificação novas aproximações e testes têm sido propostos.

Retornaremos a esses pontos adiante, mas inicialmente como em cursos de probabilidade que apresentam as distribuições

buições com suas características e que antecedem cursos de inferência estatística onde aprendemos a usá-las, nas próximas seções apresentamos um resumo dos modelos de séries temporais na classe $Y = \psi(L)\varepsilon_t$, bem como suas propriedades.

Modelos Autorregressivos, Médias Móveis e Mistos

3.1 Modelos Autorregressivos - AR(p)

Estes são modelos da forma (4) com $\theta(L) = 1$ e $\phi(L)$ de grau p em L , isto é, $\pi_i = 0, i > p$ e escreve-se

$$\phi(L)Y_t = \epsilon_t \quad \text{ou} \quad Y_t = \phi^{-1}(L)\epsilon_t$$

i.e.,

$$Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = \epsilon_t \quad (6a)$$

$$Y_t = \frac{1}{1 - \phi_1 L - \dots - \phi_p L^p} \epsilon_t \quad (6b)$$

As propriedades deste modelo geral serão indicadas a partir da análise dos casos particulares AR(1) e AR(2).

i) Modelo AR(1)

Neste caso temos

$$Y_t - \phi Y_{t-1} = \epsilon_t$$

ou

$$Y_t = \frac{1}{1 - \phi L} \epsilon_t = (1 + \phi L + \phi L^2 + \dots) \epsilon_t = \epsilon_t + \phi \epsilon_{t-1} + \phi^2 \epsilon_{t-2} + \dots,$$

isto é, o modelo AR(1) é equivalente a um modelo MA(∞) com a restrição $\psi_i = \phi^i$ nos parâmetros.

De (5), os momentos são

$$\sigma_Y^2 = \gamma_0 = (1 + \phi^2 + \phi^4 + \dots) V\{\epsilon_t\} = \frac{1}{1 - \phi^2} \sigma_\epsilon^2$$

$$\begin{aligned} \gamma_1 &= E\{Y_t Y_{t-1}\} = E\{\phi Y_{t-1}^2 + \varepsilon_t Y_{t-1}\} = \phi \gamma_0 = \frac{\phi}{1 - \phi^2} \sigma_\varepsilon^2 \\ &\vdots \\ \gamma_k &= \dots = \frac{\phi^k}{1 - \phi^2} \sigma_\varepsilon^2 \end{aligned}$$

A função de autocorrelação é

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi^k \quad k \geq 0.$$

Finalmente a condição de estacionariedade é

$$\sum_{i=0}^{\infty} \psi_i = \sum_{i=0}^{\infty} \phi^i < \infty \Rightarrow |\phi| < 1.$$

Portanto estacionariedade é equivalente a $-1 < \phi < 1$.

Em termos do polinômio $1 - \phi L$ isto é equivalente a impor que a equação

$$1 - \phi L = 0$$

tenha raiz $L = \frac{1}{\phi}$ maior que 1 em valor absoluto.

ii) Modelo AR(2)

Neste caso

$$Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} = \varepsilon_t$$

ou

$$Y_t = \frac{1}{1 - \phi_1 L - \phi_2 L^2} \varepsilon_t$$

e da mesma forma o modelo AR(2) é equivalente a um modelo MA(∞) com restrições $\psi_i = f_i(\phi_1, \phi_2)$ nos parâmetros.

Obtêm-se, após alguma álgebra,

$$\sigma_Y^2 = \frac{(1 - \phi_2)}{(1 + \phi_2)(1 - \phi_1 - \phi_2)(1 + \phi_1 - \phi_2)} \sigma_E^2$$

$$Y_1 = \phi_1 Y_0 + \phi_2 Y_{-1}$$

$$Y_2 = \phi_1 Y_1 + \phi_2 Y_0$$

⋮

⋮

$$Y_k = \phi_1 Y_{k-1} + \phi_2 Y_{k-2}$$

A função de autocorrelação portanto satisfaz

$$\beta_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}, \quad k \geq 1,$$

conhecida como equações de Yule-Walker. Observe que esta equação pode ser escrita

$$\phi(L)\rho_k = (1 - \phi_1 L - \phi_2 L^2)\rho_k = 0$$

que é uma equação de diferenças finitas cuja solução depende das raízes de

$$1 - \phi_1 L - \phi_2 L^2 = 0$$

e será portanto soma de exponenciais ou senoides amortecidas.

A condição de estacionariedade é dada por

$$\phi_2 > -1, \quad \phi_1 + \phi_2 < 1 \quad \text{e} \quad -\phi_1 + \phi_2 < 1.$$

Em termos do polinômio $1 - \phi_1 L - \phi_2 L^2$, esta condição equivale a impor que a equação $1 - \phi_1 L - \phi_2 L^2 = 0$ tenha raízes fora do círculo unitário.

iii) Modelo AR(p)

Para o caso geral do modelo (6) tem-se: que o modelo AR(p) é equivalente a um modelo MA(∞) com restrição nos parâmetros.

A função de autocorrelação satisfaz as equações de Yule-Walker

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \quad (7)$$

ou

$$(1 - \phi_1 L - \dots - \phi_p L^p) \rho_k = \phi(L) \rho_k$$

e sua solução depende das raízes de $\phi(L) = 0$. Portanto a função de autocorrelação será a soma de exponenciais e/ou senoides amortecidas.

A condição de estacionariedade é que as raízes de $\phi(L) = 0$ estejam fora do círculo unitário.

2.2 Estacionariedade e Estabilidade

A condição de estacionariedade do modelo AR(p)

$$Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = \epsilon_t$$

é que as raízes de

$$\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p = 0 \quad (7a)$$

estejam fora do círculo unitário.

Se pensarmos no modelo AR(p) como uma equação de diferenças não homogêneas com coeficientes constantes, a equação característica associada é.

$$x^p - \phi_1 x^{p-1} - \dots - \phi_p = 0 \quad (7b)$$

e a equação é estável se as raízes da equação característica estão dentro do círculo unitário.

As duas condições são equivalentes bastando notar que as equações são iguais se substituirmos em (7b) x por $1/L$ e multiplicarmos por L^p .

2.3 Modelos Médias Móveis - MA(q)

Estes são modelos na forma (3) com $\phi(L) = 1$ e $\theta(L)$ de grau q em L , isto é, $\psi_i = 0$ para todo $i > q$. Escreve-se

$$Y_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} = \theta_q(L) \epsilon_t \quad (8)$$

As propriedades deste modelo geral serão indicadas a partir da análise dos casos particulares MA(1) e MA(2).

i) Modelo MA(1)

Neste caso temos

$$Y_t = \epsilon_t + \theta_1 \epsilon_{t-1} = (1 + \theta L) \epsilon_t$$

ou

$$\frac{1}{1 + \theta L} Y_t = (1 - \theta L + \theta^2 L^2 + \dots) Y_t = \epsilon_t$$

isto é, o modelo MA(1) é equivalente a um modelo AR(∞) com a restrição $\pi_i = (-1)^i \theta_1^i$ nos parâmetros. De (5), os momentos são

$$\sigma_y^2 = \gamma_0 = (1 + \theta_1^2) \sigma_\varepsilon^2$$

$$\gamma_1 = \theta_1 \sigma_\varepsilon^2$$

$$\gamma_k = 0, \quad k > 1$$

A função de autocorrelação é

$$\rho_0 = 1, \quad \rho_1 = \frac{\theta_1}{1 + \theta_1^2}, \quad \rho_k = 0, \quad k > 1$$

O modelo MA(1) é sempre estacionário pois

$$\sum_{i=1}^{\infty} \psi_i = 1 + \theta_1^2 < \infty.$$

ii) Modelo MA(2)

Neste caso temos

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} = (1 + \theta_1 L + \theta_2 L^2) \varepsilon_t$$

ou

$$\frac{1}{1 + \theta_1 L + \theta_2 L^2} Y_t = \varepsilon_t$$

e da mesma forma o modelo MA(2) é equivalente a um modelo AR(∞) com restrições $\pi_i = g_i(\theta_1, \theta_2)$ nos parâmetros.

De (5) obtêm-se a função de autocorrelação

$$\rho_0 = 1$$

$$\rho_1 = \frac{\theta_1(1 + \theta_2)}{1 + \theta_1^2 + \theta_2^2}$$

$$\rho_2 = \frac{\theta_2}{1 + \theta_1^2 + \theta_2^2}$$

$$\rho_k = 0, \quad k > 2$$

e da mesma forma o modelo é sempre estacionário.

1) Modelo MA(q)

Para o caso geral do modelo (8) tem-se que o modelo MA(q) é equivalente a um AR(∞) com restrição nos parâmetros.

A função de autocorrelação e a variância são

$$\sigma_Y^2 = \gamma_0 = (1 + \sigma_1^2 + \dots + \sigma_q^2) \sigma_\varepsilon^2$$

$$\rho_k = \frac{\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2} \quad k = 1, \dots, q \quad (9)$$

$$\rho_k = 0, \quad k > q.$$

e o modelo é sempre estacionário.

4. Inversibilidade e Identificabilidade

O modelo MA(1),

$$Y_t = \varepsilon_t + \theta \varepsilon_{t-1}$$

de ser escrito

$$Y_t = \theta Y_{t-1} - \theta^2 Y_{t-2} + \dots + (-\theta)^j Y_{t-j} - (-\theta)^{j+1} \varepsilon_{t-j+1} + \varepsilon_t$$

Para que o modelo não dependa muito do que ocorreu no passado distante ($Y_{t-j}, \varepsilon_{t-j-1}$, j grande) é necessário que $|\theta| < 1$. Se $|\theta| < 1$ podemos escrever o modelo na forma $AR(\infty)$

$$Y_t = \theta Y_{t-1} - \theta^2 Y_{t-2} + \dots + \varepsilon_t$$

ou

$$\pi(L)Y_t = \varepsilon_t.$$

Um modelo $MA(1)$ com $|\theta| > 1$ é estacionário mas não pode ser escrito na forma $AR(\infty)$. Entretanto sua função de autocorrelação

$$\rho_1 = \frac{\theta}{1 + \theta^2}$$

pode ser reproduzida pelo processo inversível com parâmetro $|\theta'| < 1$ onde $\theta' = \frac{1}{\theta}$ pois

$$\rho_1 = \frac{\theta'}{1 + \theta'^2} = \frac{1/\theta}{1 + 1/\theta^2} = \frac{\theta}{\theta^2 + 1}$$

Portanto para uma mesma função de autocorrelação temos dois modelos. Um inversível e outro não inversível.

No caso geral do modelo $MA(q)$, temos 2^q modelos com mesma função de autocorrelação, mas somente um é inversível. Neste caso a condição para inversibilidade é que

$$\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q = 0$$

tenha raízes fora do círculo unitário.

No caso particular do MA(1) $\theta Y_t + \epsilon_t = Y_t$

$$1 + \theta L = 0, \quad |L| = \left| \frac{-1}{\theta} \right| > 1 \Rightarrow |\theta| < 1$$

Portanto, restringindo atenção a modelos invertíveis não temos problemas de identificabilidade e além disso temos previsões mais eficientes (ver Harvey, 1981, pg. 161).

2.5 Modelos Autorregressivos - Médias Móveis - ARMA(p,q)

É um modelo (3) ou (4) com $\phi_p(L) Y_t = \theta_q(L) \epsilon_t$

Isto é,

$$\phi_p(L) Y_t = \theta_q(L) \epsilon_t \quad (10a)$$

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} \quad (10b)$$

Vejamos inicialmente o caso ARMA(1,1)

i) Modelo ARMA (1,1)

Neste caso temos

$$Y_t = \phi_1 Y_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1}$$

e o modelo é equivalente respectivamente $(\theta(L))$ invertível

$$\text{AR}(\infty): \frac{(1 - \phi L)}{(1 - \theta L)} Y_t = (1 - \theta L + \theta^2 L^2 - \dots)(1 - \phi L) Y_t = \epsilon_t$$

$$\text{MA}(\infty): Y_t = \frac{(1 + \phi L)}{(1 - \phi L)} \epsilon_t = (1 + \phi L + \phi^2 L^2 + \dots)(1 + \theta L) \epsilon_t$$

Os momentos são

$$Y_0 = \phi_1 Y_1 + Y_0(Y, \epsilon) + \theta_1 Y_1(Y, \epsilon)$$

$$Y_1 = \phi_1 Y_0 + \theta_1 Y_0(Y, \epsilon)$$

$$Y_2 = \phi_1 Y_1$$

$$\vdots$$

$$Y_k = \phi_1 Y_{k-1} = \phi_1^{k-1} Y_1$$

onde

$$Y_0(Y, \epsilon) = E\{Y_t \epsilon_t\} = \sigma_\epsilon^2$$

$$Y_1(Y, \epsilon) = E\{Y_t \epsilon_{t-1}\} = (\phi + \theta) \sigma_\epsilon^2$$

substituindo obtêm-se

$$Y_0 = \frac{1 + \theta^2 + 2\theta\phi}{1 - \phi^2} \sigma_\epsilon^2$$

$$Y_1 = \frac{(1 + \phi\theta)(\phi + \theta)}{1 - \phi^2} \sigma_\epsilon^2$$

$$Y_k = \phi_1^{k-1} Y_1$$

A função de autocorrelação do modelo ARMA(1,1) se parece com a do AR(1). A diferença é que enquanto para o AR(1) ela é uma exponencial amortecida desde lag 0, para um ARMA(1,1) ela é exponencial amortecida após lag 0, i.e., a partir de lag 1.

Finalmente, o modelo é estacionário se $|\phi| < 1$ e invertível se $|\theta| < 1$.

ii) Modelo ARMA(p,q)

Para o caso geral do modelo (10) tem-se que o modelo ARMA(p,q) é equivalente a um AR(∞) com restrição nos parâmetros e a um MA(∞) com restrição nos parâmetros (se $\theta(L)$ inversível).

Os momentos satisfazem a equação

$$\gamma_k = \phi_1 \gamma_{k-1} + \dots + \phi_p \gamma_{k-p} + \gamma_k(\gamma, \varepsilon) + \gamma_{k+1}(\gamma, \varepsilon) + \dots + \gamma_{k+q}(\gamma, \varepsilon) \quad (11)$$

onde: $\gamma_\ell(\gamma, \varepsilon) = E\{Y_t, \varepsilon_{t-\ell}\}$

$$\gamma_\ell(\gamma, \varepsilon) = 0 \quad \ell < 0$$

$$\gamma_\ell(\gamma, \varepsilon) \neq 0 \quad \ell \geq 0$$

Para $k > q$ a equação simplifica

$$\gamma_k = \phi_1 \gamma_{k-1} + \dots + \phi_p \gamma_{k-p} \quad k > q \quad (12)$$

Dividindo por γ_0 obtêm-se

$$\phi(L) \rho_k = 0 \quad k > q$$

Logo os q primeiros valores ρ_1, \dots, ρ_q dependem dos parâmetros MA (θ 's) e dos parâmetros AR (ϕ 's) através de (11). Os p valores $\rho_{q-(p-1)}, \rho_{q-1}, \rho_q$ fornecem os valores iniciais de $\phi(L) \rho_k = 0, k > q$. Se $q - p < 0$ toda a função de autocorrelação $\rho_j, j = 0, 1, \dots$, é uma mistura de exponenciais e/ou senoides amortecidas, cuja forma é ditada por $\phi(L)$. Se $q - p \geq 0$ haverá $q - p + 1$ valores $\rho_0, \rho_1, \dots, \rho_q$ que não seguem este padrão, e a partir deles o padrão é seguido.

Finalmente, a condição de estacionariedade é que as raízes de $\phi(L) = 0$ estejam fora do círculo unitário e a condição de inversibilidade é que as raízes de $\theta(L) = 0$ estejam fora do círculo unitário.

3. Modelos com Média não Nula, Não Estacionários e Sazonais

3.1 Modelos com Média não Nula

Até o momento supomos que $E\{Y_t\} = \mu = 0$. Se entre tanto tem-se uma série estacionária com $\mu \neq 0$, subtrai-se a média da série. Neste caso o modelo é escrito

$$\phi_p(L)(Y_t - \mu) = \theta_q(L)\varepsilon_t$$

ou ainda

$$\phi(L)Y_t = \theta_0 + \theta_q(L)\varepsilon_t \quad (13)$$

onde $\theta_0 = \phi(L)\mu = (1 - \phi_1 - \dots - \phi_p)\mu$.

3.2 Modelos Não Estacionários - ARMA(p,d,q)

i) Transformações

Quando a série em estudo não é estacionária, o que ocorre na maioria das vezes, procura-se através de transformações tornar a mesma estacionária.

Uma transformação muito utilizada em séries de demanda e em hidrologia é a transformação logaritmica. Razão para utilização desta transformação é que mudança no logaritmo se aproxima a mudanças percentuais†

$$\log y_t - \log y_{t-1} = \log(y_t/y_{t-1}) \approx (1 + y_t/y_{t-1}) = 1 + r.$$

Além disto, esta transformação é útil quando a variância é proporcional a média. Entretanto, não há razão

para sempre usar-se esta transformação bem como qualquer outra como $\sqrt{Y_t}$ ou $1/Y_t$.

Um enfoque mais consistente e flexível, que não só permite fazer a transformação, bem como descobrir qual a transformação mais adequada consiste em utilizar a transformação de Box-Cox definida por

$$Y_t^{(\lambda)} = \begin{cases} \frac{(Y_t + \alpha)^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \log(Y_t + \alpha) & \lambda = 0 \end{cases} \quad (14)$$

Na sua forma mais simples o procedimento consiste em escolher diversos valores de λ . (Em geral $\lambda = 0.9, -0.8, \dots, 0.8, 0.9$ e α tal que $Y_t + \alpha > 0$) e analisar as diversas séries transformadas escolhendo para λ aquele valor que otimiza certo critério. Após estimar a transformação conveniente obtêm-se $Y_t^{(\lambda)}$ estacionária, que é então modelado por um ARMA.

Certos cuidados devem ser tomados ao utilizar esta transformação tendo em vista que a maioria dos programas de computador disponíveis não faz a previsão da série original corretamente. Para ilustração consideremos o caso mais simples da transformação logarítmica em que a série analisada foi $Z_t = \log Y_t$ e obteve-se a previsão \hat{Z}_t . A previsão correta para Y_t é $\exp\{\hat{Z}_t + \frac{1}{2} \sigma_Z^2\}$ e não somente a transformação inversa $\exp\{\hat{Z}_t\}$. A situação se complica muito no caso de outras transformações; Nelson e Granger

(1979) mostraram empiricamente que, ao menos para sêries econômicas, o trabalho computacional necessário para obter a previsão correta não é muito compensador em termos de ganho em precisão na previsão.

Tendência Linear e Polinomial Determinística

Suponha-se uma sêrie com variação linear no nível (média) representada pelo modelo

$$Y_t = \alpha + \beta t + u_t$$

com variância e covariância constantes. Obtêm-se uma sêrie estacionária tomando-se a diferença

$$\nabla Y_t = Y_t - Y_{t-1} = \beta + u_t - u_{t-1} = \beta + u_t$$

É fácil verificar que ∇Y_t é estacionária e, portanto, pode ser modelada por

$$\phi(L)(\nabla Y_t - \beta) = \theta(L)\epsilon_t$$

ou

$$\phi(L)\nabla Y_t = \theta_0 + \theta(L)\epsilon_t.$$

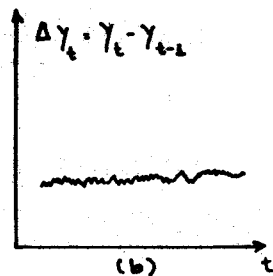
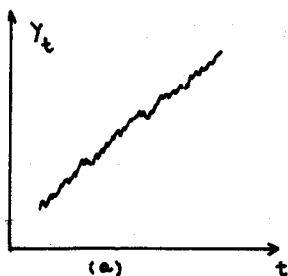


Fig. III - Gráficos: (a) Sêrie original $Y_t = \alpha + \beta t + u_t$ e (b) Sêrie de diferenças $\nabla Y_t = \beta + u_t$.

Se a tendência seguir um polinômio de grau d , toma-se a diferença $\nabla^d = (1 - L)^d$ de grau d para obter-se estacionariedade e o modelo é

$$\phi(L)\nabla^d y_t = \theta_0 + \theta(L)\varepsilon_t \quad (15)$$

iii) Não Estacionariedade Homogênea

Muitas séries temporais se comportam como se não tivessem uma média constante; entretanto elas apresentam homogeneidade no sentido de que, com exceção do nível local, ou nível local e tendência, qualquer parte da série tem comportamento semelhante.



Fig. II2 - Não estacionariedade homogênea (a) - em nível, (b) - em nível e em tendência.

Analisemos inicialmente o caso de não estacionariedade homogênea em nível. Se temos apenas dois níveis,

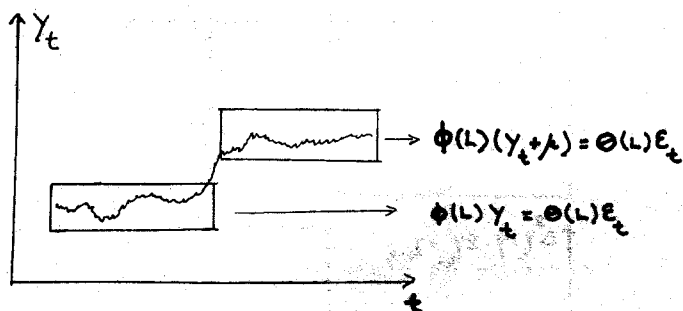


Fig. II3 - Não estacionariedade homogênea em nível.

temos os modelos

$$\phi(L)(Y_t + \mu) = \theta(L)\epsilon_t \quad \text{e} \quad \phi(L)Y_t = \theta(L)\epsilon_t$$

para cada segmento respectivamente. Logo, tem-se

$$\phi(L)Y_t = \phi(L)(Y_t + \mu),$$

i.e., $\phi(L)\mu = 0.$

Para que isto ocorra é necessário que

$$\phi(L) = \phi'(L)(1 - L) = \phi'(L)\nabla$$

pois $(1 - L)\mu = 0.$

Portanto o modelo para toda a série é

$$\phi'(L)\nabla Y_t = \theta(L)\epsilon_t. \quad (16)$$

A não estacionariedade homogênea em tendência é obtida relacionando-a aos resultados de tendência linear e não estacionariedade em nível e pode ser vista facilmente na Figura II4.

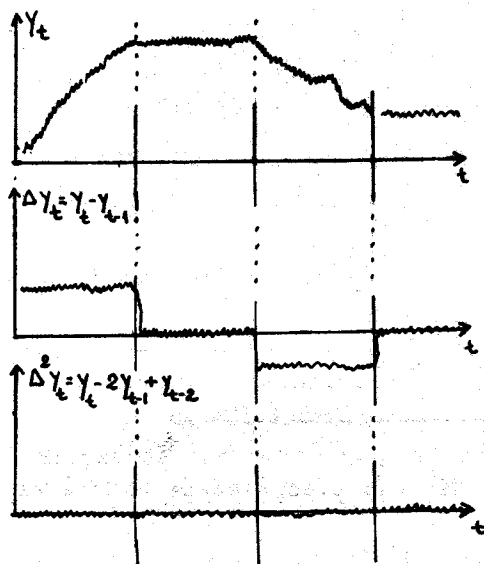


Fig. II4 - Não estacionariedade homogênea em nível e tendência.

Portanto o modelo é

$$\phi(L)\nabla^2 Y_t = \theta(L)\varepsilon_t \quad (17)$$

iv) Comentários:

- O modelo geral, combinando a transformação de Box-Cox

$$\phi(L)\nabla^d Y_t^{(\lambda)} = \theta_0 + \theta(L)\varepsilon_t, \quad (18)$$

é denominado autorregressivo integrado médias móveis. A razão do nome integrado é que ele é obtido de uma soma do modelo estacionário. Por exemplo, se $Z_t = \nabla Y_t$ é estacionário

$$Z_t = (1 - L)Y_t$$

e portanto

$$\begin{aligned} Y_t &= \frac{1}{1-L} Z_t = (1 + L + L^2 + \dots)Z_t \\ &= Z_t + Z_{t-1} + Z_{t-2} + \dots \end{aligned}$$

- De (ii) acima, vemos que o modelo (18) representa uma série com tendência determinística polinomial de grau d se incluirmos $\theta_0 \neq 0$ e, de (iii), vemos que (18) representa uma série com tendência estocástica se omitirmos o termo constante, isto é $\theta = 0$.

- Em muitas aplicações, principalmente em economia, é importante obter-se estimação de fatores como tendência e sazonalidade. Esses fatores são de difícil interpretação no caso de se utilizar diferenças. Recentemente tem-se voltado a modelar séries explicitando esses fatores. (Por exemplo, Harvey, 1983, Vishwakarma, 1983).

- A necessidade de transformação e diferenciação é verificada por inspeção visual do gráfico da série que indicará indícios de não estacionariedade. Por exemplo, tendência e/ou variância mudando com o nível etc.

Além disso se a função de autocorrelação amostral ρ_{λ} cuspitar a se amortecer a zero, i.e. se decrescer linearmente, temos indícios de que é necessário tomar diferenças. Anderson (1975, pg. 116) sugere que d pode ser determinado como o valor para o qual as séries $Y_t, \nabla Y_t, \nabla^2 Y_t$ etc tenha menor variância. Ver também Yajima (1982).

Embora a série de diferenças de uma série estacionária seja ainda estacionária, diferenciar demasiadamente uma série complica a estrutura da autocorrelação e aumenta a variância da série diferenciada.

3.3 Modelos Sazonais - SARIMA(p,d,q) x (P,D,Q)

Estes modelos correspondem a generalização do modelo ARIMA para séries com forte componente sazonal.

Tradicionalmente métodos para modelar séries sazonais por exemplo, observadas mensalmente, utilizam modelos de regressão do tipo

$$Y_t = \beta_0 + \sum_{j=1}^6 \left(\beta_{1j} \operatorname{sen} \frac{2\pi jt}{12} + \beta_{2j} \operatorname{sen} \frac{2\pi jt}{12} \right) + \varepsilon_t \quad (19a)$$

ou

$$Y_t = \sum_{j=1}^{12} \beta_j x_j + \varepsilon_t, \quad x_j = \begin{cases} 1 & \text{mês } j \\ 0 & \text{outros meses.} \end{cases} \quad (19b)$$

Estes modelos supõem os coeficientes β 's constantes e ε_t não correlacionados e são capazes de reproduzir comportamento sazonal com amplitude e fase constantes. Portanto são apropriados para descrever sazonalidade em séries estáveis como temperatura e outras variáveis meteorológicas. Entretanto em Economia e na maioria de outras áreas é mais realístico utilizar modelos que permitam amplitudes e fases adaptativas isto é, β 's variando. Além disso, é mais realístico supor erros correlacionados.

Se as séries têm sazonalidade s , os seguintes modelos são generalizações imediatas dos modelos ARIMA e levado em consideração a dependência de observações afastadas por s períodos.

Tem-se

$$\begin{aligned}
 Y_t &= \eta_t + \theta_1 \eta_{t-s} + \dots + \theta_Q \eta_{t-Qs} \\
 &= (1 + \theta_1 L^s + \dots + \theta_Q L^{Qs}) \eta_t = \theta(L) \eta_t
 \end{aligned} \quad (20)$$

A autocorrelação seria diferente de zero somente nos lag $s, 2s \dots Qs$; se η_t é ruído branco. Isto significaria que não há dependência em períodos adjacentes $t-1, t, t+1$, o que não é razoável na maioria dos casos.

ii) Modelo AR(P)_s

Tem-se

$$Y_t = \phi_1 Y_{t-s} + \dots + \phi_P Y_{t-Ps} + \eta_t$$

ou

$$(1 - \phi_1 L^s - \dots - \phi_P L^{Ps}) Y_t = \eta_t$$

$$\phi(L^s) Y_t = \eta_t \quad (21)$$

Por exemplo, para $P = 1$

$$Y_t = \phi_1 Y_{t-s} + \eta_t$$

e a autocorrelação decresce exponencialmente a zero nos lags $s, 2s, 3s, \dots$ se o η_t é ruído branco e é zero nos outros casos intermediários.

iii) Modelo ARMA(P,Q)_s

Tem-se, por exemplo, para $P = 2, Q = 2$

$$Y_t = \phi_1 Y_{t-s} + \phi_2 Y_{t-2s} + \eta_t + \theta_1 \eta_{t-s} + \theta_2 \eta_{t-2s} \quad (22)$$

No caso geral

$$\phi(L^s)Y_t = \theta(L^s)\eta_t.$$

Novamente se η_t é ruído branco as autocorrelações somente são não nulas nos lags múltiplos de s .

iv) Modelo Multiplicativo SARIMA(p,d,q) x (P,D,Q)_s

Vimos que se os η_t são ruído branco o modelo não é razoável pois supõe que períodos adjacentes são não correlacionados. O mais razoável é supor que $\{\eta_t\}$ seja correlacionado e seja modelado em geral por um ARIMA. Portanto, o modelo geral fica

$$\begin{cases} \phi(L^s)\nabla_s^D Y_t = \theta(L^s)\eta_t \\ \phi(L)\nabla^d \eta_t = \theta(L)\varepsilon_t \end{cases} \quad (23a)$$

ou

$$\phi(L^{12})\phi(L)\nabla_s^D \nabla^d Y_t = \theta(L^s)\theta(L)\varepsilon_t \quad (23b)$$

onde

$$\nabla_s^D = (1 - L^s)^D$$

$$\nabla^d = (1 - L)^d$$

$$L^s Y_t = Y_{t-s}$$

$$\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$$

$$\theta(L) = 1 + \theta_1 L - \dots + \theta_q L^q$$

$$\phi(L^s) = 1 - \phi_1 L^s - \dots - \phi_p L^{ps}$$

$$\theta(L^s) = 1 + \theta_1 L^s - \dots + \theta_q L^{qs}$$

A função de autocorrelação do modelo obedece às mesmas regras do modelo ARIMA para os lags até s e nos lags múltiplos de s . Nos lags adjacentes aos lags múltiplos de s temos um outro comportamento, para cada modelo, e este padrão de comportamento é apresentado em Box-Jenkins (1976, pg. 329-333) para diversos modelos. Em geral entretanto esta parte não é muito importante na fase de identificação.

v) Exemplo

Apresenta-se agora um modelo que tem sido frequentemente utilizado para dados mensais (ver Kenny e Durbin; 1982). A razão de sua grande adequabilidade à maioria dos dados será indicada. Trata-se do modelo SARIMA $(0,1,1) \times (0,1,1)_{12}$

$$\begin{aligned} Z_t &= \nabla_{12} \nabla Y_t = (1 - \theta L^{12})(1 - \theta L)\varepsilon_t \\ &= (1 - \theta L - \theta L^{12} + \theta \theta L^{13})\varepsilon_t \end{aligned} \quad (24)$$

A função de autocovariância é

$$\gamma_0 = (1 + \theta^2)(1 + \theta^2)\sigma_\varepsilon^2$$

$$\gamma_1 = -\theta(1 + \theta^2)\sigma_\varepsilon^2$$

$$\gamma_{11} = \gamma_{13} = \theta\theta\sigma_\varepsilon^2$$

$$\gamma_{12} = -\theta(1 + \theta^2)\sigma_\varepsilon^2$$

$$\gamma_k = 0 \quad k = 2, \dots, 10, 14, \dots$$

Box-Jenkins (1976, pg. 329-333) fornece as autocovariâncias para diversos modelos multiplicativos e alguns

modelos mais gerais não multiplicativos. Por exemplo, o modelo

$$Z_t = (1 - \theta L - \theta_1 L^{12} - \theta_2 L^{13}) \varepsilon_t$$

é não multiplicativo e a diferença em relação ao exemplo anterior é que não temos a restrição $\gamma_{11} = \gamma_{13}$.

Finalmente, vejamos a razão da grande adequabilidade do modelo (24). Suponha que temos uma série mensal sazonal com tendência linear. Um modelo em geral adequado é gerado por

$$Y_t = a + bt + f(t) + \varepsilon_t \quad (25)$$

onde $f(t)$ é uma função periódica de período 12, por exemplo as funções em (19).

Tomando ∇Y_t para eliminar a tendência,

$$\nabla Y_t = b + f(t) - f(t-1) + \varepsilon_t - \varepsilon_{t-1}.$$

Tomando $\nabla^{12} Y_t$ para eliminar sazonalidade,

$$\begin{aligned} \nabla^{12} \nabla Y_t &= (\varepsilon_t - \varepsilon_{t-1}) - (\varepsilon_{t-12} - \varepsilon_{t-13}) = \\ &= (1 - L^{12})(1 - L)\varepsilon_t. \end{aligned}$$

Logo (25) é o caso particular de (24) com $\theta = 1$ e $\theta = 1$.

iv) Modelo com diferença sazonal dividida

Muitas séries apresentam uma sazonalidade que não é múltipla inteira da unidade (hora, mês, ano) de observação.

Um exemplo clássico é a série de manchas solares (Izenman, 1983) que tem uma periodicidade média de 11,11 meses.

Nestes casos, Conradsen e Spliid (1981) sugerem a utilização da diferença

$$\nabla_{\alpha} Y_t = Y_t - (1 - \alpha)Y_{t-s} - \alpha Y_{t-s-1} \quad (27)$$

Por exemplo, analisando uma série de marés horárias eles observaram uma sazonalidade de 12,42 horas e a diferença

$$\nabla_{12,42} Y_t = Y_t - 0.58Y_{t-12} - 0.42Y_{t-13}$$

era mais adequada do que $\nabla_{12} Y_t = Y_t - Y_{t-12}$ ou $\nabla_{13} Y_t = Y_t - Y_{t-13}$.

4. Agregação

4.1 Introdução

Existem dois tipos de agregação. O primeiro ocorre quando a série observada é a soma de duas ou mais séries independentes. Essas séries individuais são chamadas de componentes não observadas, e se cada uma segue um modelo ARMA a questão de interesse é saber o modelo para a série agregada.

O segundo tipo de agregação é agregação no tempo. Os dados disponíveis se referem a uma unidade de tempo maior que a unidade de tempo empregada no modelo. Por exemplo, um intervalo natural na especificação do modelo pode ser mês, porém só temos dados disponíveis numa base anual. Se a variável em estudo é fluxo, como o Produto Nacional Bruto, então ela é considerada como tendo sido sujeita a agregação temporal. Quando a variável é estoque, como moeda disponível, o problema é diferente. Uma variável estoque é essencialmente contínua e se as observações foram feitas em intervalos de tempo menos frequentes do que o modelo requer o problema é de observações perdidas. A questão de interesse neste segundo caso também é saber que modelo segue a série agregada quando cada série desagregada segue um modelo ARMA.

A seguir apresenta-se um resumo dos resultados relevantes obtidos por Granger e Morris (1976). Outras referências recentes e extensões são apresentadas em Rose (1977), Tiao e Guttman (1980)

Agregação de Diferentes Séries

Considere-se duas séries X_t e Z_t estacionárias, independentes e com média zero. Se X é $ARMA(p_1, q_1)$ e Z_t é $ARMA(p_2, q_2)$ então

$$Y_t = X_t + Z_t$$

onde

$$p_3 \leq p_1 + p_2, \quad q_3 = \max(p_1 + q_2, p_2 + q_1). \quad (28)$$

Alguns casos específicos de interesse são

+ ruído branco = $ARMA(p, p)$

+ $AR(p_2)$ = $ARMA[p_1 + p_2, \max(p_1, p_2)]$

+ $MA(q_2)$ = $ARMA[0, \max(q_1, q_2)]$

+ ruído branco = $ARMA[p, \max(p, q)]$

+ $MA(q)$ = $ARMA(p, p + q)$.

Agregação Temporal e Observações Perdidas

Considere-se inicialmente que Y_t é uma variável fluída que segue um $ARMA(p, q)$ e que é observada agregada a cada m períodos. A variável observada é

$$Y_t^* = \sum_{j=0}^{m-1} Y_{t-j} \quad t = m, 2m, 3m, \dots$$

Mostra-se que Y_t^* segue um $ARMA(p, q^*)$ onde

$$q^* = [(p+1)(m-1) + q]/m \quad (29)$$

arredondado para o inteiro mais próximo.

Quando m cresce, a série agregada tende para ARMA($p, p+1$) se $q \geq p+1$ e ARMA(p, q) se $q < p+1$. Entretanto, as autocorrelações entre observações sucessivas tendem também a se tornar pequenas e a série agregada tende a se comportar como ruído branco.

Se Y_t é uma variável estoque que segue um ARMA(p, q) para uma determinada unidade de tempo, porque é observada cada período m de tempo a observação de Y_t segue um ARMA(p, q) onde

$$q^* = [p(m-1) + q]/m \quad (30)$$

arredondada para o inteiro mais próximo. Se o intervalo m de observações aumenta o modelo das observações é ARMA(p, p) se $q \geq p$ e ARMA($p, p-1$) se $q < p$.

5. Identificação

5.1 Introdução

No início do capítulo mencionamos que a parte mais difícil para utilização da metodologia é a fase de identificação. Até o momento a ferramenta disponível para identificar o modelo de uma série temporal é a função de autocorrelação. Entretanto, o uso de somente essa função torna difícil senão impossível a identificação em casos reais. Ao estimar a função de autocorrelação de uma série temporal a mesma é sujeita a grande variabilidade amostral e isto torna difícil sua identificação. Para ilustrar, uma possível função de autocorrelação amostral pode-se apresentar graficamente como na Fig. II-5

(a)

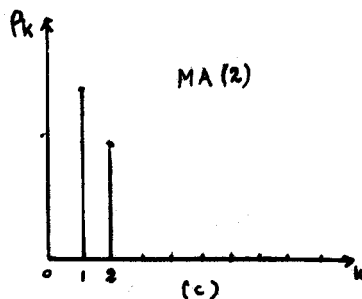
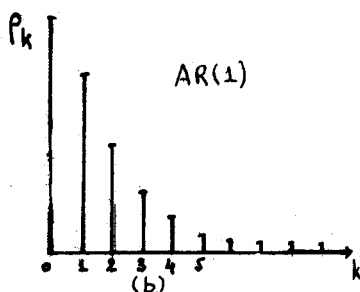
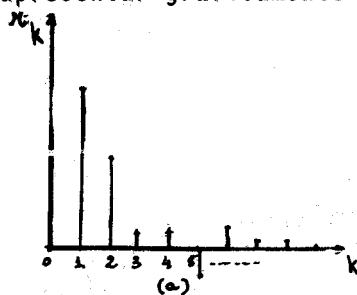


Fig. II5 - Função de autocorrelação (a) estimada, (b) teórica do modelo AR(1) e (c) teórica do modelo MA(2).

As Fig. II-5 (b) e (c) apresentam a função de autocorrelação dos modelos AR(1) e AR(2). Certamente ficaríamos em dúvida entre estes modelos, para representar a série.

Para ajudar na identificação foram propostas outras funções e critérios. Esses métodos de identificação se classificam de duas maneiras: métodos de reconhecimento de padrões e métodos automáticos. Até a Seção 5.6 serão tratados os métodos de reconhecimento de padrões e na Seção 5.7 os métodos automáticos.

5.2 Função de Autocorrelação Parcial

A autocorrelação parcial, ϕ_{kk} , mede a correlação entre Y_t e Y_{t+k} que não seja devido a dependência linear de Y_t e Y_{t+k} em $Y_{t+1}, \dots, Y_{t+k-1}$, que é determinada respectivamente por

$$Y_{t+k}^+ = \alpha_1 Y_{t+k+1} + \dots + \alpha_{k-1} Y_{t+1}$$

$$e \quad Y_t^+ = \beta_1 Y_{t+k+1} + \dots + \beta_{k-1} Y_{t+1}$$

então

$$\phi_{kk} = \text{Coor} \{ (Y_t - Y_t^+) (Y_{t+k} - Y_{t+k}^+) \}. \quad (31a)$$

A partir desta definição obtêm-se a expressão usual

$$\phi_{kk} = \frac{|P_k^*|}{|P_k|} \quad (31b)$$

onde

$$P_k = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-2} \\ \rho_{k-1} & \rho_{k-2} & \dots & \dots & 1 \end{bmatrix} \quad (32)$$

a matriz de autocorrelação e P_k^* é obtida substituindo a última coluna de P_k por

$$p_k = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix} \quad (33)$$

Uma definição mais operacional é obtida escrevendo-se as equações de Yule-Walker de um processo AR(k)

$$\rho_j = \phi_1 \rho_{j-1} + \dots + \phi_k \rho_{j-k} \quad j > 0$$

na forma matricial

$$\begin{bmatrix} \rho_1 \\ \vdots \\ \rho_k \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \dots & \rho_{k-2} \\ \vdots & \vdots & \dots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_k \\ \vdots \\ \phi_k \end{bmatrix}$$

$$p_k = P_k \phi. \quad (34)$$

A solução do sistema pela regra de Cramer fornece

$$\phi_k = \frac{|P_k^+|}{|P_k|} = \phi_{kk}$$

onde ϕ_{kk} é o último coeficiente de um modelo AR(k) ajustado

aos dados. Esta definição permite imediatamente obter o comportamento de ϕ_{kk} para cada um dos modelos: AR, MA e ARMA

- i) Suponha-se que o modelo verdadeiro é um AR(p) e que definamos a função de autocorrelação parcial ϕ_{kk} , $k=1,2,\dots$. Para isso ajustamos sucessivamente

$$Y_t = \phi_{11}Y_{t-1} + \epsilon_t \quad \Rightarrow \phi_{11}$$

$$Y_t = \phi_{21}Y_{t-1} + \phi_{22}Y_{t-2} + \epsilon_t \quad \Rightarrow \phi_{22}$$

$$\dots \dots \dots \vdots$$

$$Y_t = \phi_{p1}Y_{t-1} + \dots + \phi_{pp}Y_{t-p} + \epsilon_t \quad \Rightarrow \phi_{pp}$$

$$Y_t = \phi_{p+1,1}Y_{t-1} + \dots + \phi_{p+1,p+1}Y_{t-p} + \epsilon_t \Rightarrow \phi_{p+1,p+1}$$

pois o modelo é AR(p). Continuando sucessivamente verificamos pela mesma razão $\phi_{p+j,p+j} = 0$ $j = 1,2,\dots$. Portanto para um modelo AR(p) a função de autocorrelação parcial é não nula somente para os p primeiros lags.

- ii) Se o modelo verdadeiro é MA(q) ou ARMA(p,q), vemos que o mesmo é equivalente a um AR(∞) com coeficientes crescentes quando o lag cresce e portanto a função de autocorrelação parcial é infinita, tendendo a zero assintoticamente, i.e. é dominada por uma mistura de exponenciais e/ou senoides amortecidas dependendo das raízes de $\theta(B)$

- ii) Para modelos sazonais o comportamento da função de autocorrelação parcial é complicada e sua utilização não é mencionada em Box-Jenkins (1976). Seu comportamento para modelos sazonais é investigada em Hamilton e Watts (1982)

5.3 Inferência Sobre as Autocorrelações e Autocorrelações Parciais

Ao apontar a necessidade de introduzir outras ferramentas auxiliares para identificar os modelos para séries temporais, além da função de autocorrelação, mencionamos que uma das dificuldades com a mesma era que a função de autocorrelação era sujeita a grande variação amostral ao ser estimada.

Antes de continuar a introduzir outros métodos de identificação faremos um sumário dos estimadores e das distribuições amostrais para a autocorrelação e autocorrelação parcial de um processo estacionário ergódico

i) Estimação

A autocovariância estatística

$$C(k) = \gamma_k = E\{(Y_t - \mu)(Y_{t+k} - \mu)\}$$

é estimada pela covariância no tempo ou amostral

$$\hat{\gamma}_k^+ = \frac{1}{T-k} \sum_{t=1}^T (Y_t - \bar{Y})(Y_{t+k} - \bar{Y}) \quad (35)$$

ou

$$\hat{\gamma}_k = \frac{1}{T} \sum_{t=1}^T (Y_t - \bar{Y})(Y_{t+k} - \bar{Y}) \quad (36)$$

onde $\bar{Y} = \frac{1}{T} \sum_{t=1}^T Y_t$ e $k = 0, 1, \dots, (T-1)$.

Mostra-se (Priestley, 1981, pg. 323) que ambos os estimadores são viciados, porém o vício de $\hat{\gamma}_k$ é maior. Entretanto $\hat{\gamma}_k$ tem uma grande vantagem sobre $\hat{\gamma}_k^+$, qual

seja, fornece estimativas que são positivas definidas, o mesmo não ocorrendo com $\hat{\gamma}_k^+$. Lembre-se que esta é uma propriedade importante de γ_k .

A autocorrelação

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\text{Cov}(Y_t, Y_{t+k})}{\text{Var}(Y_t)}$$

é estimado por

$$\hat{\rho}_k = r_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \quad (37)$$

De forma análoga, poderíamos definir $\hat{\rho}_k^*$. Entretanto, pelas mesmas razões expostas para $\hat{\gamma}_k$, é preferível usar $\hat{\rho}_k$.

A autocorrelação parcial ϕ_{kk} pode ser estimada como o último coeficiente de um modelo AR(k) ajustado aos dados, ou através das equações de Yule-Walker

$$\rho_j = \phi_1 \rho_{j-1} + \dots + \phi_k \rho_{j-k} \quad j > 0$$

substituindo ρ_j por suas estimativas r_j e resolvendo para obter $\hat{\phi}_{kk}$. No caso de usar as equações de Yule-Walker, $\hat{\phi}_{kk}$ pode ser obtido recursivamente pela fórmula de Durbin (1960)

$$\begin{aligned} \hat{\phi}_{k+1,j} &= \hat{\phi}_{k,j} - \hat{\phi}_{k+1,k+1} \hat{\phi}_{k,k-j+1} \quad j = 1, \dots, k \\ \hat{\phi}_{k+1,k+1} &= \frac{r_{k+1} - \sum_{j=1}^k \hat{\phi}_{kj} r_{k+1-j}}{1 - \sum_{j=1}^k \phi_{kj} r_j} \end{aligned} \quad (38)$$

1) Distribuições Amostrais

Seja Y_t um processo linear da forma

$$Y_t = \mu + \sum_{u=-\infty}^{\infty} \psi_u \varepsilon_{t-u}$$

então (Priestley, 1981, pg. 337)

a) \bar{Y}_T é assintoticamente normal com média e variância

$$E(\bar{Y}) = \mu, \quad V(\bar{Y}) = \frac{1}{T} \sigma_\varepsilon^2 \left[1 + 2 \sum_{j=1}^{\infty} r_j \right] \quad (39)$$

b) r_j é assintoticamente normal com média e variância

$$E(r_j) = \rho_j$$

$$V(r_j) = \frac{1}{T} \left[\sum_{v=-\infty}^{\infty} \rho_v^2 + \rho_{v+j} \rho_{v-j} - 4\rho_j \rho_v \rho_{v-j} + 2\rho_v^2 \rho_j^2 \right] \quad (40)$$

este resultado é devido a Bartlett.

c) Para um modelo MA(q)

$$V(r_j) = \frac{1}{T} \left[1 + 2 \sum_{v=1}^q \rho_v^2 \right] \quad j > q \quad (41)$$

d) Para um modelo AR(p), Quenouille mostrou (Priestley, 1981, pg. 371) que $\hat{\phi}_{kk}$ ($k > p$) são aproximadamente independentes com média e variância

$$E(\hat{\phi}_{kk}) = 0 \quad V(\hat{\phi}_{kk}) = \frac{1}{T} \quad k > p. \quad (42)$$

Além disso, para T grande, $\hat{\phi}_{kk}$ tem distribuição aproximadamente normal.

e) Para um ruído branco

$$r_j \sim N\left(0, \frac{1}{T}\right) \quad (43)$$

Finalmente, testes de hipóteses e intervalos de confiança para ρ_j podem ser imediatamente obtidos a partir dos resultados (39) a (43). Por exemplo, para testar se Y_t é ruído branco, basta observar se $|r_j| < 2/\sqrt{T}$, $j > 0$. Para testar se Y_t é AR(p) basta verificar se $|\hat{\phi}_{kk}| < 2/\sqrt{N}$, $k > p$.

5.4 Autocorrelações Inversas

Cleveland (1972) definiu a função de autocorrelação inversa (FACI) ρI_k como a função de autocorrelação associada com o recíproco (inverso) do espectro da série. Para continuar no domínio do tempo, em vez de utilizar Transformada de Fourier (função característica), Chatfield (1979) utilizou Transformada-Z (função geratriz) para definir autocorrelação inversa. Essas definições podem ser especificadas de uma maneira alternativa equivalente. Dado um modelo SARIMA(p,d,q) x (P,D,Q)_s como (23)

$$\Phi(L^S)\phi(L) - \nabla_S^D \nabla^d Y_t = \Theta(L^S)\theta(L)\varepsilon_t$$

a função de autocorrelação inversa é definida como a função de autocorrelação do modelo SARIMA(q,d,p) x (Q,D,P)_s

$$\Theta(L^S)\theta(L)\nabla_S^D \nabla^d Y_t = \Phi(L^S)\phi(L)\varepsilon_t. \quad (44)$$

A função de autocorrelação inversa e propriedades são abordadas além de em Cleveland (1972) e Chatfield (1979) em Hipel, McLeod e Lennox (1977), McLeod, Hipel e Lennox (1977), Hosking (1980) e Bhansali (1980, 1983). A seguir, apresenta-

-se um resumo desses resultados.

i) FACI do modelo AR(p)

A FACI do modelo AR(p) em (6)

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t$$

tem a mesma forma que a FAC do modelo inverso MA(p)

$$Y_t = \varepsilon_t + \phi_1 \varepsilon_{t-1} + \dots + \phi_p \varepsilon_{t-p} \quad (45)$$

isto é, de (9) tem-se

$$\rho I_k = \begin{cases} \frac{\phi_k + \phi_1 \phi_{k-1} + \dots + \phi_{p-k} \phi_p}{1 + \phi^2 + \dots + \phi_p^2} & k = 1, \dots, p \\ 0 & k > p \end{cases} \quad (46)$$

A FAC do modelo AR(p) é uma mistura de exponenciais e/ou senoides amortecidas e, portanto, de difícil interpretação; entretanto a FACI tem somente p valores não nulos e é, portanto, de fácil interpretação. A função de autocorrelação parcial, FACP, neste caso também é de fácil interpretação pois também tem somente p valores não nulos. Logo, a FACI e a FACP tem o mesmo padrão de comportamento. Entretanto a FACI tem a vantagem de que, se alguns dos parâmetros intermediários do modelo AR são zero, a FACI nos lags correspondentes também é zero (Hipel, McLeod e Lennox, 1977 e Cleveland, 1977), o mesmo não ocorrendo com a FACP. Este resultado é importante principalmente na identificação de modelos sazonais.

ii) FACI do modelo MA(q)

A FACI do modelo MA(q) em (8)

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

tem a mesma forma que a FAC do modelo inverso AR(q)

$$Y_t = \theta_1 Y_{t-1} + \dots + \theta_q Y_{t-p} + \varepsilon_t \quad (47)$$

e satisfaz as equações de Yule-Walker inversas

$$\rho I_k = \theta_1 \rho_{k-1} + \dots + \theta_q \rho_{k-p}$$

ou $\theta(L)\rho I_k = 0$.

iii) FACI do modelo ARMA(p,q)

FACI do modelo ARMA(p,q) em (10)

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

tem a mesma forma que a FAC do modelo inverso ARMA(q,p)

$$Y_t = \theta_1 Y_{t-1} + \dots + \theta_q Y_{t-q} + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \dots + \phi_p \varepsilon_{t-p}$$

e satisfaz equações análogas a (11) e (12).

iv) Função de autocorrelação parcial inversa

Outra função que também ajuda na identificação é a função de autocorrelação parcial inversa, FACPI, (ϕI_{kk}) definida para um modelo SARIMA(p,d,q) x (P,D,Q)_S como a FACP do modelo inverso SARIMA(q,d,p)(Q,D,P)_S. Esta função, tem comportamento semelhante a FAC, da mesma forma que a FACI tem comportamento semelhante a FACP.

Uma forma equivalente de definir a FACPI, ϕ_{kk} é utilizando as equações de Yule-Walker inversas em forma matricial análoga a (34):

$$\begin{bmatrix} \rho I_1 \\ \vdots \\ \rho I_k \end{bmatrix} = \begin{bmatrix} 1 & \rho I_1 & \dots & \rho I_{k-1} \\ \rho I_1 & 1 & \dots & \rho I_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho I_{k-1} & \rho I_{k-2} & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi I_1 \\ \vdots \\ \phi I_{kk} \end{bmatrix} \quad (49)$$

v) Inferência sobre as autocorrelações inversas

Cleveland (1972) sugeriu dois métodos assintoticamente equivalentes para estimar ρI_k . O primeiro, através de estimativa consistente do espectro. O segundo, ajustando um modelo $AR(\ell)$ de ordem grande à série utilizando as equações de Yule-Walker e substituindo essas estimativas em (46). Hipel, McLeod e Lennox (1977) sugerem ℓ entre 10 e 40 para séries não sazonais e Chatfield (1979) o valor para o qual ρ_r sejam quase nulas.

A função de autocorrelação parcial inversa ϕ_{kk} é então estimada substituindo nas equações de Yule-Walker inversas ρI_j por $\hat{\rho} I_j$ e resolvendo para ϕ_{kk} .

A distribuição amostral de $\hat{\rho} I_j$ foi investigada em Hosking (1980) e Bhansali (1981) e a distribuição amostral de $\hat{\phi}_{kk}$ em Bhansali (1983). Resultados análogos a (39), (40) de Bartlett, (41) e (42) de Quenouille são obtidos. Por exemplo:

- Se Y_t segue um modelo $AR(p)$ tem-se que $\hat{\rho} I_k$ são in

dependentes e assintoticamente

$$\hat{\rho}I_k \sim N(0, \frac{1}{T}) \quad k > p \quad (50)$$

um resultado análogo ao de Quenouille.

- Se Y_t segue um modelo MA(q) tem-se assintoticamente

$$\hat{\phi}I_{kk} \sim N\left(0, \frac{1}{T}\left(1 + 2 \sum_{v=1}^q \rho I_v^2\right)\right), \quad k > q \quad (51)$$

onde a variância é obtida da expressão de $V(\hat{\rho}I_j)$, que é análoga a de Bartlett em (40), com ρ_j substituído por ρI_j .

Portanto testes de hipóteses e intervalos de confiança podem ser obtidos a partir desses resultados.

vi) A combinação dos resultados sobre FAC, FACP, FAI e FACP permite identificar claramente modelos AR e MA puros, entretanto o mesmo não ocorre para modelos mistos ARMA. Neste caso as funções não têm uma forma de fácil identificação. A seguir veremos algumas técnicas que procuram superar esta dificuldade.

5.5 Arranjos R^* e S^* e o Método "Corner"

i) Arranjos R^* e S^*

Gray, Kelley e McIntire (1978) mostraram como a estrutura de dependência de um modelo ARMA pode ser obtido transformando a função de autocorrelação ρ_k em duas funções, chamadas de arranjos R e S. Posteriormente

Woodward e Gray (1981) definiram funções R^* e S^* (apenas alterando a ordem dos elementos do arranjo) que ilustram mais claramente as características do modelo ARMA.

As funções R^* e S^* são escritas

$$R_j^*[\rho_i] = \frac{P_j[\rho_{i-j+1}]}{P_j[1, \rho_{i-j+1}]} = R_{ij}^*(\rho) \quad (52)$$

$$S_j^*[\rho_i] = \frac{P_{j+1}[1, \rho_{i-j+1}]}{P_j[\rho_{i-j+1}]} = S_{ij}^*(\rho) \quad (53)$$

onde

$$P_j[\rho_i] = \begin{vmatrix} \rho_i & \rho_{i+1} & \cdots & \rho_{i+j-1} \\ \rho_{i+1} & \rho_{i+2} & & \rho_{i+j} \\ \rho_{i+j-1} & \rho_{i+j} & & \rho_{i+2j-2} \end{vmatrix} \quad (54)$$

e

$$P_{j+i}[1, \rho_i] = \begin{vmatrix} 1 & & & & 1 & & & & & & & & 1 \\ \cdots & \cdots & \cdots & \cdots & \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & & P_j(\rho_i) & & & & & & & & \rho_{i+j} \\ & & & & & & & & & & & & \vdots \\ & & & & & & & & & & & & \rho_{i+2j-1} \end{vmatrix} \quad (55)$$

com $P_0(\rho_i) = 1$.

Expressões similares são também definidas quando ρ_i em (52) e (53) é substituído por $(-1)^i \rho_i$, neste caso os arranjos (52) (53) são denotados $R_j^*[(-1)^i \rho_i]$ e $S_j^*[(-1)^i \rho_i]$ e os determinantes (54) e (55) são denotados $P_j^*[(-1)^i \rho_i]$ e $P_{j+i}^*[1, (-1)^i \rho_i]$.

Os arranjos R^* e S^* podem ser mais facilmente obtidos através do seguinte esquema recursivo

$$R_{j+1}^*[\rho_i] = R_j^*[\rho_i] \left[\frac{S_j^*[\rho_i]}{S_j^*[\rho_{i-1}]} - 1 \right] \quad (56a)$$

$$S_j^*[\rho_i] = S_{j-1}^*[\rho_i] = \left| \frac{R_j^*[\rho_{i+1}]}{R_j^*[\rho_i]} - 1 \right| \quad (56b)$$

para $j = 1, 2, \dots$ e $k = 0, 1, 2, \dots$ com $S_0^*[\rho_i] = 1$ e $R_1^*[\rho_i] = \rho_i$.

Estimativas de R^* e S^* são obtidas substituindo ρ_k por sua estimativa r_k e são colocados na forma de uma tabela de dupla entrada com linhas i e colunas j .

Os arranjos R^* e S^* tem um comportamento padrão que permite a identificação de p e q em modelos ARMA bem como o grau de diferenciação e identificação de modelos sazonais. Diferentes enfoques podem ser utilizados para identificar p e q tendo em vista as diversas propriedades dos arranjos. Em geral Gray, Kelly e McIntire (1968) sugerem: que S^* é melhor para identificar p , R^* é melhor para identificar q e que ρ_i deve ser usado para dados com alta variação de curto prazo (alta frequência) e $(-1)^i \rho_i$ para dados com baixa variação a curto prazo (baixa frequência). Os autores também sugerem uma estatística $D(k, j)$ função dos R_{ij}^* e S_{ij}^* cujo valor máximo (em i, j) indica automaticamente a ordem mais adequada do modelo.

O método foi generalizado em Tucker (1982) e discutido também em Woodward e Gray (1978), Anderson (1980), Gray e Woodward (1981) e Woodward e Gray (1981) e é muito bem

resumido e apresentado em Salas e Obeysekera (1982).

1) O método "Corner"

Este método proposto por Bequin, Gourieroux e Monfort (1980), (ver também Gooijer e Heuts, 1981 e Bequin, Gourieroux e Monfort, 1981) consiste em utilizar o determinante

$$P(i,j) = \begin{vmatrix} \rho_i & \cdot & \rho_{i-1} & \cdot & \dots & \rho_{i-j+1} \\ \rho_{i+1} & \cdot & \rho_i & \cdot & \dots & \rho_{i-j+2} \\ \vdots & & \vdots & & & \vdots \\ \rho_{i+j-1} & \cdot & \rho_{i+j-2} & \cdot & \dots & \rho_i \end{vmatrix} \quad (57)$$

na seguinte forma:

- escolha um inteiro (grande) $k > \max(p,q)$ e calcule o determinante $P(i,j)$, $i,j = 1,2,\dots,k$.
- tabule $P(i,j)$ numa tabela de dupla entrada com linhas i e coluna j .
- a série segue um modelo ARMA se o arranjo P tem a estrutura

$i \backslash j$		$P+1$	\dots	k	
1	$P(1,1)$	$P(1,2)$	$P(1,p)$	$P(1,p+1)$	$P(1,k)$
2	$P(2,1)$				
\vdots					
q	$P(q,1)$				
$q+1$		0	0	0	
\vdots		\vdots	\vdots	\vdots	
k	$P(k,1)$	0	0	0	

(58)

O comportamento do arranjo não se altera se substitui-se ρ_k por $(-1)^k \rho_k$. Mostra-se que $-1 \leq P(i,j) \leq 1$... A estimação é feita substituindo ρ_k por r_k e o comportamento padrão para um ARMA é feito por inspeção visual ou através de teste estatístico χ^2 proposto pelos autores e baseado na distribuição amostral de r_j .

Para modelos MA(q) tem-se o comportamento

i \ j	1	...	k
q	X	X	X
q+1	0	0	0
⋮	⋮		
k	0	0

e como $P(h,1) = \rho_h$, $h < q$ tem-se o critério usual, isto é, verificar a FAC.

Para modelos AR(p) tem-se o comportamento

i \ j	p	p+1	k
1	X	0	0
	X	0	0
	⋮		
k	X	0	0

e lembrando (31b) e (32) verifica-se que

$$\phi_{kk} = (-1)^{h-1} \frac{P(1,h)}{P(0,h)} \quad h < p \quad (59)$$

e tem-se o critério usual, isto é, verificar a FACP, e o método é uma generalização do procedimento baseado na FACP.

A vantagem deste método é sua maior simplicidade em relação aos arranjos R^* e S^* .

5.6 Função de Autocorrelação Parcial Generalizada e Função de Autocorrelação Estendida

a) Função de Autocorrelação Parcial Generalizada

Para um modelo ARMA, de (11) e (12) as equações de Yule-Walker generalizadas são

$$\rho_k = \phi \rho_{k-1} + \dots + \phi_p \rho_{k-p} \quad k \leq q$$

$$\rho_k = \phi \rho_{k-1} + \dots + \phi_p \rho_{k-p} \quad k \geq q+1$$

A função de autocorrelação parcial generalizada (FACPG) é definida por

$$\phi_{kk,i}^{(i)} = \frac{|P^*(i,k)|}{|P(i,k)|} \quad (60)$$

de maneira similar a (34) onde $P(i,k)$ é dada em (57), isto é,

$$P(i,k) = \begin{vmatrix} \rho_i & \rho_{i-1} & \dots & \rho_{i-k+1} \\ \rho_{i+1} & \rho_i & \dots & \rho_{i-k+2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{i+k-1} & \rho_{i+k-2} & \dots & \rho_i \end{vmatrix}$$

e $P^+(i,k)$ é obtida substituindo a última coluna de $P(i,k)$ por

$$\rho^{(i)} = \begin{bmatrix} \rho_{i+1} \\ \rho_{i+2} \\ \vdots \\ \rho_{i+k} \end{bmatrix} \quad (61)$$

Portanto, de forma análoga a autocorrelação parcial ϕ_{kk} que é a solução de Cramer das equações de Yule-Walker para o último coeficiente de um modelo AR(k), aqui tem-se que a autocorrelação parcial generalizada $\phi_{kk}^{(i)}$ é a solução de Cramer das k equações de Yule-Walker generalizadas (i+1 até i+k) para o último coeficiente de um modelo ARMA(k,i).

Colocando $\phi_{kk}^{(i)}$ em uma tabela de dupla entrada com linhas i e colunas k, tem-se o seguinte padrão de comportamento para um modelo ARMA(p,q)

i \ k	1	...	p-1	p	p+1	p+2	
0	$\phi_{11}^{(0)}$		$\phi_{p-1,p-1}^{(0)}$	$\phi_{pp}^{(0)}$		
⋮	⋮						
q-1	$\phi_{11}^{(q-1)}$		$\phi_{p-1,p-1}^{(q-1)}$			
q	$\phi_{11}^{(q)}$		$\phi_{p-1,p-1}^{(q)}$	ϕ_p	0	0	0 (62)
q+1	$\phi_{11}^{(q+1)}$		$\phi_{p-1,p-1}^{(q+1)}$	ϕ_p	u	u	u
⋮				ϕ_p	u	u	u
					⋮	⋮	⋮

u - não definido

isto é, na linha q, após coluna p os valores são zero e na coluna p, após linha q os valores são constantes

e igual a ϕ_p .

A relação com os arranjos S^* é dada por

$$\phi_{kk}^{(i)} = (-1)^{k\ell+1} \frac{S_k^* [(-1)^i \rho_i]^\ell}{S_k^* [(-1)^{(-i-1)\ell} \rho_{-i-1}]^\ell} \quad (63)$$

onde $\ell = 0$, usou-se ρ_i e $\ell = 1$, usou-se $(-1)^i \rho_i$ e estimativas são obtidas substituindo ρ_i por r_i , nas equações de Yule-Walker ou em (63). Distribuições amostrais dos estimadores somente recentemente iniciaram a ser investigadas por Glasbey (1982) e Newbold e Bos (1983).

A função de autocorrelação parcial generalizada foi sugerida inicialmente por G.M. Jenkins (ver Granger e Newbold, 1977, pg. 108) e utilizada em Woodward e Gray (1981a), Jenkins e Alavi (1981) e Salas e Obeysebera (1982).

ii) Função de Autocorrelação Estendida

A função de autocorrelação (FACE) proposta recentemente por Tiao e Tsay (1981, 1982, 1983) e Tsay e Tiao (1982) é uma forma alternativa e de usar FACPG, $\phi_{kk}^{(i)}$ em (60).

Sabe-se que, para um modelo AR(p) estacionário, as estimativas dos parâmetros ϕ_{kk} obtidas por mínimos quadrados ou resolvendo as equações de Yule-Walker são assintoticamente consistentes e equivalentes (Chatfield, 1975, pg. 68). Para modelos ARMA(p,q) estacionários a FACPG pode ser obtida através das equações de Yule-Walker generalizadas ou dos arranjos S^* por (63).

A FACE é uma forma alternativa de usar $\phi_{kk}^{(i)}$ através de regressões iterativas. As vantagens da FACE são: é utilizada para modelos não estacionários, seus valores estão entre -1 e 1, sua distribuição amostral é conhecida e suas propriedades de consistência são mais estáveis (Tsay e Tiao, 1981).

Considere-se o modelo ARMA(p,q) não estacionário

$$\phi(L)U(L)Y_t = \theta_0 + \theta(L)\varepsilon_t$$

onde $U(L) = 1 - U_1L - \dots - U_dL^d$ com zeros no círculo unitário e $\phi(L)$ tem grau p' , note que $p = p' + d$ (por exemplo $U(L) = (1-L)^d$ para um modelo ARIMA(p',d,q)). As propriedades dos estimadores de mínimos quadrados ordinários (MQO) dos parâmetros AR ao ajustarmos um modelo AR(k) são

Modelo	Modelo Ajustado	MQO para parâmetros AR
ARMA(p,0)	AR(p)	Consistente
ARMA(p,q)	AR(d)	Consistente para U(L)
	AR(k) k > d	Tendencioso

Logo os estimadores MQO são consistentes para os parâmetros AR somente nos casos do modelo ser AR(p) estacionário ou não estacionário e no caso de ser ARMA(p,q) não estacionário com $\phi(L) = 1$, isto é, a parte AR é U(L).

Estimativas de ϕ_j consistentes podem ser obtidas das regressões iterativas a seguir.

- 0 - ésima regressão AR(p) iterativa

$$Y_t = \phi_{1p}^{(0)} Y_{t-1} + \dots + \phi_{pp}^{(0)} Y_{t-p} + \epsilon_{p,t}^{(0)}$$

e as estimativas de MQO de ϕ_j são consistentes se $q = 0$

- 1ª regressão AR(p) iterativa

$$Y_t = \phi_{1p}^{(1)} Y_{t-1} + \dots + \phi_{pp}^{(1)} Y_{t-p} + \beta_{p,t-1}^{(1)} \epsilon_{p,t-1}^{(0)} + \epsilon_{p,t}^{(1)}$$

e as estimativas de MQO de ϕ_j são consistentes se $q \leq 1$, ou $\phi(L) = 1$.

- 2ª regressão iterativa AR(p)

$$Y_t = \phi_{1p}^{(2)} Y_{t-1} + \dots + \phi_{pp}^{(2)} Y_{t-p} + \beta_1^{(2)} \epsilon_{p,t-2}^{(0)} + \beta_2^{(2)} \epsilon_{p,t-1}^{(1)} + \epsilon_{p,t}^{(2)}$$

e as estimativas de MQO de ϕ_j são consistentes se $q \leq 2$ ou $\phi(L) = 1$.

Na prática, as ordens (p,q) de um ARMA(p,q) são desconhecidas e portanto considera-se a j-ésima regressão iterativa de ordem k:

$$Y_t = \phi_k^{(j)} Y_{t-1} + \dots + \phi_{kk}^{(j)} Y_{t-k} + \beta_{k,t-j}^{(j)} \epsilon_{k,t-j}^{(0)} + \beta_{k,t-j+1}^{(j)} \epsilon_{k,t-j+1}^{(1)} + \dots + \beta_{k,t-1}^j \epsilon_{k,t-1}^{(j-1)} + \epsilon_{k,t}^{(j)} \quad (64)$$

$t = k+j+1, \dots, n, \quad j = 0, \dots, \quad k = 1, 2, \dots$

Baseado no resultado acima define-se a FACE da seguinte forma:

- a) Se $p = 0$, Y_t é $MA(q)$ e a FAC ρ_k é zero para $k > q$. Isto é,

$$\rho_k^{(0)} = 0 \quad \text{se } k > q \quad \text{e } p = 0 \quad (65)$$

- b) Se $p = 1$, Y_t é $ARMA(1, q)$ e se temos uma estimativa consistente de ϕ_1

$$Y_t - \hat{\phi}_1 Y_{t-1}$$

é $MA(q)$ assintoticamente e sua FAC tem a propriedade (65).

Entretanto para obter a estimativa eficiente $\hat{\phi}_1$ o número de regressões iterativas depende de q que é desconhecido, tem-se então as considerações

- Se $q = 0$ então $\phi_{11}^{(j)}$, $j \geq 0$ é consistente para ϕ_1 e portanto

$$W_t = Y_t - \hat{\phi}_{11}^{(0)} Y_{t-1}$$

é ruído branco e então

$$\rho_k(W_t^0) = 0 \quad k > 0$$

- Se $q = 1$ então $\phi_{11}^{(j)}$, $j \geq 0$ é consistente para ϕ_1 e portanto

$$W_t^{(1)} = Y_t - \hat{\phi}_{11}^{(1)} Y_{t-1}$$

é $MA(1)$ e então

$$\rho_k(W_t^1) = 0 \quad k > 1$$

- Se $q = q_0$, $\phi_{11}^{(j)}$, $j \geq q$ é consistente para ϕ_1 e portanto

$$W_t^{(q_0)} = Y_t - \hat{\phi}_{11}^{(q_0)} Y_{t-1}$$

e $MA(q)$ e então

$$\rho_k(W_t^{(q_0)}) = 0 \quad k > q_0$$

Baseado nestas considerações define-se a 1ª FACE de Y_t como

$$\rho_k^{(1)} = \rho_k(W_t^k) \quad k = 1, 2, \dots \quad (66)$$

onde $W_t^{(k)} = Y_t - \hat{\phi}_{11}^k Y_{t-1}$

tem-se então

$$\rho_k^{(1)} = 0 \quad k > q \quad \text{e} \quad p = 1.$$

De forma análoga define-se a p_0 -ésima FACE de Y_t como

$$\rho_k^{(p_0)} = \rho_k \left(Y_t - \sum_{\ell=1}^{p_0} \hat{\phi}_{\ell p_0} Y_{t-\ell} \right) \quad (67)$$

e então

$$\rho_k^{(p_0)} = 0 \quad k > q \quad \text{e} \quad p = p_0 \quad (68)$$

Estimativas da FACE são obtidas tomando-se valores p_0 e q_0 pre-especificados, calculando-se $\hat{\phi}_{\ell p}^{(j)}$ recursivamente (ver Tiao e Tsay, 1981) das estimativas de MQO de regressão estepwise $AR(p)$ $AR(p+1)$, ... $AR(p+q)$. Os erros padrões das estimativas são obtidos aproximadamente do re

sultado (40) de Bartlett e uma aproximação útil é

$$V(r_j^{(i)}) = \frac{1}{n - i - j} \quad (69)$$

A FACE amostral $r_j^{(i)}$ é colocada em uma tabela de dupla entrada com linhas i e colunas j . É conveniente também apresentar uma tabela análoga consistindo do símbolo indicador X denotando valores fora do intervalo de ± 2 desvios padrões e 0 para valores dentro do intervalo. Para um modelo ARMA(p, q) a tabela se apresentaria como

AR \ MA	0	1	q-1	q	q+1		
0	X	X		X	X	X	X
1	X	X		X	X	X	X
p	X	X		0	0	0	0
⋮							
p+1	X	X		X	0	0	0
				X	X	0	0
				X	X	X	0

(70)

5.7 Critérios Automáticos

Os diversos métodos de identificação apresentados requerem intervenção e julgamento do analista. Nesta Seção apresenta-se alguns dos critérios de seleção de modelos que diminuem significativamente a necessidade do julgamento do analista e podem ser automatizados em computador.

Anteriormente trabalhos nesta direção consistiam em ajustar sucessivamente modelos com crescente número de parâme

tros até obter resultado insignificante em testes de hipóteses sobre o último modelo. Entretanto, o problema de escolha de modelos no contexto de séries temporais é melhor abordado como um problema de decisão em lugar de teste de hipótese. O trabalho pioneiro nesta direção é devido a Akaike (1969). A idéia geral é que todos os critérios de qualidade do ajustamento melhoram mais ou menos regularmente quando aumenta-se o número de parâmetros do modelo. Entretanto, aumentar demasiadamente o número de parâmetros é indesejável pois a precisão das estimativas diminui e o modelo se torna de pouca utilidade. Para resolver esses dois objetivos, a princípio conflitantes, os métodos de escolha procuram penalizar os critérios de ajustamento pelo número de parâmetros incluídos. Os métodos são portanto uma formalização do critério da parcimônia na escolha de modelos. Exemplos de tal procedimento seriam

$$\hat{\sigma}_e^2(k)[1 + \chi(k)] \quad (71)$$

onde $\hat{\sigma}_e^2(k)$ é a variância amostral do resíduo ao ajustar um modelo com k parâmetros e $\chi(k)$ é uma função crescente de k , ou

$$-2 \log L_M + 2k \quad (72)$$

onde L_M é o máximo da verossimilhança do modelo com k parâmetros.

Desde o trabalho pioneiro de Akaike diversos trabalhos sobre o assunto foram desenvolvidos. A Figura II5 apresenta alguns dos critérios mais utilizados. Observe que a propriedade de não consistência de muitos dos critérios em ge

ral é devido a que o critério foi desenvolvido supondo que a série é gerada por um modelo infinito (por exemplo $AR(\infty)$) e que se deseja apenas um modelo que o aproxime (ver Poskitt e Tremayne, 1983b). Maiores detalhes podem ser vistos em Hipel (1981), Andel (1982), Hannan (1980), Poskitt and Tremayne (1983b). Em particular, Rissanen (1984) sugere dois novos critérios e Poskitt e Tremayne (1983a) usando o argumento Bayesiano mostram que priors para os modelos, fornecem os diferentes critérios.

Autores	Modelo	Critério a ser Minimizado	Propriedade
Akaike	AR(p)	$FPE(p) = \hat{\sigma}_p^2 \left(1 + \frac{p}{T}\right)$	Não consistente
Parzen	AR(p)	$CAT(p) = \begin{cases} - \left(1 + \frac{1}{T}\right) & p = 0 \\ \left\{ \frac{p}{T} \sum_{i=1}^p \hat{\sigma}_i^{-2} \right\} \hat{\sigma}_p^{-2} & p = 1, 2, 3, \dots \end{cases}$	Não consistente
Shibata	AR(p)	$S(p) = T \ln \hat{\sigma}_p^2 + T \ln(T + 2p)$	Não consistente
McClave	AR(p)	$M(p) = (T - p - 1) \frac{\hat{\sigma}_p^2 - \hat{\sigma}_{p+1}^2}{\hat{\sigma}_{p+1}^2}$	Não consistente
Akaike	ARMA(p,q)	$AIC(p,q) = T \ln \hat{\sigma}_{p,q}^2 + 2(p+q)$	Não consistente
Akaike, Rissanen e Schwartz	ARMA(p,q)	$BIC(p,q) = T \ln \hat{\sigma}_{p,q}^2 + (p+q) \ln T$	Consistente
Hannan-Quinn	ARMA(p,q)	$\phi_c(p,q) = T \ln \hat{\sigma}_{p,q}^2 + (p+q) c \ln \ln T \quad c > 2$	Consistente
Gray-Kelley-MacIntire	ARMA(p,q)	$D(p,q) = - \left \frac{S_{p+1}^{*(\rho_{-q-p-1})}}{S_p^{*(\rho_{-q-p})}} \right / v \quad \text{onde}$ $v = \left[\frac{S_p^{*(\rho_{-q-p+1})} + S_{p+1}^{*(\rho_{-q-p})}}{S_p^{*(\rho_{-q-p+2})} + S_p^{*(\rho_{-q-p+1})}} \right]^2 + \sum_{i=1}^3 \left[\frac{R_{p+1}^{*(\rho_{-q-p-i})}}{R_q^{*(\rho_{-q-p+i+1})}} \right]^2 + \left[\frac{R_{p+1}^{*(\rho_{-q-p+1})}}{R_p^{*(\rho_{-q-p+1+1})}} \right]^2$ <p>onde $S_{-1}^{*(\rho_q)} = 0$ e $R_0^{(\rho_q)} = 1$ e a expressão é válida para $(-1)^m \rho_m$ no lugar de ρ_m.</p>	Consistente

6. Estimação e Verificação

6.1 Estimação dos Parâmetros

Após especificar a forma do modelo, isto é, o valor de (p,q) no modelo ARMA, é necessário estimar seus parâmetros. Nesta Seção apresenta-se somente um resumo dos resultados para estimação já que esta fase é obtida automaticamente em computadores. Maiores detalhes podem ser vistos em Box Jenkins (1976), Priestley (1981) e Abraham e Ledolter (1983)

Considere-se o modelo ARMA(p,q) escrito como

$$\varepsilon_t = Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (73)$$

Supondo ε_t independentes, com distribuição normal com média zero e variância σ_ε^2 , a distribuição de probabilidade conjunta de $\varepsilon = (\varepsilon_1, \dots, \varepsilon_t)'$ é

$$f(\varepsilon/\phi, \theta, \sigma_\varepsilon^2) = (2\pi\sigma_\varepsilon^2)^{-n/2} \exp \left\{ -\frac{1}{2\sigma_\varepsilon^2} \left(\sum_{t=1}^T \varepsilon_t^2 \right) \right\} \quad (74)$$

onde $\phi = (\phi_1, \dots, \phi_p)'$ e $\theta = (\theta_1, \dots, \theta_q)'$.

Dadas as observações $Y = (Y_1, \dots, Y_t)$ a função de verossimilhança para os parâmetros $(\phi, \theta, \sigma_\varepsilon^2)$ é obtida de (74) (Nbold, 1974).

$$L(\phi, \theta, \sigma_\varepsilon^2/Y) = g_1(\phi, \theta, \sigma_\varepsilon^2) \exp \left\{ -\frac{1}{2\sigma_\varepsilon^2} S(\phi, \theta) \right\} \quad (75)$$

onde g_1 é uma função dos parâmetros $(\phi, \theta, \sigma_\varepsilon^2)$ e

$$S(\phi, \theta) = \sum_{t=1-p-q}^T E^2(u_t/Y, \phi, \theta, \sigma_\varepsilon^2) \quad (76)$$

Em (76), $E(u_t/Y, \phi, \theta, \sigma_\varepsilon^2)$ é a esperança condicional de u_t da dos Y , ϕ , θ e σ_ε^2 , e

$$u_t = \begin{cases} \varepsilon_t & t = 1, 2, \dots, T \\ g_2(\varepsilon^*, Y^*) & t \leq 0 \end{cases} \quad (77)$$

onde g_2 é uma função linear dos valores não observados $\varepsilon^* = (\varepsilon_{1-q}, \dots, \varepsilon_{-1}, \varepsilon_0)'$ e $Y^* = (Y_{1-p}, \dots, Y_{-1}, Y_0)$ necessários para calcular $\varepsilon_1, \dots, \varepsilon_T$ na equação (73). As funções g_1 e g_2 dependem do modelo particular

i) Estimadores de Máxima Verossimilhança Exata

Estimativas de máxima verossimilhança para os parâmetros são obtidos maximizando a função (75). Entretanto a expressão (75) não pode ser obtida numa forma simples e diversos algoritmos têm sido propostos para obtê-la. McLeod (1977) apresenta uma aproximação e Dent, (1977), Ali (1977) Ansley (1979), Ljung e Box (1979) sugerem outros algoritmos, sendo o de Ansley (1979) considerado o mais eficiente. Alternativos algoritmos usando Filtro de Kalman foram estudados por Gardner, Harvey e Phillips (1980) e Harvey e Phillips (1976), Pearlman (1980) e Solo (1982).

O uso da função de verossimilhança exata é aconselhável não só por não envolver aproximações mas também porque, no caso de raízes do polinômio MA perto da região de inversibilidade, os outros métodos tendem a produzir não convergência nos métodos iterativos de maximização.

ii) Maxima Verossimilhana Condicionada ou Mınimos Quadrados Condicionados

Se fixarmos valores iniciais ϵ^* e Y^* , por exemplo tomar $\epsilon^* = 0'$ e $Y^* = (Y_1, \dots, Y_p)$, (75) se torna funao de verossimilhana condicionada

$$L_C(\phi, \theta, \sigma_\epsilon^2) = (2\pi\sigma_\epsilon^2)^{-n/2} \exp -\frac{1}{2\sigma_\epsilon^2} \sum_{t=p+1}^T S_C(\phi, \theta) \quad (77)$$

onde

$$S_C(\phi, \theta) = \sum_{t=p+1}^T \epsilon_t^2 \quad (78)$$

e a soma dos quadrados condicionada. Maximizar (77) em (ϕ, θ) e equivalente a minimizar (78).

iii) Mınimos Quadrados Nao Condicionado

Quando o numero T de observaoes e suficientemente grande a funao $g_1(\phi, \theta, \sigma_\epsilon^2)$ em (75) e dominada pelo termo exponencial e portanto uma estimativa pode ser obtida abandonando a funao $g_1(\phi, \theta, \sigma_\epsilon^2)$ e maximizando em (ϕ, θ) o termo exponencial, ou equivalentemente minimizando

$$S(\phi, \theta). \quad (79)$$

Neste caso e necessario obter estimativas para ϵ^* e Y^* para obter $g_2(\epsilon^*, Y^*)$. Existem dois metodos disponiveis: mınimos quadrados e previsao inversa (ver Abraham e Ledolter, 1983 pg. 253). O metodo de previsao para tras consiste em notar que o modelo inverso

$$Y_t = \phi_1 Y_{t+1} + \dots + \phi_p Y_{t+p} + \epsilon_t + \theta_1 \epsilon_{t+1} + \dots + \theta_q \epsilon_{t+q} \quad (80)$$

tem a mesma função de autocorrelação que o modelo (73). Isto significa que o modelo ARMA é reversível no tempo, isto é, a série Y_1, Y_2, \dots, Y_T (do início para o fim) e a série Y_T, Y_{T-1}, \dots, Y_1 (do fim para o início) seguem o mesmo modelo ARMA.

Portanto, usando a série Y_T, Y_{T-1}, \dots, Y_1 e o modelo ARMA identificado podemos fazer previsões de Y_0, Y_1, \dots e obter $\varepsilon_0, \varepsilon_{-1}, \dots$.

Comparação entre os diversos métodos de estimação é apresentada em Aigner (1971) Dent e Min (1978), Beveridge (1981) e Damsleth (1983).

v) Estimação Não Linear

A função de verossimilhança exata (75) é não linear nos parâmetros assim como as funções $S_c(\phi, \theta)$, e $S(\phi, \theta)$ quando o modelo possui termos MA. Nestes casos os estimadores são obtidos por métodos iterativos de otimização, em geral os algoritmos mais usados são os de Marquardt e Powell. (ver Bard, 1974). Valores iniciais para esses algoritmos podem ser obtidos das expressões das funções de autocorrelações substituindo ρ_j por r_j .

Para ilustrar considere-se a estimação de θ no modelo MA(1)

$$Y_t = \varepsilon_t + \theta \varepsilon_{t-1}$$

Neste caso sabe-se que $\rho_1 = \frac{\theta}{1 + \theta^2}$ e para que o mo

delo seja inversível $-1 < \theta < 1$. Além disso o modelo é equivalente ao modelo $AR(\infty)$

$$Y_t = \theta Y_{t-1} - \theta^2 Y_{t-2} + \dots + \varepsilon_t$$

isto é,

$$\varepsilon_t = Y_t - \theta Y_{t-1} + \theta^2 Y_{t-2} - \theta^3 Y_{t-3} + \dots$$

Neste caso que temos um só parâmetro pode-se empregar um processo de otimização por busca. Para isso escolhe-se valores, por exemplo,

$$\theta = -0.9, -0.8, \dots, 0.8, 0.9$$

e suponha-se que substituindo r_1 por ρ_1 na FAC obtém-se $\theta = 0.5$, que servirá para iniciar o processo. Calcula-se então

$$\hat{\varepsilon}_t(0.5) = Y_t - 0.5Y_{t-1} + 0.5^2 Y_{t-2} - 0.5^3 Y_{t-3}$$

isto é, abandona-se termos muito pequenos (aqui supomos $0.5^4 Y_{t-4}$ em diante, pequenos).

Conhecendo-se Y_t , calcula-se $S_C(0.5) = \sum_{i=4}^T \hat{\varepsilon}_i^2(0.5)$, toma-se outro valor de θ , θ_2 perto de 0.5 e calcula-se $S_C(\theta_2)$ e assim sucessivamente até varrer todos os valores. Escolhe-se como estimativa de mínimos quadrados condicionados de θ , aquele valor para o qual $S_C(\theta)$ é mínimo.

Procedimento análogo é feito para um modelo $ARMA(1,1)$, por exemplo

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$$

Suponha-se $Y_0 = 0$ e $\varepsilon_0 = 0$, as equações para ε_t seriam

$$\varepsilon_1 = Y_1$$

$$\varepsilon_2 = Y_2 - \phi_1 Y_1 - \theta \varepsilon_1$$

$$\vdots$$

$$\varepsilon_T = Y_T - \phi_1 Y_{T-1} - \theta \varepsilon_{T-1}$$

Neste caso escolheremos uma malha de valores no retângulo $-1 < \phi, \theta < 1$.

Para modelos com mais parâmetros outros procedimentos de otimização seriam utilizados.

A estimação dos parâmetros de um modelo $AR(p)$ por mínimos quadrados não envolve otimização não linear resolvendo-se num problema usual de regressão linear

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t$$

e as estimativas são consistentes. Uma alternativa proposta por Durbin (1959, 1960) para estimar os parâmetros de modelos MA e ARMA consiste em estimar um modelo AR de ordem grande e explorar a dualidade entre os modelos MA e ARMA de ordem finita com os $AR(\infty)$.

v) Variância das estimativas

Uma vez obtido as estimativas $\hat{\eta} = (\hat{\phi}, \hat{\theta})$ a estimativa de σ_ε^2 é

$$\hat{\sigma}_\epsilon^2 = \frac{S(\hat{\eta})}{T} = \frac{S(\hat{\phi}, \hat{\theta})}{T} \quad (81)$$

Dos resultados clássicos de máxima verossimilhança a matriz de variância-covariância de $\hat{\eta}$ (inverso da matriz de informação de Fisher) é

$$I^{-1}(\hat{\phi}, \hat{\theta}) = 2\hat{\sigma}_\epsilon^2 \left(\frac{\partial^2 S(\hat{\eta})}{\partial \eta_i \partial \eta_j} \right) \quad (82)$$

6.2 Verificação da Adequabilidade

Após identificar e estimar eficientemente os parâmetros do modelo, verifica-se a adequabilidade do mesmo antes de usá-lo em previsão. Nesta Seção apresenta-se alguns testes mais usados. Uma revisão sobre o assunto é apresentada em Newbold (1982) e comparações empíricas sobre alguns testes é apresentada em Clark e Godolphin (1983).

i) Testes de Sobreajustamento

Estes testes são baseados na estimação de um modelo mais elaborado, testando-se a significância dos parâmetros adicionais. Em geral adiciona-se um parâmetro adicional de cada vez, evitando-se desta forma incluir parâmetros que possam se anular. (Por exemplo, $1 - \phi L$ e $1 - \theta L$ com $\phi = \theta$).

Poskitt e Tremayne (1980) evitam este problema de estimabilidade utilizando o teste dos multiplicadores de Lagrange para construir testes de ARMA(p,q) contra ARMA(p+k₁,q+k₂).

(i) Testes Sobre os Resíduos

Dado o modelo ARMA(p,q) pode-se calcular

$$\varepsilon_t = \theta^{-1}(L)\phi(L)Y_t \quad \text{e} \quad \hat{\rho}_k(\varepsilon_t)$$

Como ε_t é ruído branco tem-se que a distribuição de

$$\hat{\rho}_k(\varepsilon_t) \sim N(0, \frac{1}{T}) \quad (83)$$

e portanto pode-se construir testes e intervalos para $\rho_k(\varepsilon_t)$. Entretanto os ϕ 's e θ 's são estimados e o que se calcula é

$$\hat{\varepsilon}_t = \hat{\theta}^{-1}(L)\hat{\phi}(L)Y_t \quad \text{e} \quad \hat{\rho}_k(\hat{\varepsilon}_t) \quad (84)$$

Embora para um número de observações grande, $\hat{\varepsilon}_t$ deva se comportar como ε_t deve-se ter cuidado ao usar a distribuição $N(0, \frac{1}{T})$. Box e Pierce (1970) mostraram que $\hat{\varepsilon}_t$ se comporta aproximadamente como um AR(p+q) se o modelo para Y_t é ARMA(p,q).

Entretanto, os seguintes testes têm sido bastante empregados com sucesso:

a) Testes individuais de $\rho_k(\hat{\varepsilon}_t)$.

Plotar faixas $\pm 2/\sqrt{T}$ no gráfico da FAC de $\hat{\varepsilon}_t$ e considerar como zero os valores de $\rho_k(\hat{\varepsilon}_t)$ dentro das faixas.

b) Teste de Portmanteau

Este teste verifica se em conjunto as primeiras M (gran

de) autocorrelações são significantes. O teste é baseado na estatística

$$Q' = n \sum_{j=1}^M \hat{\rho}_j^2(\hat{\varepsilon}_t) \sim \chi_{M-p-q}^2 \quad (85)$$

onde M é escolhido de tal forma que $\rho_j \approx 0$, $j > M$, e a hipótese é rejeitada se Q é maior que o valor de χ^2 tabelado.

Diversos autores verificaram que a distribuição de Q se afasta muito da distribuição χ^2 . Ljung e Box (1979) então sugeriram como suficientemente correta a estatística

$$Q = n(n+2) \sum_{j=1}^M (n-j)^{-1} \hat{\rho}_j^2(\hat{\varepsilon}_t). \quad (86)$$

c) Teste do periodograma acumulado

Este teste verifica se a função de distribuição espectral amostral se ajusta a função de distribuição do ruído branco. Neste teste calcula-se o periodograma

$$I(r) = \frac{2}{T} \left\{ \left[\sum_{i=1}^T \hat{\varepsilon}_i \cos 2\pi \frac{ri}{T} \right]^2 + \left[\sum_{i=1}^T \hat{\varepsilon}_i \sin 2\pi \frac{ri}{T} \right]^2 \right\} \quad i=1 \quad (87)$$

para $r = 1, \dots, \left[\frac{T-2}{2} \right]$ e o periodograma acumulado

$$C(j) = \sum_{r=1}^j I(r) / N \hat{\sigma}_{\varepsilon}^2 \quad i = 1, \dots, \left[\frac{T-2}{2} \right] \quad (88)$$

Plota-se $C(j)$ contra $\frac{j}{T}$. Se a série $\hat{\varepsilon}_t$ é ruído branco o plot deve seguir a reta que une os pontos $(0,0)$ e $(0.5,1)$. A adequabilidade do ajustamento é verificada usando o teste (bandas) de Smirnov-Kolmogorov. Para isso calcula-se retas distanciadas $\pm D$ da reta teórica

onde $D = K_\alpha / \sqrt{\ell}$, α é o nível de significância, $\ell = \frac{T-2}{2}$ (T par) ou $\ell = \frac{T-1}{2}$ (T ímpar) e

α	10	5	1
K_α	1.22	1.36	1.63

d) Portmanteau dos Resíduos ao Quadrado

Recentemente McLeod e Li (1983) sugeriram usar a estatística Q' em (85) com as autocorrelações de $\hat{\epsilon}_t^2$, isto é, $\hat{\rho}_k(\hat{\epsilon}_t^2)$ no lugar de $\rho_k(\epsilon_t)$.

6.3 Uso dos Resíduos para Alterar o Modelo

Suponha que a análise dos resíduos indicou não adequabilidade do modelo original

$$\phi(L)Y_t = \theta(L)\epsilon_t$$

e que o modelo identificado para o resíduo foi

$$\phi'(L)\epsilon_t = \theta'(L)\eta_t \quad (89)$$

Substituindo-se $\epsilon_t = \phi^{-1}(L)\theta'(L)\eta_t$, o novo modelo tentativo é

$$\phi(L)\phi'(L)Y_t = \theta(L)\theta'(L)\eta_t. \quad (100)$$

7. Previsão

7.1 Representações e Notação

O objetivo final da modelagem é a utilização do modelo em previsão. Como modelos $ARIMA(p,d,q)$ e $SARIMA(p,d,q)$ $(P,D,Q)_s$ podem ser considerados respectivamente como modelos $ARMA(p+d,q)$ e $ARMA(p+d+D+Ps, q+Qs)$ os resultados serão apresentados para o modelo ARMA.

Um modelo $ARMA(p,q)$ estacionário e inversível pode ser representado nas três formas alternativas

$$i) \phi(L)Y_t = \theta(L)\varepsilon_t \quad (101)$$

$$ii) Y_t = \psi(L)\varepsilon_t \quad \text{onde} \quad \psi(L) = \phi^{-1}(L)\theta(L) \quad (102)$$

$$iii) \pi(L)Y_t = \varepsilon_t \quad \text{onde} \quad \pi(L) = \theta^{-1}(L)\phi(L) \quad (103)$$

A forma mais simples de escrever o modelo é a forma (i) e (iii). A forma (ii) é usada para resultados teóricos.

A seguinte notação será usada

$\hat{Y}_t(\ell)$ - previsão de $Y_{t+\ell}$ feita no período t

$$E[Y_{t+\ell}] = E\{Y_{t+\ell}/Y_t, Y_{t-1}, \dots\}$$

$e_t(\ell) = Y_{t+\ell} - \hat{Y}_t(\ell)$ - erro de previsão ℓ passos a frente.

7.2 Variância da Previsão e Intervalos de Confiança

Estamos interessados em prever Y_{t+l} como função linear dos valores presente e passados Y_t, Y_{t+1}, \dots que por sua vez podem ser escritos como funções lineares dos resíduos presente e passados $\epsilon_t, \epsilon_{t-1}, \dots$ (pois $Y_t = \psi(B)\epsilon_t$). Logo

$$\hat{Y}_t(l) = \psi_l^+ \epsilon_t + \psi_{l+1}^+ \epsilon_{t-1} + \dots$$

Deseja-se então obter ψ_{l+j}^+ , $j = 0, 1, \dots$ tal que minimize o erro quadrático médio de previsão

$$EQM = E\{e_t^2(l)\} = E\{Y_{t+l} - \hat{Y}_t(l)\}^2 \quad (104)$$

Tem-se

$$Y_{t+l} = \epsilon_{t+l} + \psi_1 \epsilon_{t+l-1} + \psi_2 \epsilon_{t+l-2} + \dots + \psi_l \epsilon_t + \psi_{l+1} \epsilon_{t-1} + \dots$$

logo

$$e_t(l) = Y_{t+l} - \hat{Y}_t(l) = \epsilon_{t+l} + \psi_1 \epsilon_{t+l-1} + \dots + \psi_{l-1} \epsilon_{t+1} + [\psi_l - \psi_l^+] \epsilon_t + [\psi_{l+1} - \psi_{l+1}^+] \epsilon_{t-1} + \dots$$

Observe-se que $E\{e_t(l)\} = 0$ logo $\hat{Y}_t(l)$ é um previsor não viciado de Y_{t+l} então

$$\begin{aligned} E\{\hat{Y}_t(l)\} &= E\{e_t^2(l)\} = E\{\epsilon_{t+l} + \dots + \psi_{l-1} \epsilon_{t+1} + [\psi_l - \psi_l^+] \epsilon_t + \dots\}^2 \\ &= \sigma_\epsilon^2 \sum_{j=0}^{l-1} \psi_j^2 + \sigma_\epsilon^2 \sum_{i=0}^{\infty} [\psi_{l+i} - \psi_{l+i}^+]^2 \quad (\psi_0 = 1) \end{aligned}$$

cujo mínimo é obtido para

$$\psi_{l+i}^+ = \psi_{l+i}$$

Tem-se então

$$V\{\hat{Y}_t(\ell)\} = \sigma_\varepsilon^2 \sum_{j=0}^{\ell-1} \psi_j^2 \quad (105)$$

e se ε_t tem distribuição normal, intervalos de confiança para $Y_{t+\ell}$ são obtidos de

$$\hat{Y}_t(\ell) \pm K_\alpha \sigma_a \left(\sum_{i=0}^{\infty} \psi_i^2 \right)^{1/2} \quad (106)$$

onde K_α é a ordenada da normal com nível de significância $\alpha\%$.

Observe que para séries estacionárias $\sum_{i=0}^{\infty} \psi_i^2 < M$ e logo a amplitude do intervalo se estabiliza enquanto que para séries não estacionárias o mesmo não ocorre.

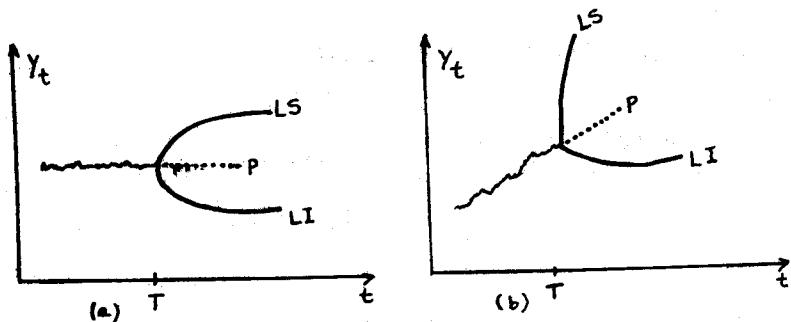


Fig. II6 - Intervalos de confiança (a) Série estacionária, (b) Série não estacionária (P - previsão, LI - limite inferior e LS - limite superior).

7.3 Cálculo e Atualização das Previsões

i) Determinação de ε_t

Os resíduos ε_t podem ser obtidos através das previ

sões e erros um passo a frente

$$e_t(1) = Y_{t+1} - \hat{Y}_t(1) = \varepsilon_{t+1} \quad (107)$$

sendo os ε_t independentes vê-se que em previsões com EQM mínimo os erros de previsão um passo a frente são independentes.

ii) Cálculo das previsões

Tem-se inicialmente

$$\begin{aligned} E[Y_{t+l}] &= E[\varepsilon_{t+l}] + \psi_1 E[\varepsilon_{t+l-1}] + \dots + \\ &+ \psi_l E[\varepsilon_t] + \psi_{l+1} E[\varepsilon_{t-1}] + \dots \\ &= \psi_l \varepsilon_t + \psi_{l+1} \varepsilon_{t-1} + \dots = \hat{Y}_t(l) \end{aligned}$$

pois
$$E[\varepsilon_{t+j}] = \begin{cases} 0 & j = 1, 2, \dots \\ \varepsilon_{t+j} & j = 0, -1, -2, \dots \end{cases}$$

A equação de previsão para

$$\begin{aligned} Y_{t+l} &= \phi_1 Y_{t+l-1} + \dots + \phi_p Y_{t+l-p} + \\ &+ \varepsilon_{t+l} + \theta_1 \varepsilon_{t+l-1} + \dots + \theta_q \varepsilon_{t+l-q} \end{aligned}$$

é

$$\begin{aligned} \hat{Y}_t(l) &= \phi_1 E[Y_{t+l-1}] + \dots + \phi_p E[Y_{t+l-p}] + \\ &+ E[\varepsilon_{t+l}] + \theta_1 E[\varepsilon_{t+l-1}] + \dots + \theta_q E[\varepsilon_{t+l-q}] \quad (108) \end{aligned}$$

fazendo-se

$$E[Y_{t+j}] = \begin{cases} Y_{t+j} & j = 0, -1, -2, \dots \\ \hat{Y}_t(j) & j = 1, 2, \dots \end{cases} \quad (109)$$

$$E[\varepsilon_{t+j}] = \begin{cases} 0 & j = 1, 2, \dots \\ \varepsilon_{t+j} = Y_{t+j} - \hat{Y}_{t+j-1}(1) & j = 0, -1, -2, \dots \end{cases}$$

iii) Aplicações

a) Para o modelo AR(p) tem-se

$$\hat{Y}_t(1) = \phi_1 Y_t + \dots + \phi_p Y_{t-p+1}$$

$$\hat{Y}_t(2) = \phi_1 \hat{Y}_t(1) + \dots + \phi_p Y_{t-p+2}$$

$$\vdots$$

$$\hat{Y}_t(p) = \phi_1 \hat{Y}_t(p-1) + \dots + \phi_p Y_t$$

$$\hat{Y}_t(p+1) = \phi_1 \hat{Y}_t(p) + \dots + \phi_p \hat{Y}_t(1)$$

Em geral

$$\hat{Y}_t(p+j) = \phi_1 \hat{Y}_t(p+j+1) + \dots + \phi_p \hat{Y}_t(j)$$

b) Para o modelo MA(q) tem-se

$$\hat{Y}_t(1) = \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q+1}$$

$$\hat{Y}_t(2) = \theta_2 \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q+2}$$

$$\dots$$

$$\hat{Y}_t(q) = \theta_q \varepsilon_t$$

$$\hat{Y}_t(q+j) = 0 \quad j = 1, 2, \dots$$

No caso de $\mu \neq 0$ soma-se θ_0 às equações acima.

c) Para o modelo ARMA(p,q)

$$\hat{Y}_t(1) = \phi_1 Y_t + \dots + \phi_p Y_{t-p+1} + \theta_1 \epsilon_t + \dots + \theta_q \epsilon_{t-q+1}$$

$$\hat{Y}_t(2) = \phi_1 \hat{Y}_t(1) + \dots + \phi_p Y_{t-p+2} + \theta_2 \epsilon_t + \dots + \theta_q \epsilon_{t-q+2}$$

Em geral para $j = 1, 2, \dots$

$$q \geq p \quad \hat{Y}_t(q+j) = \phi_1 \hat{Y}_t(q+j-1) + \dots + \phi_p \hat{Y}_t(q+j-p) \quad (110)$$

$$q < p \quad \hat{Y}_t(p+j) = \phi_1 Y_t(p+j-1) + \dots + \phi_p \hat{Y}_t(j).$$

iv) Atualização das previsões

Quando observa-se Y_{t+1} , as previsões $\hat{Y}_t(1), \dots, \hat{Y}_t(\ell)$ podem ser facilmente atualizadas para a nova origem $t+1$ usando a relação

$$\hat{Y}_{t+1}(j) = \hat{Y}_t(j+1) + \psi_j \epsilon_{t+1} \quad j = 1, \dots, \ell \quad (111)$$

onde $\epsilon_{t+1} = Y_{t+1} - \hat{Y}_t(1) = e_t(1)$.

Este resultado é obtido lembrando-se que

$$\hat{Y}_t(j+1) = \psi_{j+1} \epsilon_t + \psi_{j+2} \epsilon_{t-1} + \psi_{j+3} \epsilon_{t-2} + \dots$$

$$\hat{Y}_{t+1}(j) = \psi_j \epsilon_{t+1} + \psi_{j+1} \epsilon_t + \psi_{j+2} \epsilon_{t-1} + \dots$$

7.4 Robustez e Parcimônia

Os resultados anteriores sobre previsão foram obtidos supondo conhecidos os parâmetros ϕ 's, θ 's e σ_ϵ^2 . Na prática esses valores são substituídos por estimativas. Entretanto, os resultados anteriores sobre previsão são robustos com relação a erros nos valores dos parâmetros, isto é, a previsão e

seus intervalos de confiança não se alteram muito se usarmos estimativas eficientes no lugar dos parâmetros. Por outro lado, Ledolter e Abraham (1981) apontaram a importância de identificar corretamente o modelo, já que a cada parâmetro desnecessário incluído no modelo a variância da predição um passo a frente aumenta na ordem de σ_{ϵ}^2/T , onde T é o número de observações.

7.5 Combinação de Previsões

Em vista de que as previsões obtidas têm erro quadrático mínimo parece não haver espaço para melhorar a previsão. Entretanto, quase sempre não é possível modelar exatamente uma situação real e é possível que uma combinação da previsão obtida com modelos ARMA com outros métodos de previsão possam diminuir o erro da previsão (ver Granger e Newbold, 1977 e Nelson, 1972).

Um exemplo de combinação de previsão obtida com um modelo ARMA com uma previsão subjetiva proporcionando melhora substancial no erro de previsão é apresentada em Pereira, Coqueiro e Perrota (1983).

7.6 Relação com Outros Métodos

Abraham e Ledolter (1983) apresentam a relação entre previsões obtidas dos modelos ARMA e com os conhecidos métodos automáticos de previsão (Pereira 1980), mostrando que estes últimos são ótimos apenas para uma subclasse dos modelos

ARMA.

A relação entre modelos ARMA e métodos de decomposição é apresentada em Cleveland e Tiao (1976).

8. Outras Metodologias

Nesta seção faz-se uma apresentação de algumas metodologias alternativas a de Box-Jenkins recentemente propostas. Serão indicados apenas os principais pontos necessários para entender os princípios básicos e como utilizá-las. Não serão discutidos os méritos e deméritos de cada uma. Entretanto, essas novas metodologias podem ser utilizadas automaticamente em computadores e com o aumento crescente da capacidade dessas máquinas certamente algumas dessas metodologias serão de grande utilidade em futuro próximo.

8.1 Metodologia Recursiva de Hannan-Rissanen

A ordem (p, q) de um modelo ARMA pode ser determinado minimizando $BIC(p, q) = T \ln \hat{\sigma}_{p, q}^2 + (p + q) \ln T$ onde $\hat{\sigma}_{p, q}^2$ é a estimativa de máxima verossimilhança da variância dos resíduos ε . Na prática é necessário fixar (P, Q) ambos suficientemente grandes para incluir os verdadeiros valores (p_0, q_0) e calcular $BIC(p, q)$ para $p \leq P$ e $q \leq Q$. Isto representa estimar $(P + 1)(Q + 1)$ modelos, por máxima verossimilhança.

Hannan e Rissanen (1982) sugerem o seguinte procedimento fortemente consistente. Estimar $\sigma_{p, q}^2$ através de uma série de regressões de Y_t em $Y_{t-1}, \dots, Y_{t-p}, \hat{\varepsilon}_{t-1}, \dots, \hat{\varepsilon}_{t-q}$ onde $\hat{\varepsilon}_t$ é obtido ajustando inicialmente um modelo AR de ordem grande aos dados. Além disso para $p = q$ a seqüência de regressões é economicamente calculada recursivamente através de uma seqüência de autorregressões bivariadas. O procedimento é o seguinte

i) Supõe-se que o modelo verdadeiro seja $ARMA(p_0, q_0)$. Defina-se P tal que presumivelmente $\max(p_0, q_0) \leq P$.

ii) Obtenha $\hat{\varepsilon}_t$ de

$$\hat{\varepsilon}_t = Y_t - \hat{a}_{M1}Y_{t-1} - \dots - \hat{a}_{MM}Y_{t-M} \quad (112)$$

onde M é obtido minimizando $BIC(p, 0)$ e \hat{a}_j obtidos recursivamente pelas equações (38) de Durbin, isto é,

$$\hat{a}_{MM} = \frac{r_M - \sum_{j=1}^{M-1} \hat{a}_{M-1,j} r_{M-j}}{\hat{\sigma}_{M-1}^2},$$

$$\hat{a}_{M,j} = \hat{a}_{M-1,j} + \hat{a}_{MM} \hat{a}_{M-1,M-j},$$

$$\hat{\sigma}_M^2 = \{1 - \hat{a}_{MM}\} \hat{\sigma}_{M-1}^2.$$

iii) Estime $P+1$ modelos $ARMA(p, p)$, $p = 0, 1, \dots, P$ através de mínimos quadrados da regressão

$$Y_t = \phi_{1p}Y_{t-1} + \dots + \phi_{pp}Y_{t-p} + \theta \hat{\varepsilon}_{t-1} + \dots + \theta_p \hat{\varepsilon}_{t-p} \quad (114)$$

obtendo $\hat{\sigma}_{pp}^2$, a estimativa dos resíduos.

iv) Determine \hat{p} como o valor que minimize

$$BIC(p, p) = T \ln \hat{\sigma}_{pp}^2 + 2p \ln T.$$

v) Hannan e Rissanen (1982) mostram que \hat{p} é um estimador consistente de $\max(p_0, q_0)$. Examine então $2\hat{p}$ modelos $ARMA(\hat{p}, r)$ e $ARMA(r, \hat{p})$ para $r = 0, 1, \dots, \hat{p} - 1$ através de regressões análogas a (114). Escolha o modelo que minimize $BIC(p, q)$ entre esses modelos.

Observe que em vez de estimar $(P+1)(P+1) = P^2 + 2P + 1$ modelos por máxima verossimilhança (em geral usando otimização não linear), o procedimento estima $P+1+2\tilde{p}(\tilde{p} \leq P)$ modelos através de regressão linear de mínimos quadrados. Hannan e Rissanen (1982) propõem ainda um procedimento baseado em autorregressões bivariadas de $w_t = (Y_t, \hat{\epsilon}_t)$ que reduz o trabalho computacional do cálculo das regressões.

8.2 Modelagem ARMA(n, n - 1) de Pandit-Wu

Essa metodologia é baseada no fato de que, se observarmos em instantes discretos um modelo AR(p) contínuo no tempo, o processo observado é um modelo ARMA(p, p - 1). Pandit (1973) e Pandit e Wu (1983) através do estudo da dinâmica de sistemas físicos sugerem o seguinte procedimento para modelagem de séries temporais:

- i) Como toda série estacionária pode ser representada por um AR(∞), estime modelos ARMA(2n, 2n - 1), $n = 1, 2, \dots$ de ordem crescente por mínimos quadrados condicionados ou máxima verossimilhança.
- ii) A cada estimação de um novo modelo calcula-se

$$F = \frac{A_1 - A_0}{4} \div \frac{A_0}{T - 4n + 4} \sim F(4, T - 4n + 4) \quad (115)$$

onde T - número de observações

A_1 - soma dos quadrados dos resíduos do modelo ARMA(2n, 2n - 1)

A_0 - soma dos quadrados dos resíduos do modelo ARMA(2n + 2, 2n + 1).

Se o valor de F excede o valor tabelado ao nível de significância $\alpha\%$ então, há uma diminuição significativa na soma dos quadrados dos resíduos ao passar-se do modelo $ARMA(2n, 2n - 1)$ para o modelo $ARMA(2n + 2, 2n + 1)$. Caso contrário, isto é, F menor que o valor tabelado, conclui-se que o modelo $ARMA(2n, 2n - 1)$ é adequado.

iii) Ao decidir parar no modelo $ARMA(2n, 2n - 1)$ pode-se usar o critério F para testar outros modelos, não testados até esta etapa, como por exemplo $ARMA(2n - 1, 2n - 2)$, $ARMA(2n - 1, m)$ $m < 2n - 1$, $AR(m)$, etc.

Observe que o critério F anterior é análogo ao do caso de um problema de regressão com r parâmetros em que se testa se δ dos parâmetros é zero através de

$$F = \frac{A_1 - A_0}{\delta} \div \frac{A_0}{N - r} \sim F(\delta, T - r).$$

Em (115) tem-se

$$r = (2n + 2) + (2n + 1) + 1 = 4n + 4$$

$$\delta = (4n + 4) - [4(n - 1) + 4].$$

8.3 Modelagem ARARMA de Parzen

Esta metodologia proposta por Parzen (1979, 1980 a, b, 1982) utiliza modelos ARARMA, assim denominados porque o modelo utilizado para a série temporal é baseado em uma análise sofisticada de modelos ARMA (modelo de curta memória) para os resíduos de um modelo extrapolativo (modelo de longa me

mória) obtido de uma autorregressão não estacionária de melhor lag.

Basicamente o procedimento é o seguinte

- i) Determina-se a memória da série estimando um modelo AR(M) como (112), onde a ordem M é determinada pelo critério CAT(p) ou AIC(p,0), (ver Fig. II5). Calculando-se os coeficientes através de (113) obtêm-se os resíduos $\hat{\epsilon}_t = \tilde{Y}_t$, isto é,

$$\tilde{Y}_t = Y_t - \hat{a}_{M1}Y_{t-1} - \dots - \hat{a}_{MM}Y_{t-M}$$

- ii) Calcula-se

$$\tilde{\sigma}_M^2 = \frac{\sum_{t=M+1}^T |\tilde{Y}_t|^2}{\sum_{t=M+1}^T |Y_t|^2} \quad (116)$$

e tem-se então:

$$\left\{ \begin{array}{ll} \text{Nenhuma memória (ruído branco)} & - \ell \tilde{\sigma}^2 + \frac{2M}{T} > -\frac{1}{T} \\ \text{Memória longa} & - \tilde{\sigma}^2 < \frac{8}{T} \\ \text{Memória curta} & - \text{outros casos} \end{array} \right. \quad (117)$$

- iii) Se a série é considerada ruído branco a análise é terminada. Se a série é considerada de memória curta, a mesma é modelada por um ARMA identificado por um dos processos da Seção 5.

- iv) Se a série é considerada de memória longa transforma a série através de sub modelos de

$$\tilde{Y}_t = Y_t - \phi_{j_1} Y_{t-j_1} - \dots - \phi_{j_M} Y_{t-j_M}$$

em geral utiliza-se

$$\tilde{Y}_t = Y_t - \phi_\ell Y_{t-\ell} \quad (118a)$$

$$\tilde{Y}_t = Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} \quad (118b)$$

$$\tilde{Y}_t = Y_t - \phi_1 Y_{t-(\ell-1)} - \phi_2 Y_{t-\ell} \quad (118c)$$

Por exemplo, se (118a) é usado, o lag ℓ é escolhido minimizando em ℓ

$$\text{Err}(\ell) = \frac{\sum_{t=\ell+1}^T (Y_t - \phi_\ell Y_{t-\ell})^2}{\sum_{t=\ell+1}^T Y_t^2} \quad (119)$$

e $\hat{\phi}_\ell$ é escolhido minimizando em ϕ

$$\sum_{t=\ell+1}^T \{Y_t - \phi Y_{t-\ell}\}^2 \quad (120)$$

O lag mais significativo é definido como aquele que minimiza $\text{Err}(\ell)$.

A análise é complementada por diversos plots de $\text{CAT}(p)$ contra p , $\hat{\sigma}_M^2$ contra M , ($\text{Err}(\ell)$ contra ℓ) densidade espectral e espectro acumulado. (ver Parzen 1980a).

8.4 Modelagem em Espaço de Estados de Akaike

Esta metodologia sugerida por Akaike (1974, 1976) baseia-se em resultados da teoria de controle e utiliza a representação de um modelo ARMA em espaço de estados (ver Akaike 1974 ou McGregor 1972).

Para o modelo ARMA(p,q)

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

de (1) (110) (111) temos

$$\hat{Y}_t = \phi^{-1}(L)\theta(L)\varepsilon_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \dots$$

$$\hat{Y}_t(p+j) = \phi_1 \hat{Y}_t(p+j-1) + \dots + \phi_p \hat{Y}_t(j) \quad (121)$$

$$\hat{Y}_{t+1}(j) = \hat{Y}_t(j+1) + \psi_j \varepsilon_{t+1} \quad j = 0, \dots$$

Mudando a notação da previsão para $\hat{Y}_t(j) = \hat{Y}_{t+j/t}$ tem-se ($k = \max(p, q+1)$).

$$\hat{Y}_{t+j+1/t+1} = \hat{Y}_{t+j+1/t} + \psi_j \varepsilon_{t+1} \quad (122)$$

e

$$\hat{Y}_{t+k/t+1} = \phi_1 \hat{Y}_{t+k-1/t} + \dots + \phi_k \hat{Y}_t/t + \psi_{k-1} \varepsilon_{t+1} \quad (123)$$

e portanto

$$Z_{t+1} = \begin{bmatrix} Y_{t+1/t+1} \\ Y_{t+2/t+1} \\ \vdots \\ Y_{t+k/t+1} \end{bmatrix} = \begin{bmatrix} 0 & & & \\ 0 & & & \\ & & I_{k-1} & \\ \phi_k & \phi_{k-1} & \dots & \phi_1 \end{bmatrix} \begin{bmatrix} \hat{Y}_t/t \\ \hat{Y}_{t+1/t} \\ \vdots \\ Y_{t+k-1/t} \end{bmatrix} + \begin{bmatrix} 1 \\ \psi_1 \\ \vdots \\ \psi_{k-1} \end{bmatrix} \varepsilon_{t+1} \quad (124)$$

onde $Y_t = (1 \ 0 \ \dots \ 0) Z_t$ (pois $\hat{Y}_t/t = Y_t$) e $\phi_j = 0 \ j > p$.

A representação (124), isto é,

$$Z_{t+1} = FZ_t + G\varepsilon_t \quad (125)$$

$$Y_t = HZ_t$$

e chamada representação em espaço de estados ou representação Markoviana do modelo ARMA.

Dada a representação em espaço de estados, Akaike (1974) mostrou que a dimensão k de Z_t de mínima representação é idêntica ao posto da matriz de Hankel

$$H_L = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \gamma_L \\ \gamma_1 & \gamma_2 & \gamma_3 & \gamma_{L+1} \\ \gamma_2 & \gamma_3 & \gamma_4 & \gamma_{L+2} \\ \dots & \dots & \dots & \dots \\ \gamma_L & \gamma_{L+1} & \gamma_{L+2} & \gamma_{2L} \end{bmatrix} \quad (126)$$

para $L(>k)$ suficientemente grande, onde $\gamma_j = \text{Cov}(Y_{t+j}, Y_t)$. Observe que H_L corresponde a matriz de covariância entre $Y_t^+ = (Y_t, Y_{t+1}, \dots, Y_{t+L-1})$ e $Y_t^- = (Y_t, Y_{t-1}, \dots, Y_{t-L+1})$ e o número de correlações canônicas não nulas entre Y_t^+ e Y_t^- é igual a dimensão de Z_t .

A título de recordação, correlação canônica entre dois conjuntos de variáveis $\{X_1, \dots, X_m\}$ e $\{Y_1, \dots, Y_n\}$ envolve a determinação de coeficientes $\{a_{1j}, \dots, a_{1m}\}$ e $\{b_{1j}, \dots, b_{1n}\}$ tais que a correlação r_1 entre as combinações lineares

$$X_1^* = \sum_{j=1}^m a_{1j} X_j, \quad Y_1^* = \sum_{j=1}^n b_{1j} Y_j$$

seja maximizada. A seguir determina-se

$$X_2^* = \sum_{j=1}^m a_{2j} X_j, \quad Y_2^* = \sum_{j=1}^n b_{2j} Y_j$$

tais que X_2^* e Y_2^* sejam não correlacionadas com X_1^* e Y_1^* e que $r = \text{Corr}(X_2^*, Y_2^*) < r_1$ seja maximizada. O processo con

tinua até que $\min(m,n)$ correlações canônicas $\{r_1, r_2, \dots\}$ e os correspondentes pares de variáveis canônicas $\{X_1^*, Y_1^*\}, \dots, \{X_2^*, Y_2^*\}, \dots$ tenham sido extraídos.

Baseado nestas considerações tem-se o seguinte procedimento:

- i) Determine as autocovariâncias amostrais $\hat{\gamma}_j$.
- ii) Ajuste um modelo AR(M) aos dados onde M é escolhido pelo critério AIC:
- iii) Defina o vetor $Y^- = (Y_t, Y_{t-1}, \dots, Y_{t-M})'$ e o vetor $Y_M^+ = (Y_t, Y_{t+1}, \dots, Y_{t+M})'$.
- iv) O vetor Z de dimensão $k+1$ é definido por $k+1$ componentes de Y_M^+ denotadas por Z^i . Supõe-se que as correlações canônicas entre $Z^- = (Z^1, \dots, Z^k)$ e Y^- são positivas e as primeiras k componentes do vetor de estados Z_t são definidas por $Z_{it} = Z^i$, $i = 1, 2, \dots, k$.
- v) Inicie o procedimento com $Z = (Y_t, Y_{t+1})'$.
- vi) Calcule a correlação canônica entre Y^- e Z.
- vii) Compute

$$DIC(\ell) = -T \ln \prod_{i=\ell+1}^{k+1} (1 - c_i^2) - 2(k+1-\ell)(M+1-\ell) \quad (127)$$

onde

c_i = i-esima maior correlação canônica

$k+1$ = dimensão do vetor Z

ℓ = número de correlações canônicas não nulas.

iii) Se $DIC(\ell - 1)$ é positivo, a última componente de Z é considerada linearmente independente e aceita como a $(k+1)$ -ésima componente $Z_{(k+1),t}$ do vetor de estado Z_t . Aumente o valor de $k+1$ de uma unidade e volte a etapa iv).

ix) Se $DIC(\ell - 1)$ é negativo a menor correlação canônica entre Y^- e Z é considerada nula, neste caso a última componente de Z é considerada linearmente dependente nas antecedentes na seqüência de preditores.

Supondo que a $(k+1)$ -ésima variável canônica é definida por

$$b_1 Z^1 + b_2 Z^2 + \dots + b_{k+1} Z^{k+1} = 0$$

onde os b 's são os pesos canônicos da $(k+1)$ -ésima variável canônica julgada com correlação canônica zero. Os b 's não podem ser zero e então

$$Z^{k+1} = -\frac{b_1}{b_{k+1}} Z^1 - \dots - \frac{b_k}{b_{k+1}} Z^k. \quad (129)$$

A análise da variável canônica fornece k , a dimensão do vetor de estados e estimativas dos parâmetros autorregressivos

$$\hat{\phi}_i = -\frac{b_i}{b_{k+1}} \quad i = 1, \dots, k. \quad (130)$$

x) Estimativas para os pesos ψ_i são obtidas através do mo

delo AR(M) ajustado na etapa ii), isto é,

$$\begin{aligned} Y_t &= \alpha_1 Y_{t-1} + \dots + \alpha_M Y_{t-M} + u_t \\ &= (1 - \alpha_1 L - \dots - \alpha_M L^M)^{-1} u_t \end{aligned}$$

$$\psi_0 = 1$$

$$\psi_1 = \alpha_1$$

$$\psi_2 = \alpha_1 \psi_1 + \alpha_2$$

$$\vdots$$

$$\psi_M = \alpha_1 \psi_{M-1} + \alpha_2 \psi_{M-2} + \dots + \alpha_{M-1} \psi_1 + \alpha_M$$

Se necessário, estime o modelo ARMA(k,k-1) por máxima verossimilhança ou através de procedimentos no domínio da frequência (Akaike 1976, Cooper e Wood 1982a, 1982b, 1982c). Mais geralmente, escolha alguns valores (i, j) próximos de (k, k-1) estimando diversos modelos ARMA(i,j) utilizando AIC para escolher o mais adequado (Priestley, 1980).

Finalmente, dado o modelo em espaço de estados, previsões e suas variâncias podem ser calculadas como na Seção 7 ou através do Filtro de Kalman como descrito em Hui e O'Reilly (1982).

8.5 Decomposição por Componentes Principais

Ao final da Seção 3.2, mencionamos que muitas aplicações requerem a obtenção da tendência e/ou sazonalidade explicitamente. Diversos métodos têm sido propostos com esta finalidade.

lidade e considerados em Kenny e Durbin (1982). Nesta Seção apresenta-se um método utilizado por Basilevsky e Hum (1979) baseado em procedimento da análise estatística multivariada clássica e conhecido como decomposição de Karhunen-Loeve.

O procedimento a ser descrito embora justificável teoricamente é bastante empírico e pode ser considerado como um método de análise exploratória dos dados. Para indicar os passos considere-se uma série de observações mensais Y_t com componentes de tendência e de sazonalidade. As etapas são:

- i) Se o período de sazonalidade é s (por exemplo 12) defina M tentativamente como um múltiplo de s (por exemplo $M = 24$ ou 36).
- ii) Sendo o número de observações $T = n + M$ definindo a série $Y(t)$ como

$$Y(t) = (Y_{-M+1}, Y_{-M+2}, \dots, Y_0, Y_1, \dots, Y_n), \quad (132)$$

construa $M + 1$ vetores defasados de comprimento $n > M$.

$$Y(t - j) = (Y_{1-j} \dots Y_{n-j}), \quad j = 0, 1, \dots, M. \quad (133)$$

- iii) Considere-se o problema como n observações de $M + 1$ variáveis. Efetua-se então uma análise de componentes principais diagonalizando a matriz de covariância (ou de correlação) destas variáveis. Isto é, transforma-se estas $M + 1$ variáveis correlacionadas em novas $M + 1$ variáveis Z_t não correlacionadas com variâncias λ_i , $i = 1, \dots, M$ tais que

$$\lambda_1 > \lambda_2 > \dots > \lambda_{M+1}$$

- iv) Plote as variáveis Z_t para verificar seu comportamento, tentando interpretar algum deles ou suas combinações como tendência, sazonalidade, ciclo, etc.

Na prática em geral não será necessário plotar todos os Z_t mas somente os principais que expliquem a maior parte da variância. Basilevsky e Hum (1979) encontraram em sua análise seis primeiros componentes explicando 66,9% da variação e interpretaram a agregação de Z_{1t} e Z_{2t} como sazonalidade, Z_{3t} como tendência e a agregação de Z_{4t} , Z_{5t} e Z_{6t} como ciclo. No caso escolheu-se $M = 32$ por conveniência computacional.

Finalmente, para fazer previsões pode-se adicionar a seguinte etapa:

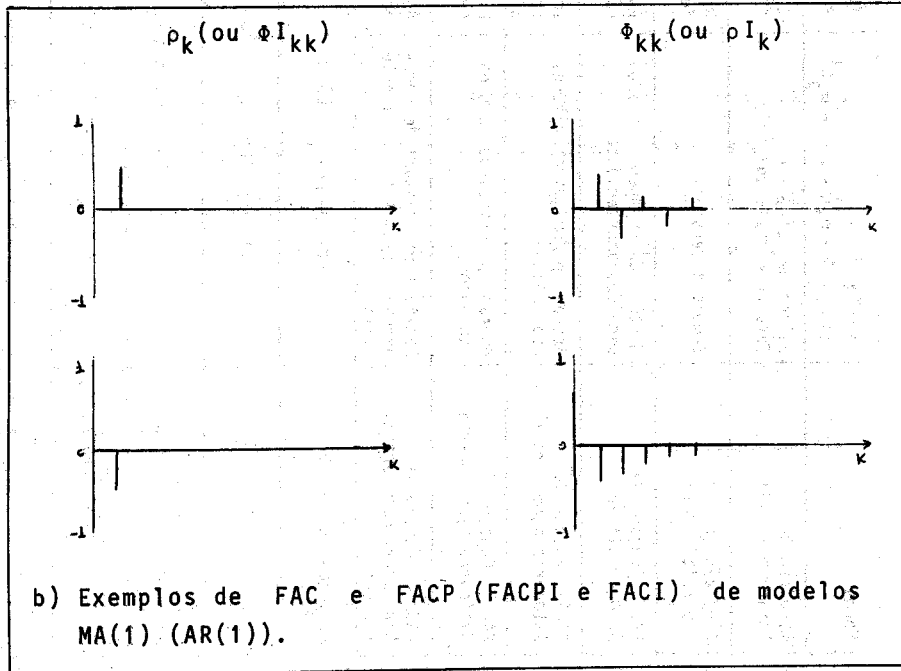
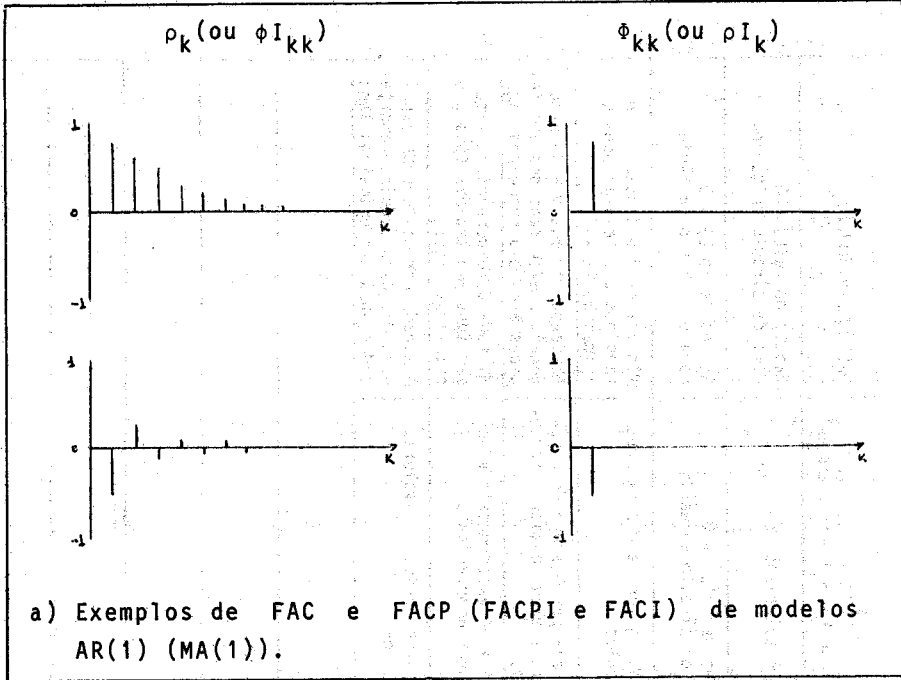
- vi) Como os Z são não correlacionados faz-se previsões dos mesmos utilizando a metodologia de Box-Jenkins. Essas previsões fornecem as previsões da tendência, sazonalidade, etc que combinadas dão a previsão para Y .

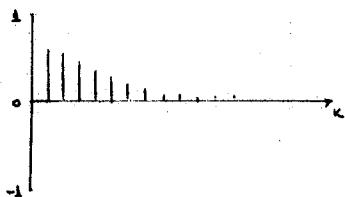
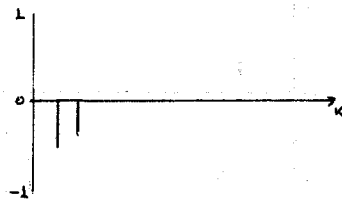
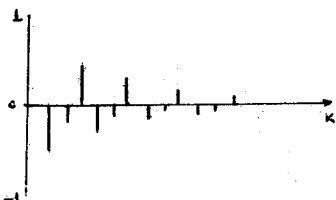
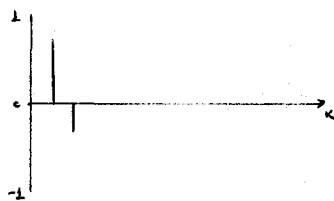
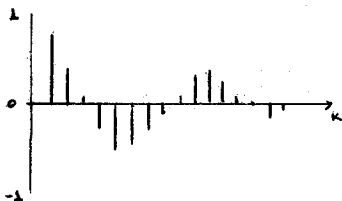
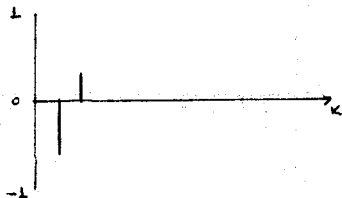
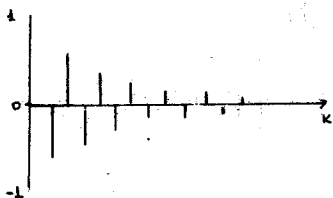
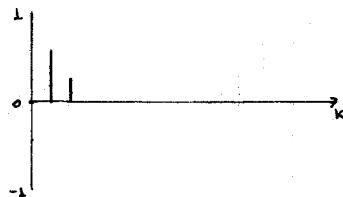
APÊNDICES

A1 - Resumo de Identificação

Modelo	AR(p)	MA(q)	ARMA(p, q)
Forma	$\phi(L)Y_t = \epsilon_t$ $\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$	$Y_t = \theta(L)\epsilon_t$ $\theta(L) = 1 + \theta_1 L + \dots + \theta_q L^q$	$\phi(L)Y_t = \theta(L)\epsilon_t$
Condição de Estacionariedade	Raízes de $\phi(L) = 0$ fora do círculo unitário.	Sempre estacionário.	Raízes de $\phi(L) = 0$ fora do círculo unitário.
Condição de Inversibilidade	Sempre inversível	Raízes de $\theta(L) = 0$ fora do círculo unitário	Raízes de $\theta(L) = 0$ fora do círculo unitário.
FAC - Função de Autocorrelação	Mistura de exponenciais e/ou senóides amortecidas.	Apenas q primeiros valores não nulos.	Mistura de exponenciais e/ou senóides amortecidas após q-p primeiros lags.
FACP - Função de Autocorrelação Parcial	Apenas p primeiros valores não nulos.	Dominada por mistura de exponenciais e/ou senóides amortecidas.	Dominada por mistura de exponenciais e/ou senóides amortecidas após p-q primeiros lags.
FACI - Função de Autocorrelação Inversa.	Apenas p primeiros valores não nulos.	Mistura de exponenciais e/ou senóides amortecidas.	Mistura de exponenciais e/ou senóides amortecidas após p-q primeiros lags.
FACPI - Função de Autocorrelação Parcial Inversa	Dominada por mistura de exponenciais e/ou senóides amortecidas.	Apenas q primeiros valores não nulos.	Dominada por mistura de exponenciais e/ou senóides amortecidas após q-p primeiros lags.
Arranjos R* e S*	ver páginas 80 a 83.		
Método Corner	ver páginas 83 a 85.		
FACPG - Função de Autocorrelação Parcial Generalizada.	ver páginas 85 a 87.		
FACE - Função de Autocorrelação Estendida	ver páginas 87 a 93.		
CrITÉrios Automáticos	ver páginas 93 a 95.		

A2 - Correlogramas

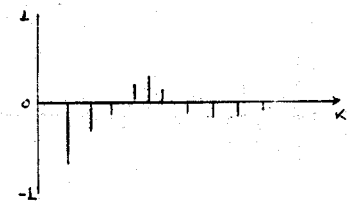
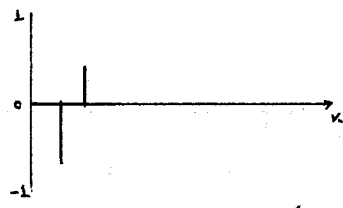
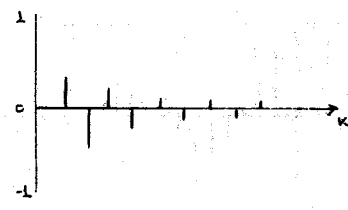
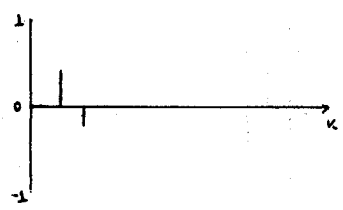
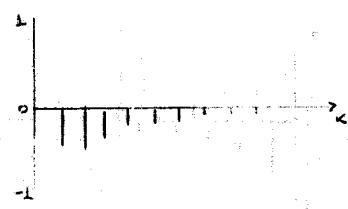
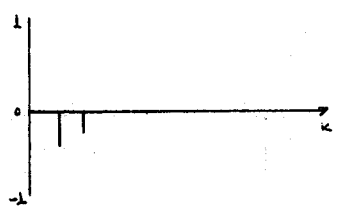
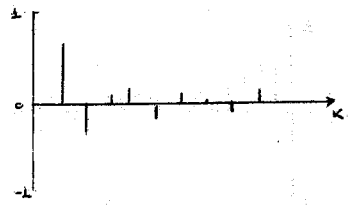
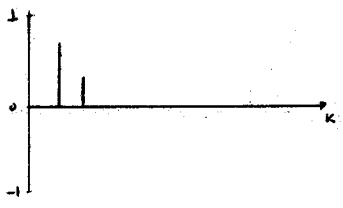


ρ_k (ou Φ_{kk})

 Φ_{kk} (ou ρI_k)


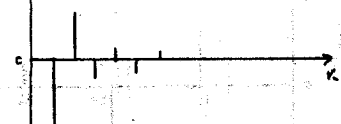
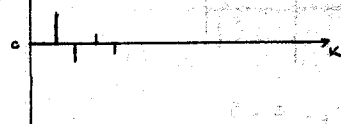
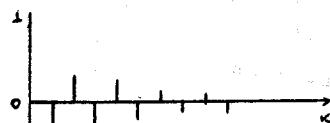
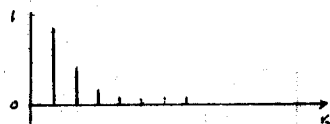
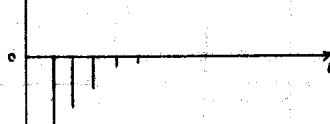
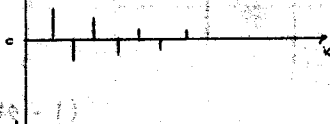
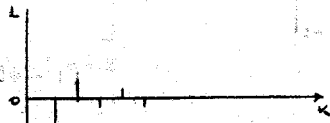
c) Exemplos de FAC e FACP (FACPI e FACI) de modelos AR(2) (MA(2)).

ρ_k (ou ϕI_{kk})

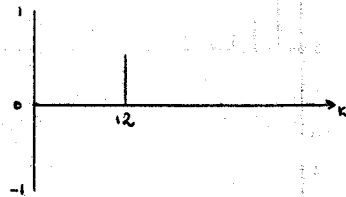
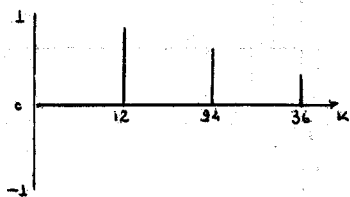
ϕ_{kk} (ou ρI_k)



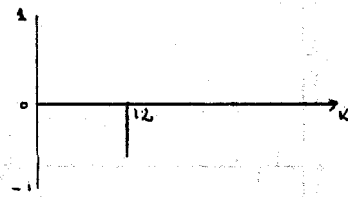
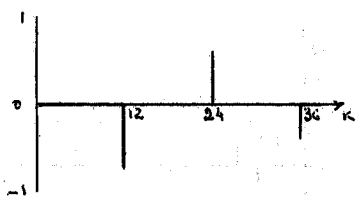
d) Exemplos de FAC e FACP (FACPI e FACI) de modelos MA(2) (AR(2)).

ρ_k (ou $\phi_{I_{kk}}$) ϕ_{kk} (ou ρ_{I_k})

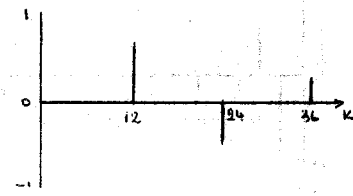
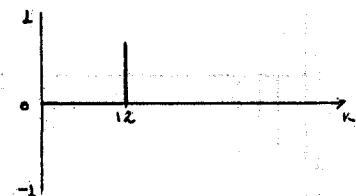
e) Exemplos de FAC e FACP (FACPI e FACI) de modelos ARMA(1,1) (ARMA(1,1)).

ρ_k (ou ϕI_{kk}) ϕ_{kk} (ou ρI_k)

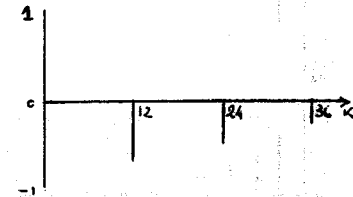
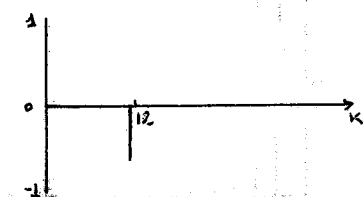
$$(1 - \phi B^{12}) Y_t = \varepsilon_t, \quad \phi > 0$$



$$(1 - \phi B^{12}) Y_t = \varepsilon_t, \quad \phi < 0$$

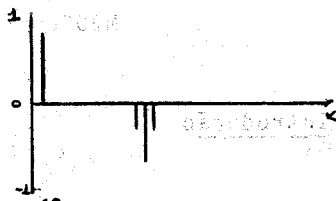
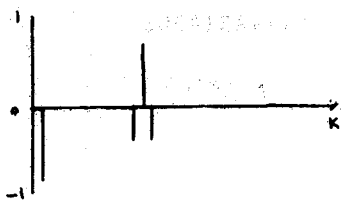


$$Y_t = (1 - \theta B^{12}) \varepsilon_t, \quad \theta > 0$$

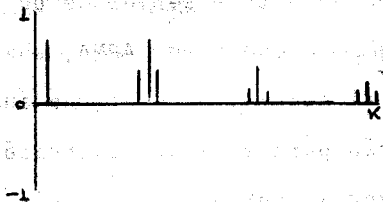
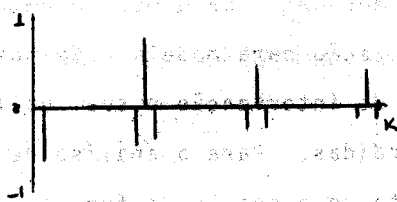


$$Y_t = (1 - \theta B^{12}) \varepsilon_t, \quad \theta < 0$$

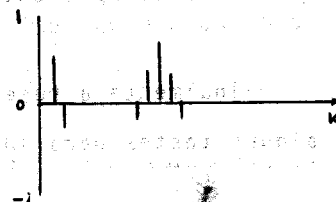
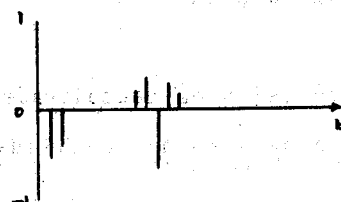
f) Exemplos de FAC e FACP (FACPI e FACI) de modelos sazonais puros.

ρ_k (ou ϕI_{kk})


$$Y_t = (1 - \theta B)(1 - \theta B^{12}) \varepsilon_t$$



$$(1 - \phi B^{12}) Y_t = (1 - \theta B) \varepsilon_t$$



$$Y_t = (1 - \theta_1 B - \theta_2 B^2)(1 - \theta B^{12}) \varepsilon_t$$

g) Exemplos de FAC (FACPI) de modelos sazonais múltiplos.
cativos.

CAPÍTULO III

MODELOS LINEARES MULTIVARIADOS

1. Introdução

Este capítulo apresenta alguns modelos para análise de séries temporais quando dispõe-se de informações adicionais ou de outras séries temporais. Apresenta-se o modelo de regressão com erro ARMA, sua utilização para modelar efeitos de calendário, o modelo de análise de intervenção e sua utilização para estimar observações perdidas. Para a análise de diversas séries temporais apresenta-se o modelo de função de transferência para sistemas sem "feedback" e o modelo vetorial ARMA para sistemas com "feedback".

Apresenta-se a metodologia de Box-Jenkins-Tiao para modelagem de modelos vetoriais ARMA, bem como as metodologias alternativas. Esses resultados são extensões imediatas dos casos univariados apresentados no capítulo anterior.

Finalmente apresenta-se o conceito de causalidade bem como alguns testes para identificar a direção da causalidade.

2. Regressão em Séries Temporais

2.1 Regressão com Erros ARMA

Considere-se a equação de regressão

$$Y = f(X, \beta) + u_t = \beta_0 + \beta_1 x_{1t} + \dots + \beta_L x_{Lt} + u_t \quad (1)$$

onde os x 's podem ser valores defasados das variáveis exógenas e considerados conhecidos. O processo u_t é um processo linear geral que pode ser representado parcimoniosamente por um modelo ARMA(p,q) (ou SARIMA no caso mais geral), isto é,

$$\phi(L)u_t = \theta(L)\varepsilon_t \quad (2)$$

Suponha-se que u_t é estacionário e inversível.

A construção de modelos da forma (1) e sua utilização para previsão pode ser feita utilizando as seguintes etapas:

i) Obtenha os resíduos u_t ,

$$\hat{u}_t = Y_t - \tilde{\beta}_0 - \tilde{\beta}_1 x_{1t} - \dots - \tilde{\beta}_L x_{Lt} \quad (3)$$

onde os $\tilde{\beta}$'s são estimativas de mínimos quadrados ordinários dos β 's.

ii) Construa o modelo ARMA para \hat{u}_t utilizando os resultados do capítulo anterior.

iii) As estimativas $\tilde{\beta}$'s embora não viciadas, são não eficientes e portanto o mesmo ocorre via \hat{u}_t para as estimativas dos ϕ 's e θ 's de (2). Por esta razão os parâ

metros (β, ϕ, θ) devem ser estimados.

Dois procedimentos podem ser utilizados

- a) Com as estimativas $\hat{\phi}$ e $\hat{\theta}$ obtidas em (ii) forme a matriz de covariância de \hat{u}_t , i.e.

$$\hat{\Sigma} = \{\hat{\gamma}_{ij}\} = \{\gamma_{i-j}\}, \quad \gamma_k = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k} \quad (4)$$

(Por exemplo, se $\hat{u}_t = \phi \hat{u}_{t-1} + \varepsilon_t$, $\gamma_k = \phi^k \sigma_\varepsilon^2$).

Obtenha a estimativa de mínimos quadrados generalizados de β

$$\tilde{\beta} = (X' \hat{\Sigma}^{-1} X)^{-1} X' \hat{\Sigma}^{-1} Y \quad (5)$$

volte às etapas (i) e (ii) com $\tilde{\beta}$.

A matriz de covariância dos estimadores é

$$V(\tilde{\beta}) = \hat{\sigma}_\varepsilon^2 (X' \hat{\Sigma}^{-1} X)^{-1}. \quad (6)$$

Este procedimento é equivalente ao método de Cochrane-Orcutt generalizado. (Pierce, 1980).

- b) Através de máxima verossimilhança (de forma análoga a Seção 6 do capítulo anterior). Em particular uma aproximação pode ser obtida a partir do modelo

$$Y_t = \beta_0 + \beta_1 x_{1t} + \dots + \beta_L x_{Lt} + \frac{\theta(L)}{\phi(L)} \varepsilon_t \quad (7)$$

minimizando as somas dos quadrados

$$S(\beta, \phi, \theta) = \sum \varepsilon_t^2 \quad (8)$$

ou a soma dos quadrados condicionados

$$S_c(\beta, \phi, \theta) = \Sigma \epsilon_t^2. \quad (9)$$

A matriz de covariâncias dos estimadores é

$$I^{-1}(\hat{\beta}, \hat{\phi}, \hat{\theta}) \approx 2\sigma_\epsilon^2 \left(\frac{\partial^2 S(\hat{\eta})}{\partial \eta_i \partial \eta_j} \right) \quad (10)$$

onde $\hat{\eta} = (\hat{\beta}, \hat{\phi}, \hat{\theta})$.

Por outro lado, se a função de verossimilhança exata for necessária, Harvey e Phillips (1979) fornecem um algoritmo para obtê-la usando Filtro de Kalman.

iv) A previsão de valores futuros de Y envolve a previsão de valores de u_t e da parte estrutural, isto é,

$$Y_t(\ell) = \sum_{j=1}^{\ell} \hat{\beta}_j \hat{x}_{jt}(\ell) + \hat{u}_t(\ell). \quad (11)$$

onde $\hat{u}_t(\ell)$ é obtida como na Seção 7 do capítulo anterior.

Outros métodos de estimação e aplicações podem ser vistos em Bhattacharyya (1974), Pukkila (1978), Harvey e Phillips (1979) e Pierce (1980).

2.2 Efeitos de Calendário

Em séries mensais muitas vezes o nível da série é afetado por variações do calendário devido a mudanças do nível de atividade resultante de mudanças na construção do calendário. Por exemplo, o total de vendas do mês de janeiro é afetado pelo número de fins de semanas que este mês contém. Outro exemplo é o de vendas de bebidas alcoólicas dos meses de fevereiro e março que são afetados pela ocorrência ou não do carnaval naqueles meses. O primeiro exemplo é um caso de e-

feito de variação dos dias úteis" e o segundo é de "efeito de feriado". Esses efeitos têm sido considerados no contexto de ajustamento sazonal (Cleveland e Devlin, 1980 e Cleveland e Grupe, 1981). Recentemente Hillmer (1982) e Bell e Hillmer (1983) consideraram este problema na identificação e previsão com modelos ARMA:

Em particular considerando o modelo (1)

$$Y_t = f(X, \beta) + u_t$$

onde $f(X, \beta)$ representava os efeitos de calendário, foi analisado o efeito de não incluir $f(X, \beta)$ tanto na fase de identificação como na de previsão. Na fase de identificação verificou-se que os resíduos de um modelo ARMA para Y_t apresentava altas correlações em lags completamente díspares e que as previsões eram melhores ao se incluir $f(X, \beta)$ no modelo.

Para o efeito de variação de dias úteis o modelo utilizado é

$$Y_t = \sum_{j=1}^7 \beta_j x_{ij} + u_t \quad (12)$$

onde x_{it} é o número de segundas feiras ($i=1$), terças feiras ($i=2$) até domingos ($i=7$) para o mês t , β_i são parâmetros e u_t obedece um modelo ARIMA ou SARIMA. Por conveniência é útil a parametrização

$$Y_t = \mu + \sum_{i=1}^6 \beta_i (X_{it} - X_{i7}) + \frac{\theta(L)}{\phi(L)\delta(L)} \epsilon_t \quad (13)$$

onde $\delta(L)$ é um polinômio de diferenças para obter estacionariedade. Por exemplo

$$\delta(L) = (1 - L)(1 - L^{12})$$

neste caso o modelo (11) é reescrito como

$$(1 - L)(1 - L^{12})Y_t = \mu + \sum_{i=1}^6 (1 - L)(1 - L^{12})(X_{it} - X_{i7}) + \frac{\theta(L)}{\phi(L)} \varepsilon_t \quad (14)$$

A equação (14) está na forma de regressão com erro ARMA e o procedimento da Seção 2.1 é então utilizado.

Dadas as estimativas de $\hat{\beta}$ e sua matriz de covariâncias $V(\hat{\beta})$, sabe-se que para um número grande de observações:

$$\hat{\beta} \sim N(0, V(\hat{\beta})) \quad (15)$$

$$\hat{\beta}'[V(\hat{\beta})]^{-1}\hat{\beta} \sim \chi_6^2 \quad (16)$$

Para testar a hipótese de que não há efeito de variação de dias úteis, isto é,

$$H_0 : \beta_1 = \dots = \beta_6 = 0$$

contra $H_1 =$ nem todos $\beta_i = 0$

tem-se que, se $\hat{\beta}'[V(\hat{\beta})]^{-1}\hat{\beta} > \chi(\alpha)$, rejeita-se H_0 , onde $\chi(\alpha)$ é o valor tabelado da χ_6^2 com $\alpha\%$ de significância.

3. Função de Transferência

3.1 Introdução

Nesta seção relaciona-se a variável Y_t não só com seu passado, mas também com o presente e passado de outras variáveis. A influência de uma variável em outra pode-se distribuir ao longo de diversos períodos. Efeitos instantâneos e defasados de uma variável input (exôgena) X_t em uma variável output (endôgena) Y_t podem ser representados por um modelo da forma

$$Y_t = v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \dots = v(L) X_t \quad (17)$$

onde $v(L) = v_0 + v_1 L + v_2 L^2 + \dots$ é chamada de função de transferência e define a dinâmica do sistema; os coeficientes v_0, v_1, v_2, \dots são chamados pesos de resposta a impulso. Teoricamente este modelo envolve um número infinito de coeficientes o que torna impossível sua estimação. Uma alternativa é usar uma forma truncada

$$Y_t = v_0 X_t + v_1 X_{t-1} + \dots + v_h X_{t-h} \quad (18)$$

onde h é escolhido de tal forma que os efeitos defasados além de h são negligíveis. Outra alternativa é escrever $v(L)$ como uma função racional

$$v(L) = \frac{w(L)L^b}{\delta(L)} \quad (19)$$

Os operadores $w(L) = w_0 - w_1 L - \dots - w_s L^s$ e $\delta(L) = 1 - \delta_1 L - \dots - \delta_r L^r$ são polinômios em L e b é um parâmetro representando o retardo entre as variáveis. Supõe-se que as raízes de $w(L) = 0$

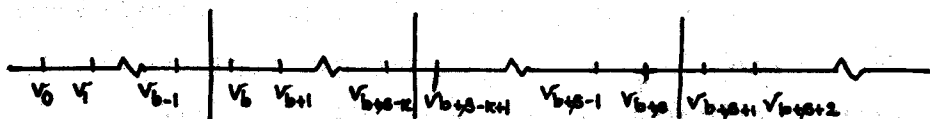
e $\delta(L) = 0$ estão sobre ou fora do círculo unitário. A relação entre os coeficientes v_k e $w = (w_0, \dots, w_s)'$ e $\delta = (\delta_1, \dots, \delta_r)'$ e b pode ser obtida igualando os coeficientes de L^j em

$$(1 - \delta_1 L - \dots - \delta_r L^r)(v_0 + v_1 L + \dots) = w_0 - w_1 L - \dots - w_s L^s \quad (20)$$

resultando em

$$\begin{aligned} v_j &= 0 & j < b \\ v_j &= \delta_1 v_{j-1} + \delta_2 v_{j-2} + \dots + \delta_r v_{j-r} + w_0 & j = b \\ v_j &= \delta_1 v_{j-1} + \delta_2 v_{j-2} + \dots + \delta_r v_{j-r} - w_{j-b} & j = b+1, \dots, b+s \\ v_j &= \delta_1 v_{j-1} + \delta_2 v_{j-2} + \dots + \delta_r v_{j-r} & j > b+s \end{aligned} \quad (21)$$

ou esquematicamente



1. b pesos: 0

2. $s-k+1$ pesos sem padrão

3. π valores iniciais p/ equação de diferenças

4. valores seguindo padrão ditado por equação de diferenças de grau π .

- Grupos de respostas ao impulso v_j .
(Grupo 2 só existe se $r \leq s$. Se $r > s+1$ Grupos 1 e 3 se misturam).

Box e Jenkins (1976, pg. 349 a 351), Heuts e Willems (1981) e Box e Tiao (1975) apresentam respostas de impulsos para alguns modelos.

Até o momento supôs-se que Y_t era completamente de

terminada por X_t e seu passado. Usualmente tal relação é acompanhada de um ruído aditivo N_t que pode ser autocorrelacionado e segue um modelo ARMA(p,q). Portanto o modelo geral a ser considerado é

$$Y_t = \frac{\omega(L)L^b}{\delta(L)} X_t + N_t = \frac{(\omega_0 - \omega_1 L - \dots - \omega_r L^r)}{(1 - \delta_1 L - \dots - \delta_r L^r)} X_{t-b} + \frac{\theta(L)}{\phi(L)} \varepsilon_t \quad (22)$$

Neste modelo supõe-se que o passado de X influencia futuros Y , mas não vice versa. Tais modelos são chamados de modelos de função de transferência com ruído ARMA. Eles são modelos de defasagens distribuídas e são generalizações dos modelos de regressão pois admitem: a) erros correlacionados; b) relações dinâmicas entre variáveis endôgenas e exôgenas; c) variáveis exôgenas estocásticas.

A construção de modelos de função de transferência é feita através do procedimento iterativo: identificação, estimação, verificação, antes do modelo ser utilizado.

3.2 Função de Correlação Cruzada

Diz-se que o processo estocástico bivariado (X_t, Y_t) é estacionário em sentido amplo (ou fracamente ou 2ª ordem) se

- i) X_t e Y_t são estacionários em sentido amplo, isto é,
 $E(X_t) = \mu_X$, $E(Y_t) = \mu_Y$, μ_X , μ_Y não dependem de t ,
 $Cov(X_t X_{t+k}) = \gamma_X(k)$, $Cov(Y_t Y_{t+k}) = \gamma_Y(k)$, são dependentes de k .

- ii) A função de covariância cruzada

$\text{Cov}(X_t, Y_{t+k}) = \gamma_{XY}(k)$ sã depende do lag k

Observando que

$$\gamma_{XY}(k) = E(X_t - \mu_X)(Y_{t+k} - \mu_Y) = E(Y_{t+k} - \mu_Y)(X_t - \mu_X) = \gamma_{YX}(-k) \quad (23)$$

ta considerar $\gamma_{XY}(k)$ $k = 0, \pm 1, \pm 2, \dots$. Alã disto, este resultado mostra que $\gamma_{XY}(k)$ nã ã uma funãõ par (simãtrica).

Define-se a funãõ de correlaãõ cruzada como

$$\rho_{XY}(k) = \frac{\gamma_{XY}(k)}{\{\gamma_X(0)\gamma_Y(0)\}^{1/2}} \quad k = 0, \pm 1, \pm 2, \dots \quad (24)$$

Se T observaãões de uma sãrie temporal sua estimativa ã dada por

$$\hat{r}_{XY} = \frac{\hat{\gamma}_{XY}(k)}{\{\hat{\gamma}_X(0)\hat{\gamma}_Y(0)\}^{1/2}} \quad k = 0, \pm 1, \pm 2, \dots \quad (25)$$

$$\hat{\gamma}_{XY} = \begin{cases} \frac{1}{T} \sum_{t=1}^{T-k} (X_t - \bar{X})(Y_{t+k} - \bar{Y}) & k = 0, 1, 2, \dots \\ \frac{1}{T} \sum_{t=1-k}^T (X_t - \bar{X})(Y_{t+k} - \bar{Y}) & k = 0, -1, -2, \dots \end{cases} \quad (26)$$

$$\hat{\gamma}_X(0) = \hat{\sigma}_X^2 = \frac{1}{T} \sum_{t=1}^T (X_t - \bar{X})^2 \quad \hat{\gamma}_Y(0) = \hat{\sigma}_Y^2 = \frac{1}{T} \sum_{t=1}^T (Y_t - \bar{Y})^2$$

\bar{X} e \bar{Y} sãõ as mãdias amostrais das sãries respectivas.

Resultados sobre variãncias e covariãncia dos estimadores da correlaãõ cruzada sãõ apresentados em Bartlett (1955), Ghosh (1976) e McLeod (1979). Em particular tem-se:

Se ε_t ã ruãido branco e Y_t ã autocorrelacionado com au

to correlação $\rho_Y(k)$ e ϵ_t, Y_t são não correlacionadas para todos os lag ($\rho_{\epsilon Y}(k) = 0 \forall k$)

$$V\{r_{\epsilon Y}(k)\} \approx \frac{1}{n-k} \quad (27)$$

$$\rho(r_{\epsilon Y}(k), r_{\epsilon Y}(k+l)) \approx \rho_Y(l)$$

Resultados como o anterior mostram que não seria conveniente tentar especificar a função de transferência $v(L)$ em (17) ou (22) a partir das correlações cruzadas amostrais $r_{XY}(k)$, pois sua distribuição dependerá das funções de autocorrelações $\rho_X(l); \rho_Y(l)$, além de $\rho_{XY}(l)$. O que se deseja é identificar a dependência entre as séries, e a dependência dentro de cada série obscurece esta identificação. O que se faz em geral é pré-branquear (transformar em ruído branco) as séries X_t e Y_t antes de calcular a função de correlação cruzada. Estes procedimentos procuram evitar a obtenção de correlações espúrias (ver Granger e Newbold, 1974 e Box e Newbold, 1971, Hipel, Li e McLeod, 1983).

3.3 Identificação Empírica

Este enfoque descrito em Hipel, Li e McLeod (1983) consiste nas seguintes etapas

i) Baseado no conhecimento do fenômeno que gera a série temporal, bem como nas propriedades do modelo de função de transferência, identifique a função de transferência $v(L)$ em (22). Por exemplo, analisando-se séries mensais onde o input é uma série mensal de precipitações e o output é a vazão

sal de um rio, e se, pelas características físicas da bacia, sabe-se que as chuvas deste mês sã afetam a vazão deste mês e do próximo mês, uma função de transferência apropriada seria

$$v(L) = \omega_0 - \omega_1 L = v_0 + v_1 L. \quad (28)$$

ii) Apõs decidir sobre a forma de $v(L)$, identifique o modelo para o ruído N_t em (22). Para isto utilize o procedimento da Seção 2, isto é, estime $v(L)$ por mínimos quadrados (linear ou não linear dependendo de $v(L)$). Identifique o modelo ARMA para os resíduos e o modelo em (22) fica completamente especificado.

3.4 Identificação de Haugh-Box

Este enfoque descrito em Haugh e Box (1977) consiste nas seguintes etapas:

i) Apõs convenientes transformações (Box-Cox, diferenças sazonais etc.) suponha que a sãrie bivariada (X_t, Y_t) é estacionãria. Identifique os modelos ARMA

$$\phi_X(L)X_t = \theta_X(L)\alpha_t \quad (29)$$

$$\phi_Y(L)Y_t = \theta_Y(L)\eta_t$$

e estime os resíduos $\hat{\alpha}_t$ e $\hat{\eta}_t$. (30)

ii) Utilizando (25) calcule a função de correlação cruzada (FCC) entre $\hat{\alpha}_t$ e $\hat{\eta}_t$.

iii) Baseado nas características da FCC entre os resíduos e usando (21) identifique a função de transferência

$$\eta_t = v'(L)\alpha_t + N_t' = \frac{w'(L)}{\delta'(L)} \alpha_t + \frac{\theta'(L)}{\phi'(L)} \varepsilon_t \quad (31)$$

Haugh e Box (1977) mostraram que a FCC entre os resíduos é diretamente proporcional a função resposta ao impulso, isto é, $v'(L) = v_0' + v_1'L + \dots$, isto é,

$$v_k' = \sigma_\alpha \sigma_\eta \rho_{\alpha\eta}(k), \quad k = 0, 1, \dots \quad (32)$$

Substituindo os parâmetros em (32) por suas estimativas obtém-se uma estimativa inicial para v_k' .

iv) A FCC da etapa (ii) pode ser usada para identificar a forma de $\theta'(L)/\phi'(L)$ em (30). Haugh e Box (1977) mostraram que $\phi'(L) = \delta'(L)$ e $\theta'(L)$ é no máximo de ordem igual a $\delta'(L)$ ou $w'(L)$.

v) Haugh e Box (1977) derivaram ainda a relação entre as funções de transferência em (22) e (31) como

$$v(L) = \frac{w(L)}{\delta(L)} = \frac{\theta_Y(L)}{\phi_Y(L)} \frac{\phi_X(L)}{\theta_X(L)} \frac{w'(L)}{\delta'(L)} \quad (33)$$

$$\frac{\theta(L)}{\phi(L)} = \frac{\theta_Y(L)}{\phi_Y(L)} \frac{\theta'(L)}{\phi'(L)} \quad (34)$$

Portanto os resultados das etapas (i) e (iii) podem ser substituídos em (33) para obter a forma de $v(L)$ e os resultados das etapas (i) e (iv) podem ser usados para obter $\theta(L)/\phi(L)$ em (34).

O modelo de função de transferência em (22) fica em

tão completamente especificado.

3.5 Identificação de Box-Jenkins

Este enfoque descrito em Box e Jenkins (1976) consiste nas seguintes etapas

- i) Identifique o modelo ARMA para X_t

$$\phi(L)X_t = \theta(L)\varepsilon_t$$

e estime os resíduos $\hat{\alpha}_t$.

- ii) Usando o filtro estimado $\hat{\theta}(L)/\hat{\phi}(L)$ da etapa (i) transforme a série Y_t por

$$\hat{\beta}_t = \frac{\hat{\phi}(L)}{\hat{\theta}(L)} Y_t \quad (35)$$

onde a série $\hat{\beta}_t$ não é ruído branco pois o filtro (30) não foi usado em (35).

- iii) Calcule a FCC entre $\hat{\beta}_t$ e $\hat{\alpha}_t$.

- iv) Box e Jenkins (1976, pg. 380) mostraram que os parâmetros da função de transferência $v(L)$ em (22) são diretamente proporcionais a FCC da etapa (iii), isto é,

$$v_k = \frac{\sigma_\beta}{\sigma_\alpha} \rho_{\alpha\beta}(k) \quad k = 0, 1, \dots \quad (36)$$

Substituindo os parâmetros em (36) por suas estimativas obtêm-se uma estimativa inicial, para v_k , e utilizando (21) identifique (b, r, s) .

v) Com a função de transferência identificada em (iv) estime N_t através e

$$\hat{N}_t = Y_t - \hat{\nu}(L)X_t \quad (37)$$

e obtenha o modelo ARMA para resíduo \hat{N}_t . O modelo de função de transferência em (22) fica completamente especificado.

3.6 Identificação por Mínimos Quadrados e Método "Corner"

Este enfoque sugerido por Liu e Hanssens (1982) utiliza principalmente mínimos quadrados linear. Sua principal vantagem é que o mesmo pode ser estendido facilmente para o caso de diversas variáveis input (exógenas).

Sem perda de generalidade considere-se o seguinte modelo de função de transferência com dois input:

$$Y_t = c + \frac{w_1(L)}{\delta_1(L)} X_{1t} + \frac{w_2(L)}{\delta_2(L)} X_{2t} + \frac{\theta(L)}{\phi(L)} \varepsilon_t \quad (38)$$

A identificação consiste nas seguintes etapas:

- i) Para efeito de acurácia numérica na estimação de mínimos quadrados identifique modelos ARMA para X_{1t} e X_{2t} .

Se os modelos não têm termos AR ou se os polinômios AR não têm raízes próximas de 1 vá para a etapa (ii).

Se os modelos com termos AR têm raízes do polinômio AR próximas de 1, escolha um filtro comum analisando os termos AR com raízes próximas de 1. Aplique este fil

tro a todas as séries input e ao output.

- ii) Escolha M e N suficientemente grande e expresse o modelo de função de transferência (38) resultante da etapa (i) como

$$Y_t^i = c + (v_{10} + v_{11}L + \dots + v_{1M}L^M) X_{1t}^i + (v_{20} + v_{21}L + \dots + v_{2N}L^N) X_{2t}^i + u_t \quad (39)$$

Estime os parâmetros por mínimos quadrados ordinários e obtenha \hat{u}_t . Nesta etapa despreze alguns dos coeficientes considerados insignificantes.

- iii) Se \hat{u}_t é ruído branco vá para a etapa (v).

Se \hat{u}_t não é ruído branco identifique um modelo ARMA para \hat{u}_t .

- iv) Com o modelo ARMA para \hat{u}_t , reestime os parâmetros de (39) através de um dos seguintes procedimentos:

Sendo Σ^{-1} a matriz de variância covariância de \hat{u}_t

- a) Estime os parâmetros de (39) através de mínimos quadrados generalizados.
- b) Como $\Sigma^{-1} = HH'$ estime os parâmetros de (39) através de mínimos quadrados ordinários da regressão de HY_t em HZ onde

$$Z = ((v_{10} + v_{11}L + \dots + v_{1M}L^M)X_{1t}, (v_{20} + v_{21}L + \dots + v_{2N}L^N)X_{2t}).$$

- c) Ljung e Box (1979) mostraram que a transformação HY_t

e HZ é aproximadamente equivalente a filtrar o output e os inputs pelo modelo ARMA de \hat{u}_t . Portanto após filtrar o input e output, reestime os parâmetros de (39) por mínimos quadrados ordinários.

d) Estime os parâmetros de (39) por máxima verossimilhança como descrito na Seção 2.1.

v) Se nenhuma filtragem foi feita na etapa (i) o ruído do modelo é aquele obtido na etapa (iii).

Se na etapa (i) foi feita uma filtragem obtenha o ruído do modelo para as variáveis originais utilizando os pesos v 's obtidos em (ii) ou (iv) e identifique o modelo ARMA para o ruído.

vi) Se um número grande de pesos v são significantes, obtenha a forma racional $w_\ell(L)/\delta_\ell(L)$ para a série $X_{\ell t}$ usando o método Corner (Seção 5.5, Capítulo anterior) em $v_\ell(L)$ da seguinte forma.

Para a variável $X_{\ell t}$ e pesos $v_{\ell j}$ ($\ell = 1, 2$): seja V_1^* o valor máximo de $|v_{\ell j}|$, $j = 1, \dots, M$.

$$\eta_j = v_{\ell j} / V_1^*$$

Construa o determinante $P(i, j)$

$$P(i,j) = \begin{vmatrix} \eta_i & \eta_{i-1} & \dots & \eta_{i-j+1} \\ \eta_{i+1} & \eta_i & \dots & \eta_{i-j+2} \\ \dots & \dots & \dots & \dots \\ \eta_{i+j-1} & \eta_{i+j-2} & \dots & \eta_i \end{vmatrix} \quad (40)$$

(o subscrito ℓ foi omitido por simplicidade)

utilizando o método Corner, a tabela resultante tem a forma

$i \backslash j$	1	2	r	r+1	k
0	0	0		0	0		0
1	0	0		0	0		
						
b-1	0	0	0	0		0
b	P(b,1)			P(b,r)	P(b,r+1)		P(bk)
⋮							
$\Delta+b+1$							
$\Delta+b$	P($\Delta+b,1$)				0	0
⋮		
k	P(k,1)				0	0

Este procedimento é aplicado para $\ell = 1,2$ obtendo-se as formas racionais $w_\ell(L)/\delta_\ell(L)$ para cada variável $\ell = 1,2$.

7 Estimação, Verificação e Previsão

Estimação

Se os ε_t têm distribuição normal, estimativas de

máxima verossimilhança dos parâmetros são obtidas minimizando a soma dos quadrados (8) ou (9) para o modelo (22).

ii) Verificação

Após estimar o modelo de função de transferência o modelo pode ser adequado se:

- a) o modelo para o ruído N_t não é adequado;
- b) a função de transferência não está correta;
- c) ambos (a) e (b) ocorreu.

Os seguintes resultados são úteis para verificar a adequabilidade do modelo

- i) Se somente o modelo para o ruído é inadequado então

$$\begin{aligned} \rho_{\epsilon}(k) &\neq 0 \text{ para algum } k, \\ \rho_{\alpha\epsilon}(k) &= 0 \quad \forall k. \end{aligned} \quad (42)$$

- ii) Se a função de transferência é inadequada então

$$\begin{aligned} \rho_{\epsilon}(k) &\neq 0 \\ \rho_{\alpha\epsilon}(k) &\neq 0 \end{aligned} \quad \text{para algum } k. \quad (43)$$

Lembrando que $V\{r_{\hat{\epsilon}}(k)\} \approx 1/T$ e de (27),

$V\{r_{\hat{\alpha\epsilon}}(k)\} \approx \frac{1}{T-k}$ pode-se construir testes para (a), (b) ou (c). Por exemplo se $|r_{\hat{\epsilon}}(k)| > \frac{2}{\sqrt{T}}$ para algum k e $|r_{\hat{\alpha\epsilon}}| < \frac{2}{\sqrt{T-k}}$, $\forall k$ então o modelo para o ruído deve ser mo

dificado.

Além disso os seguintes testes de portmanteau podem ser empregados na forma usual (Lyung e Box, 1978). Para testar se os resíduos são autocorrelacionados calcula-se

$$Q = m(m+2) \sum_{k=1}^M \frac{r_{\hat{\epsilon}}^2(k)}{m-k} \sim \chi_{M-p-q}^2 \quad (44)$$

(p, q - parâmetros do modelo para N_t).

Para testar se os resíduos e os input são correlacionados calcula-se

$$Q = m(m+2) \sum_{k=0}^M r_{\hat{\alpha}\hat{\epsilon}}^2(k)/(m-k) \sim \chi_{M-r-d}^2, \quad (45)$$

onde m é o número de resíduos.

Além disso, se $\hat{r}_{\hat{\alpha}\hat{\epsilon}}(k) \neq 0$ para algum valor negativo de k , tem-se indicação de feedback e um modelo vetorial deve ser considerado.

iii) Previsão

Após a etapa de verificação, se o modelo é considerado adequado, o mesmo pode ser então utilizado para previsão do output Y_t .

Box e Jenkins (1976) mostram que a previsão de mínimos quadrados de Y_{t+l} é dada por

$$\hat{Y}_t(\ell) = E[Y_{t+\ell}] = E(Y_{t+\ell}/X_t, Y_t, X_{t-1}, Y_{t-2}, \dots) \quad (46a)$$

$$= \hat{\psi}(\ell)\hat{X}_t(\ell) + \hat{\psi}(\ell)\hat{\varepsilon}_t(\ell) \quad (46b)$$

$$= \frac{\hat{\omega}(\ell)}{\hat{\delta}(\ell)}\hat{X}_t(\ell) + \frac{\hat{\theta}(\ell)}{\hat{\phi}(\ell)}\hat{\varepsilon}_t(\ell) \quad (46c)$$

onde $\hat{Y}_t(\ell)$ é a previsão de $Y_{t+\ell}$ ($\hat{Y}_t(\ell) = Y_{t+\ell}$ se $\ell \leq 0$), $\hat{X}_t(\ell)$ é a previsão de $X_{t+\ell}$ ($\hat{X}_t(\ell) = X_{t+\ell}$ se $\ell \leq 0$) e $\hat{\varepsilon}_t(\ell)$ é a previsão de $\varepsilon_{t+\ell}$ ($\varepsilon_t(\ell) = \varepsilon_t$ se $\ell \leq 0$ e $\varepsilon_t(\ell) = 0$ se $\ell > 0$).

A variância da previsão é dada por

$$V\{\hat{Y}_t(\ell)\} = \sigma_{\varepsilon}^2 \sum_{j=b}^{\ell-1} \psi_j^2 + \sigma_{\varepsilon}^2 \sum_{j=0}^{\ell-1} \psi_j^2 \quad (47)$$

Outros resultados sobre previsão com função de transferência são apresentados em Pierce (1975).

4. Análise de Intervenção

O modelo de função de transferência foi visto como uma generalização dos modelos de regressão que permite relações dinâmicas entre as variáveis input e output. Nesta seção apresenta-se a correspondente generalização dos modelos de regressão com variáveis 0 e 1 (dummy).

4.1 Modelos para Intervenção

Frequentemente o comportamento de uma série temporal é afetado por determinados eventos ocasionais em certos instantes de tempo. Por exemplo, o nível de vendas de um setor industrial pode ser afetado por uma greve, a taxa de câmbio pode ser afetada por uma maxidesvalorização, a vazão de um rio pode ser afetada por um incêndio ocorrido na sua bacia hidrográfica ou por uma represa construída no seu percurso, o número de acidentes fatais na estrada pode ser afetado pela introdução de nova lei sobre o uso de cintos de segurança. Tais exemplos são chamados de intervenções e um modelo para séries com este tipo de alteração é dado em Box e Tiao (1975).

A intervenção pode afetar a série temporal de diversas formas. Ela pode alterar o nível da série abruptamente ou após alguma defasagem no tempo, pode alterar a tendência ou outras formas mais complicadas de efeitos. O modelo para intervenção fornece resposta para questões do tipo: Qual a magnitude do efeito? Qual a duração do efeito? O efeito é constante ou está aumentando?

O modelo de função de transferência com input particular é utilizado para modelar intervenções na série temporal. O modelo de análise de intervenção é obtido escolhendo para input da função de transferência (22) a função degrau $S_t(T_0)$ ou a função pulso $P_t(T_0)$ onde

$$S_t(T_0) = \begin{cases} 1 & t \geq T_0 \\ 0 & t < T_0 \end{cases} \quad P_t(T_0) = \begin{cases} 1 & t = T_0 \\ 0 & t \neq T_0 \end{cases} \quad (48)$$

Observe que a função pulso é obtida da função degrau tomando a primeira diferença, $P_t(T_0) = (1-L)S_t(T_0)$. O modelo dinâmico com M inputs X_{it} sendo $M-l$ intervenções é

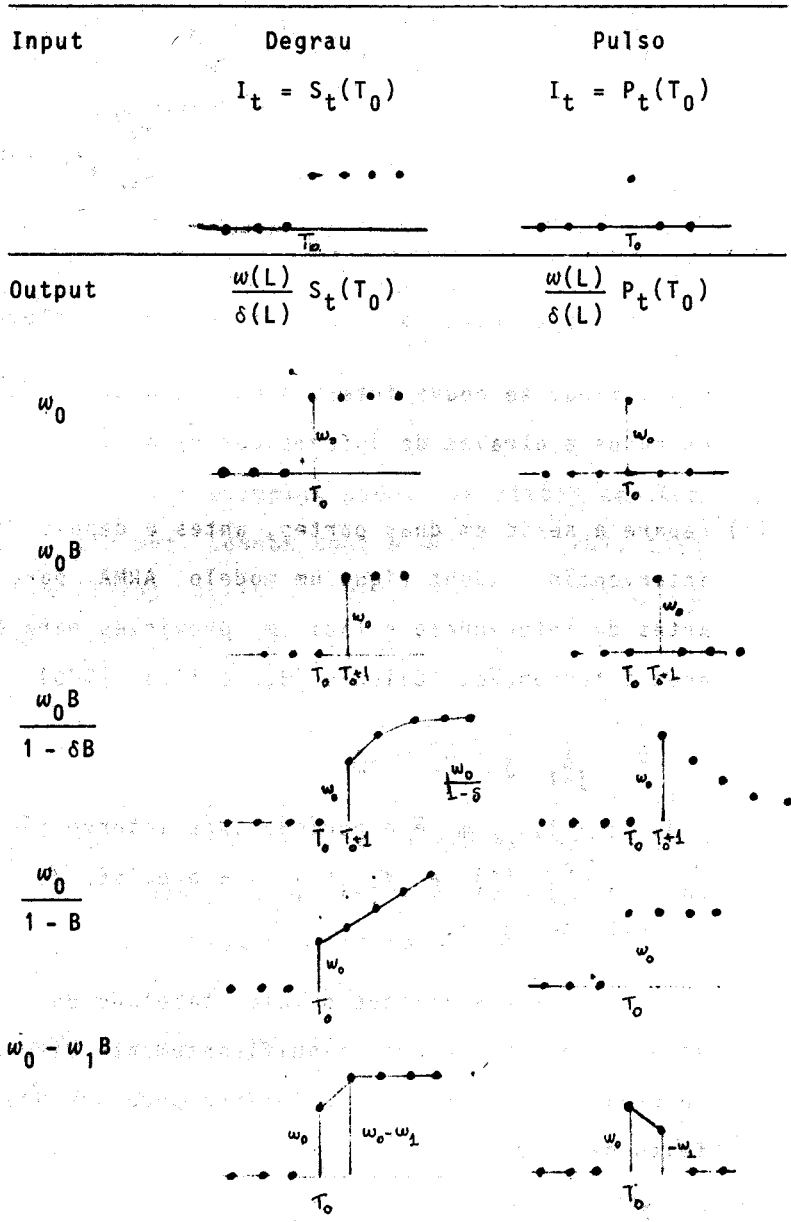
$$Y_t = \sum_{i=1}^l \frac{w_i(L)}{\delta_i(L)} X_{it} + \sum_{i=l+1}^M \frac{w_i(L)}{\delta_i(L)} I_i + \frac{\theta(L)}{\phi(L)} \varepsilon_t \quad (49)$$

onde I_t é a função pulso ou a função degrau.

Sem perda de generalidade considere-se o caso em que não há inputs X_{it} e tem-se somente uma intervenção

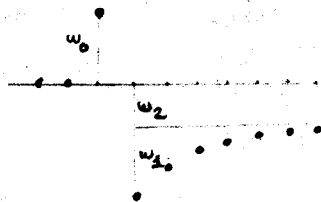
$$Y_t = \frac{w(L)}{\delta(L)} I_t + \frac{\theta(L)}{\phi(L)} \varepsilon_t. \quad (50)$$

Na seção anterior, a identificação da forma da função de transferência $v(L) = w(L)/\delta(L)$ era obtida em geral analisando os dados. Na análise de intervenção esta forma é postulada utilizando o conhecimento do fenômeno. A figura apresenta alguns casos especiais de resposta a I_t e que facilitam a construção de modelos de intervenção.



Outros comportamentos podem ser obtidos combinando alguns dos padrões da tabela, por exemplo

$$\left\{ \omega_0 + \frac{\omega_1 B}{1 - \delta B} + \frac{\omega_2 B^2}{1 - B} \right\} P_t(T_0)$$



As etapas para construir um modelo de intervenção são:

- i) Identifique se houve intervenção e quando, inspecionando os dados e através de informações exógenas.
- ii) Separe a série em duas partes, antes e depois da possível intervenção. Identifique um modelo ARMA para a série antes da intervenção e faça m previsões para o período após intervenção. Calcule (Box e Tiao, 1976)

$$Q = \frac{\sum_{j=1}^m \hat{\varepsilon}_j^2}{\hat{\sigma}_\varepsilon^2} \sim \chi_m^2$$

onde $j = 1, \dots, m$ é o período após intervenção e $\hat{\varepsilon}_j = Y_j - \hat{Y}_{j-1}(1)$ e $\hat{Y}_{j-1}(1)$ é a previsão de Y_j feita no instante $j-1$.

Se Q for maior que o valor tabelado da χ_m^2 as previsões são consideradas significativamente diferentes das observações. A razão da diferença pode ser devida ao efeito da intervenção.

- iii) Faça uma lista de supostas mudanças que a intervenção é plausível de causar. Modele a intervenção tendo em vista esta lista e os gráficos da resposta a I_t .

Uma outra maneira é tentar diversos modelos e testar

a adequabilidade dos mesmos.

- iv) Utilize o processo iterativo identificação, estimação e verificação para função de transferência para obter um modelo adequado.
- v) Em geral o interesse é medir o próprio efeito da intervenção, porém o modelo pode perfeitamente ser usado para previsão.

Maiores detalhes e exemplos podem ser vistos em Box e Tiao (1975, 1976), Hipel, Lennox, Unny e McLeod (1975), Bhat t acharyya e Layton (1979), Abraham (1980), Agreen (1981), Sa boia (1976), Pino e Morettin (1980) e Thury (1982).

1.2 Observações Perdidas

Uma dificuldade para a análise de séries temporais é a ocorrência de observações perdidas, isto é, em alguns instantes de tempo os dados não foram observados por algum motivo. É necessário então estimar esses valores para analisar a série temporal. O modelo de análise de intervenção pode ser usado com este objetivo. (Hipel, McLeod e Noakes, 1982).

Suponha que no instante $t = T_0$ a observação Y_{T_0} não foi feita, sua estimação pode ser feita através do modelo de intervenção

$$Y_t = w_0 P_t(T_0) + N_t.$$

Utilizando os procedimentos da Seção 4.1 a observação perdida é estimada como \hat{w}_0 , o modelo fornece também o erro padrão da estimativa e portanto intervalos de confiança são obtidos.

No caso de M observações perdidas o modelo seria

$$Y_t = \sum_{i=1}^M w_{0i} P_t(T_{0i}) + N_t.$$

Outros enfoques para o problema de observações perdidas pode ser visto em Lettenmaier (1980) e Harvey e Pierse (1982).

5. Modelos Vetoriais ARMA

5.1 Introdução

Nesta seção será apresentada a extensão dos modelos ARMA para séries temporais multivariadas. Para esta extensão utiliza-se os anagramas MARMA, derivado de multivariado autorregressivo-médias móveis ou ARMAV, derivado de autorregressivo-médias móveis vetorial (ver Granger, 1982).

As propriedades dos modelos ARMA vetoriais e sua utilização na análise de séries multivariadas foi considerada por vários autores e os resultados, na maioria das vezes são extensões naturais do caso univariado (ver por exemplo Quenouille (1957), Hannan (1970), Granger e Newbold (1977), Kendall, Stuart e Ord (1983), Jenkins (1979), Jenkins e Alavi (1981), Jenkins e McLeod (1982), Box e Tiao (1977), Tiao e Box (1981), Tiao e Tsay (1983), Abraham (1981), Ledolter (1978), Umashankar e Ledolter (1983), Chan e Wallis (1978), Indjehagopian (1979), Maraval (1981), Heuts e Rens (1980), Ord (1982), Terasvirta (1982), Thury (1980, 1982), Camacho e McLeod (1983), Aaker, Carman e Jacobson (1982), Moriaty e Salamon (1980), Bhattacharyya (1982), Heyse e Wey (1983), Petto (1983) e Lanna e Indursky (1983) e suas referências).

Sejam $Y_t' = (Y_{1t}, \dots, Y_{mt})$ um vetor de m séries temporais e $\varepsilon_t' = (\varepsilon_{1t}, \dots, \varepsilon_{mt})$ um vetor de m ruídos brancos com matriz de covariância cruzada

$$E(\varepsilon_t \varepsilon_{t+k}') = \begin{cases} \Sigma & k = 0 \\ 0 & k \neq 0 \end{cases} \quad (51)$$

Suponha, que (possivelmente após transformações e ou diferenciação), Y_t é estacionário, isto é, cada Y_{it} é estacionário e que as covariâncias cruzadas $\text{Cov}(Y_{it}, Y_{jt+k})$ são dependam de $k(i,j)$. A extensão multivariada do Teorema de Wold é que Y_t pode ser representada por

$$\begin{aligned} Y_t &= \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots = \\ &= (I_m + \psi_1 L + \psi_2 L^2 + \dots) \varepsilon_t = \Psi(L) \varepsilon_t \end{aligned} \quad (52)$$

onde ψ_i são matrizes $(m \times m)$ de parâmetros, $\psi_0 = I_m$, e $\Psi(L)$ um polinômio matricial infinito em L .

Em (52), Y_t é um modelo linear discreto multivariado ou $MA(\infty)$ vetorial. Uma representação alternativa obtida por substituição é

$$Y_t - \pi_1 Y_{t-1} - \pi_2 Y_{t-2} - \dots = \varepsilon_t$$

$$\text{ou} \quad (I_m - \pi_1 L - \pi_2 L^2 - \dots) Y_t = \varepsilon_t$$

$$\text{isto é, } \Pi(L) Y_t = \varepsilon_t \quad (53)$$

onde os π_i são matrizes $(m \times m)$ de parâmetros, $\pi_0 = I_m$ e $\Pi(L)$ um polinômio matricial infinito em L . Em (53) Y_t está na forma de um modelo $AR(\infty)$ vetorial.

Aproximando os polinômios matriciais infinitos pelo produto de dois polinômios matriciais finitos, isto é,

$$Y_t = \Psi(L)\varepsilon_t = \Phi^{-1}(L)\Theta(L)\varepsilon_t \quad (54)$$

$$\Pi(L)Y_t = \Theta^{-1}(L)\Phi(L)Y_t = \varepsilon_t \quad (55)$$

tem-se a representação

$$\Phi(L)Y_t = \Theta(L)\varepsilon_t \quad (56)$$

que é o modelo ARMA(p,q) vetorial, onde

$$\Psi(L) = I_m + \psi_1 L + \psi_2 L^2 + \dots$$

$$\Pi(L) = I_m - \pi_1 L - \pi_2 L^2 - \dots$$

$$\Theta(L) = I_m + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q$$

$$\Phi(L) = I_m - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$$

I_m , matriz identidade de ordem m , ψ_i , π_i , θ_i e ϕ_i são matrizes $(m \times m)$ de parâmetros e $\theta_0 = \phi_0 = I_m$.

5.2 Algumas Propriedades

Considere o modelo ARMA(p,q) vetorial (56)

$$\Phi(L)Y_t = \Theta(L)\varepsilon_t. \quad (56)$$

O modelo é dito estacionário se os zeros do polinômio determinantal $|\Phi(L)|$ estão fora do círculo unitário. Neste caso o modelo pode ser escrito na forma (52).

O modelo é dito inversível se os zeros do polinômio determinantal $|\Theta(L)|$ estão fora do círculo unitário. Neste caso o modelo pode ser escrito na forma (53).

A escolha da forma dos polinômios $\Phi(L)$ e $\Theta(L)$ e da matriz Σ foi feita para evitar redundância (identificabilidade ou estimabilidade) dos parâmetros. Em geral, com $\Theta_0 = I_m$, redundância é evitada se escolhe-se entre as parametrizações

$$a) \Phi = I_m \text{ e } \Sigma \text{ não diagonal}$$

ou

$$b) \Phi \neq I_m \text{ e } \Sigma \text{ diagonal}$$

(57)

em geral (a) é utilizada por estatísticos e engenheiros e (b) por economistas.

A matriz de covariância cruzada de lag k para Y_t é

$$\begin{aligned} \Gamma(k) &= E(Y_t Y_{t+k}') \\ &= \{\gamma_{ij}(k)\} \end{aligned} \quad \begin{aligned} k &= 0, \pm 1, \pm 2, \dots \\ i, j &= 1, 2, \dots, m \end{aligned} \quad (58)$$

e a matriz de correlação cruzada é

$$\rho(k) = \{\rho_{ij}(k)\} \quad \begin{aligned} k &= 0, \pm 1, \pm 2, \dots \\ i, j &= 1, 2, \dots, m \end{aligned} \quad (59)$$

onde $\rho_{ij}(k) = \gamma_{ij}(k) / \{\gamma_{ii}(0)\gamma_{jj}(0)\}^{1/2}$.

Note que $\Gamma(-k) = \Gamma'(k)$ e $\rho(-k) = \rho'(k)$

onde Γ' , ρ' são as matrizes transpostas de Γ e ρ , respectivamente.

Quando o modelo é estacionário mostra-se que:

$$\Gamma(k) = \Gamma(k-1)\Phi_1' + \dots + \Gamma(k-p)\Phi_p' - \Gamma_{Y_\epsilon}(k) + \Gamma_{Y_\epsilon}(k-1)\Theta_1' + \dots + \Gamma_{Y_\epsilon}(k-q)\Theta_q' \quad (60)$$

para $k \geq 0$, $\Gamma_{Y_\Sigma}(j) = E(Y_t \varepsilon_{t+j}')$ e para $k > q$ tem-se

$$\Gamma(k) = \Gamma(k-1)\phi_1' + \dots + \Gamma(k-p)\phi_p' \quad (61)$$

Se $q = 0$, isto é, o modelo é AR(p) tem-se a generalização multivariada das equações de Yule-Walker:

$$\Gamma(0) = \Gamma(-1)\phi_1' + \dots + \Gamma(-p)\phi_p' + \Sigma \quad (62)$$

$$\Gamma(k) = \Gamma(k-1)\phi_1' + \dots + \Gamma(k-p)\phi_p' \quad k \geq 1$$

Se $p = 0$, isto é, o modelo é MA(q) tem-se

$$\Gamma(0) = \Sigma + \theta_1 \Sigma \theta_1' + \dots + \theta_q \Sigma \theta_q'$$

$$\Gamma(k) = \Sigma \theta_k + \theta_1 \Sigma \theta_{1+k}' + \dots + \theta_{q-k} \Sigma \theta_q' \quad k = 1, \dots, q. \quad (63)$$

$$\Gamma(k) = 0 \quad k > p.$$

O problema da unicidade entre modelos e suas matrizes de covariância cruzadas é mais complexo no caso multivariado. Aqui não basta exigir inversibilidade; outras restrições são necessárias (ver Priestley, 198, pg. 801). Quando $\phi_0 = \theta_0 = I_m$ e Σ é não singular e com as condições de estacionariedade e inversibilidade satisfeitas, uma condição suficiente para uma representação única é

$$\text{Rank} [\phi_p, \theta_q] = m. \quad (64)$$

A representação (56) do modelo vetorial ARMA na forma (54), isto é,

$$Y_t = \phi^{-1}(L)\theta(L)\varepsilon_t$$

pode ser escrita

$$|\Phi(L)|Y_t = \text{Adj}[\Phi(L)]\Theta(L)\varepsilon_t \quad (65)$$

onde Adj denota matriz adjunta e $|\Phi(L)|$ o (polinômio escalar) determinante de $\Phi(L)$. Portanto todo modelo vetorial ARMA pode ser escrito na forma

$$\Phi^*(L)Y_t = \Theta^*(L)\varepsilon_t \quad (66)$$

onde

- a) A matriz autorregressiva $\Phi^*(L)$ é diagonal com elementos idênticos e iguais ao determinante de $|\Phi(L)|$;
- b) A matriz média móvel $\Theta^*(L)$ tem ordem máxima $q^* = q_1 \times q$ onde q_1 é a ordem máxima da matriz $\text{Adj}[\Phi(L)]$.

O resultado (66) fornece relações úteis na identificação de modelos para séries multivariadas e modelos econométricos.

Em particular tem-se

- i) Para o modelo vetorial MA(q), cada série individual segue um modelo univariado MA(q).
- ii) Para o modelo vetorial AR(p), a máxima ordem das séries individuais é ARMA(mp, (m-1)p).
- iii) Em geral cada série individual segue um modelo ARMA. Pode parecer que a parte autorregressiva das séries indivi

duais sejam idênticas. Isto não ocorre em geral, devido a possíveis cancelamentos com as partes médias móveis. Se não houver cancelamento cada série segue um modelo $ARMA(m, (m-1)(p+q))$ caso contrário esta é a maior ordem das séries individuais.

5.3 Modelos com Médias não Nulas e Sazonais

i) Modelos com Médias não Nula

Se o vetor Y_t tem média não nula $\mu' = (\mu_1, \dots, \mu_m)$ o modelo (56) é modificado para

$$\Phi(L)Y_t = C + \Theta(L)\varepsilon_t \quad (67)$$

onde $C' = (c_1, \dots, c_m) = \Theta(L)\mu'$.

ii) Modelos Sazonais

No caso univariado, sazonalidade foi introduzida através do modelo multiplicativo ARMA ou SARIMA. A atração do modelo multiplicativo é que em geral ele é mais parcimonioso comparado com modelos multiplicativos nos quais os parâmetros são introduzidos em lags selecionados (ver Seção 3.3 do capítulo anterior). A falta de parcimônia no modelo resulta em um aumento do número de parâmetros, um problema que se torna crucial no caso multivariado. O modelo multiplicativo sazonal vetorial é

$$\Phi(L)\Phi(L^S)Y_t = C + \Theta(L)\Theta(L^S)\varepsilon_t \quad (68)$$

onde $\Phi(L)$, $\Phi(L^S)$, $\Theta(L)$ e $\Theta(L^S)$ são polinômios matriciais de grau p , p_s , q , q_s , respectivamente.

Para obter maior parcimônia, Jenkins e Alavi (1981) propuseram modelos multiplicativos diagonais. Esses modelos são utilizados quando a série Y_{it} é afetada pelo que ocorre s meses atrás na série, porém não é afetada pelo que ocorreu naqueles meses com a série Y_{jt} , portanto o comportamento sazonal é restrito aos elementos das diagonais das matrizes autorregressivas ou médias móveis, isto é, $\phi(L^s)$ e $\theta(L^s)$ são diagonais. Um exemplo de modelo sazonal diagonal é

$$\begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = \begin{pmatrix} (1 - \theta_{11}L)(1 - \theta_{11}L^s) & \theta_{12}L \\ \theta_{21}L & (1 - \theta_{22}L)(1 - \theta_{22}L^s) \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

5.4 Função de Transferência e Análise de Intervenção

i) Função de Transferência

Algumas vezes as m séries $Y_t' = (Y_{1t}, \dots, Y_{mt})$ apresentam relações especiais na qual algumas funcionam como input (exógenas) das demais e não há feedback entre elas. Em particular, esta relação pode ser escrita na forma de função de transferência

$$Y_{ht} = \sum_{i=1}^{h-1} \frac{w_i(L)L^{b_i}}{\delta_i(L)} Y_{it} + \frac{\theta_h(L)}{\phi_h(L)} \varepsilon_{ht} \quad (69)$$

onde $w_i(L)$, $\delta_i(L)$, $\theta_h(L)$ e $\phi_h(L)$ são polinômios em L , os b_i são inteiros não negativos e ε_{ht} são ruídos brancos e independentes.

O modelo (69) pode ser escrito na forma da equação

(56) onde as matrizes $\Phi(L)$ e $\Theta(L)$ são triangulares inferiores.

Por exemplo, considere-se o caso MA(1) com $m=2$ e Θ_1 triangular inferior

$$\begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = \begin{pmatrix} 1 - \theta_{11}L & \\ \theta_{21}L & 1 - \theta_{22}L \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

que pode ser escrito na forma alternativa

$$\begin{aligned} Y_{1t} &= (1 - \theta_{11})\varepsilon_{1t} \\ Y_{2t} &= \frac{w_0 - w_1L}{1 - \theta_{11}L} Y_{1t} + (1 - \theta_{22})a_t \end{aligned} \quad (70)$$

onde ε_{2t} é escrito como $\varepsilon_{2t} = \beta\varepsilon_{1t} + a_t$, ε_t e a_t independentes, $w_0 = \beta$, $w_1 = \beta\theta_{22} - \theta_{21}$. A forma (70) é um caso especial da função de transferência (69).

A função de transferência (69) pode ser generalizada pela abordagem por séries temporais multivariadas permitindo que Φ e Θ sejam triangulares inferiores por bloco. Por exemplo,

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} I - \Theta_{11}L & \\ \Theta_{21}L & I - \Theta_{22}L \end{pmatrix} \begin{pmatrix} \varepsilon_{Xt} \\ \varepsilon_{Yt} \end{pmatrix}$$

onde $X_t^i = (X_{1t}, \dots, X_{nt})$ e $Y_t^i = (Y_{1t}, \dots, Y_{mt})$, com dimensões apropriadas para os outros elementos etc. Portanto as séries input X_t bem como as séries output Y_t podem ter relações de feedback internamente.

Destes resultados vê-se que o modelo vetorial ARMA generalizam os resultados de função de transferência. A identificação de modelos de função de transferência com múltiplos inputs e múltiplos outputs é tratada em Jenkins (1979), Jenkins e McLeod (1982) e Heuts e Rens (1980).

ii) Análise de Intervenção

O modelo de análise de intervenção para o caso multivariado e M intervenções nos instantes T_1, \dots, T_M utiliza o modelo

$$Y_t = R(L)I_t + \phi^{-1}(L)\theta(L)\varepsilon_t \quad (71)$$

onde $R(L)$ é uma matriz $(m \times k)$ de funções racionais de L com elementos

$$R_{ij}(L) = \frac{\omega_{ij}(L)}{\delta_{ij}(L)} \quad (72)$$

e $I_t' = (I_t(T_1), \dots, I_t(T_M))$ é um vetor de variáveis indicadoras, onde $I_t(T_j)$ toma a forma de um pulso ou de um degrau.

Maiores detalhes e aplicação são apresentados em Abraham (1980).

5.5 Identificação

Os métodos de identificação para séries multivariadas a serem apresentados são generalização do caso univariado basando na maioria das vezes substituir as funções escalares da Seção 5 do capítulo anterior por correspondentes matrizes. No

caso multivariado, entretanto, para facilitar a interpretação utiliza-se no lugar dos valores numéricos uma função indicadora com "+" para indicar que os valores são maiores que 2 desvios padrões, "-" para os valores menores que -2 desvios padrões e "." para indicar que os valores estão dentro de ± 2 desvios padrões (ver Tiao e Box, 1981 e Tiao e Tsay, 1983).

i) Matrizes de Correlações Cruzadas

A matriz de correlação cruzada amostral $R(k)$ definida com elementos

$$r_{ij}(k) = \hat{\rho}_{ij}(k) = \frac{\hat{\gamma}_{ij}(k)}{\sqrt{\hat{\gamma}_{ii}(0)\hat{\gamma}_{jj}(0)}} \quad (73)$$

onde

$$\gamma_{ij} = \sum_{t=1}^{T-k} (Y_{it} - \bar{Y}_i)(Y_{jt+k} - \bar{Y}_j)$$

$$\gamma_{ii}(0) = \hat{\sigma}_i^2 = \sum_{t=1}^T (Y_{it} - \bar{Y}_i)^2$$

$$\bar{Y}_i = \frac{1}{T} \sum_{t=1}^T Y_{it}$$

é particularmente útil na identificação de modelos vetoriais MA(q) pois, de (63), $\rho_{ij}(k) = 0$, $k > q$.

Para determinar o ponto na qual as autocorrelações se anulam pode-se usar os resultados de Bartlett (1955) sobre as distribuições de $r_{ij}(k)$ ou simplesmente usar uma aproximação supondo $\text{Var}\{r_{ij}(k)\} \approx \frac{1}{T}$.

Em vez de fazer os gráficos das correlações cruzadas e escrever os valores numéricos em $R(k)$, Tiao e Box (1981)

sugerem utilizar a função indicadora considerando o desvio padrão igual a $1/\sqrt{T}$. Um resultado típico para um modelo MA(2) para $m=2$ séries seria

R(k) em termos da função indicador					
lag k	1	2	3	4
Ind[R(k)]	$\begin{vmatrix} + & - \\ - & . \end{vmatrix}$	$\begin{vmatrix} + & + \\ - & - \end{vmatrix}$	$\begin{vmatrix} : & : \\ : & : \end{vmatrix}$	$\begin{vmatrix} : & : \\ : & : \end{vmatrix}$

ii) Matriz de Correlações Cruzadas Parciais

Da expressão (62) vê-se que as propriedades da correlação cruzada são complicadas para modelos vetoriais AR(p). Neste caso faz-se uso das matrizes de correlações cruzadas parciais ϕ_{kk} definidas como a última matriz de coeficientes ϕ_k de um modelo vetorial AR(k) para a série.

A estimativa de $\hat{\phi}_{kk}$ é sucessivamente obtida ou por regressão multivariada de mínimos quadrados ou resolvendo as equações de Yule-Walker generalizadas (62) ou finalmente com os resultados de De Jong (1976) que generalizou as fórmulas de Durbin (1960) em (38) do capítulo anterior.

Para modelos vetoriais AR(p), $\phi_{kk} = 0$, $k > p$ e para determinar o ponto a partir do qual as autocorrelações cruzadas parciais se anulam, comparam-se os elementos da matriz $\hat{\phi}_{kk}$ estimada com o desvio padrão aproximado $1/\sqrt{T}$ e utiliza-se a função indicadora. Um resultado típico para um modelo AR(1) para $m=2$ séries seria

$\hat{\phi}_{kk}$ em termos da função indicadora

lag k	1	2	3	4
Ind[$\hat{\phi}_{kk}$]	$\begin{vmatrix} - & - \\ - & + \end{vmatrix}$	$\begin{vmatrix} : & : \\ : & : \end{vmatrix}$	$\begin{vmatrix} : & : \\ : & : \end{vmatrix}$	$\begin{vmatrix} : & : \\ : & : \end{vmatrix}$

Para ajudar a determinação da ordem do modelo autor regressivo, pode-se empregar a razão de verossimilhança para testar a hipótese $\phi_\ell = 0$ contra a alternativa $\phi_\ell \neq 0$ quando um $AR(\ell)$ é ajustado. Se

$$S(\ell) = \sum_{t=1}^T (Y_t - \phi_1 Y_{t-1} - \dots - \phi_\ell Y_{t-\ell})(Y_t - \phi_1 Y_{t-1} - \dots - \phi_\ell Y_{t-\ell})' \quad (74)$$

é a matriz da soma dos quadrados e produtos cruzados dos resíduos após ajustar um modelo $AR(\ell)$ a razão de verossimilhança é a razão dos determinantes

$$U = |S(\ell)| / |S(\ell-1)| \quad (75)$$

e a estatística

$$M(\ell) = -(N - \frac{1}{2} - \ell m) \ln U$$

sob a hipótese nula tem distribuição χ^2 com m^2 graus de liberdade e $N = T - p - 1$, se a média é não nula e incluída e $N = T - p$ caso contrário.

iii) Matriz de Arranjos R^* , S^* e método Corner

Para a identificação de modelos vetoriais $ARMA(p, q)$ Tiao e Box (1981) sugerem estender os resultados de Gray, Kelley e McIntire (1978) e Bequin, Gourieroux e Monfort (1980) para o caso multivariado.

Das relações (60) e (61) para um modelo vetorial ARMA (p,q) as matrizes de covariância cruzadas $\Gamma(\ell)$ são relacionadas aos parâmetros autorregressivos ϕ 's por

$$\begin{bmatrix} \Gamma(i) & \Gamma(i+1) & \dots & \Gamma(i-p+2) & \Gamma(i-p+1) \\ & \Gamma(i+1) & & & \vdots \\ & \vdots & & & \vdots \\ \Gamma(i+p-2) & \dots & \Gamma(i+1) & \Gamma(i) & \Gamma(i-1) \\ \Gamma(i+p-1) & & & \Gamma(i+1) & \Gamma(i) \end{bmatrix} \begin{bmatrix} \phi' \\ \vdots \\ \vdots \\ \phi'_{p-1} \\ \phi'_p \end{bmatrix} =$$

$$= \begin{bmatrix} \Gamma(i+1) \\ \vdots \\ \vdots \\ \Gamma(i+p-1) \\ \Gamma(p+i) \end{bmatrix} \quad (76)$$

$i = q > q+1, \dots$

$$\text{ou } \begin{vmatrix} A(p,i) & b(p,i) \\ g'(p,i) & \Gamma(i) \end{vmatrix} \begin{vmatrix} \phi_{p-1} \\ \phi'_p \end{vmatrix} = \begin{vmatrix} c(p,i) \\ \Gamma(p+i) \end{vmatrix} \quad (77)$$

Considere agora as matrizes (m x m)

$$D(\ell, i) = \begin{vmatrix} d_{kj}(\ell, i) \end{vmatrix} \quad \begin{array}{l} \ell = 1, 2, \dots \\ i = 0, 1, \dots \end{array}$$

onde $d_{kj}(\ell, i)$ é o determinante

$$d_{kj}(\ell, i) = \det \begin{vmatrix} A(\ell, i) & c_j(\ell, i) \\ g'_k(\ell, i) & \gamma_{kj}(\ell+i) \end{vmatrix}$$

onde $c_j(\ell, i)$ é a j -ésima coluna de $C(\ell, i)$, $g_k(\ell, i)$ é a k -ésima linha de $g'(\ell, i)$ e $\gamma_{kj}(\ell + i)$ é o (k, j) -ésimo elemento de $\Gamma(\ell + i)$. Mostra-se que se

$$D(\ell, i) = 0 \quad \ell > p \quad \text{e} \quad i \geq q \quad (78)$$

tem-se então uma generalização do método Corner.

A apresentação deste resultado é feita como no caso univariado apenas considerando agora no lugar de escalar, matrizes $(m \times m)$ na tabela correspondente.

iv) Matriz de Correlação Cruzada Parcial Generalizada

Para identificação de modelos vetoriais ARMA(p,q), Jenkins e Alavi (1981) propõem utilizar a matriz de correlação cruzada parcial q -condicionada ou generalizada. Esta matriz denotada por $\phi_{kk}^{(i)}$ consiste na obtenção do último coeficiente autorregressivo ao ajustar um modelo ARMA(k,i). O procedimento é uma extensão da função de autocorrelação parcial generalizada para o caso multivariado e consiste em supor sucessivamente $q=1,2,\dots$ resolver para a última matriz ϕ_k , para $h = q+1, q+2, \dots, q+k$ o sistema (61), isto é,

$$\Gamma(h) = \Gamma(h-1)\phi_1' + \Gamma(h-2)\phi_2' + \dots + \Gamma(h-k)\phi_k' \quad (79)$$

Para um modelo vetorial ARMA(p,q), as matrizes $\hat{\phi}_{kk}^{(q)}$ se anulam para $k > p$ e este resultado é utilizado na identificação de p e q .

A apresentação do resultado pode ser feito em uma

tabela como no caso univariado considerando no lugar de escalar, matrizes $(m \times m)$. Em particular, a função indicadora pode ser usada.

v) Matriz de Correlação Cruzada Estendida

Para identificar modelos vetoriais ARMA(p,q) Tiao e Tsay (1983) propuseram estender a função de autocorrelação estendida para o caso multivariado. O procedimento é análogo ao caso univariado bastando considerar regressões iterativas multivariadas no lugar de regressões iterativas univariadas e considerar as matrizes de correlações cruzadas dos resíduos multivariados no lugar de funções de autocorrelações dos resíduos.

A apresentação dos resultados é análoga ao caso univariado e pode-se utilizar a função indicadora com desvio padrão aproximado $1/\sqrt{T-i-j}$ para a matriz da posição (i,j) da tabela.

Para $m=2$ um resultado típico para um modelo bivariado ARMA(p,q) é obtido considerando em (70) do capítulo anterior cada escalar 0 como uma matriz

$$\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix}$$

e cada escalar X como uma matriz

$$\begin{vmatrix} X & X \\ X & X \end{vmatrix}$$

onde X não podem ser todos ".".

Em particular para um modelo bivariado ARMA(1,2) se obteria

MA AR	0	1	2	3	4
0	$\begin{vmatrix} + & + \\ + & + \end{vmatrix}$	$\begin{vmatrix} + & + \\ + & + \end{vmatrix}$	$\begin{vmatrix} - & - \\ - & - \end{vmatrix}$	$\begin{vmatrix} - & - \\ - & - \end{vmatrix}$	$\begin{vmatrix} & \\ & \end{vmatrix}$
1	$\begin{vmatrix} + & + \\ + & + \end{vmatrix}$	$\begin{vmatrix} + & - \\ - & + \end{vmatrix}$	$\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix}$	$\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix}$	$\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix}$
2	$\begin{vmatrix} - & - \\ - & - \end{vmatrix}$	$\begin{vmatrix} - & - \\ + & + \end{vmatrix}$	$\begin{vmatrix} - & - \\ - & - \end{vmatrix}$	$\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix}$	$\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix}$
					$\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix}$

vi) Critérios Automáticos

A generalização dos critérios de identificação automática são discutidos em Priestley (1981, pg. 690).

A Figura III1 apresenta alguns dos critérios.

Autores	Modelo	Critério a ser minimizado
Akaike	AR(p)	$AIC(p) = T \ln \hat{\Sigma}_\epsilon^{(p)} + 2m^2 p$
Parzer	AR(p)	$CAT(p) = \text{traço} \left\{ \frac{m}{T} \sum_{j=1}^p \left[\hat{\Sigma}_\epsilon^{(j)} \right]^{-1} \right\} - \left[\hat{\Sigma}_\epsilon^{(p)} \right]^{-1}$
Hannan	ARMA(p,q)	$\phi = \ln \hat{\Sigma}_\epsilon + (p,q) C_T / T$

Fig. III1. Critérios de Identificação Automática.

onde:

$\hat{\Sigma}_\epsilon(j)$ = estimativa da matriz de covariância dos resíduos ao ajustar um modelo vetorial AR(j), obtida resolvendo para Σ em (62),

$$\tilde{\Sigma}_\epsilon(j) = \frac{T}{T - mj} \hat{\Sigma}_\epsilon(j),$$

$\hat{\Sigma}_\epsilon$ = estimativa da matriz de covariância dos resíduos de um modelo ARMA(p,q) ajustado,

C_T = uma seqüência que $\rightarrow \infty$ quando $T \rightarrow \infty$.

5.6 Estimação e Previsão

Identificado o modelo, a estimativa dos parâmetros $\bar{\epsilon}$ é obtida supondo ϵ_t com distribuição normal multivariada, i.e., $N_m(0, \cdot)$ e maximizando a função de verossimilhança. Em seguida testes de verificação da adequabilidade do modelo, são efetuados.

Duas funções objetivo podem ser consideradas na etapa de otimização: a função de verossimilhança condicionada e a função de verossimilhança exata,

i) Função de Verossimilhança Condicional

Para um modelo ARMA(p,q) a série $\bar{\epsilon}$ é considerada como consistindo de $T - p$ observações Y_{p+1}, \dots, Y_T . A função de verossimilhança $\bar{\epsilon}$ é determinada a partir de $\epsilon_{p+1}, \dots, \epsilon_T$, usando os valores Y_1, \dots, Y_p e considerando $\epsilon_{p-q-1} = \dots = \epsilon_p = 0$. Sendo

$$(\phi, \theta, \Sigma) = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \Sigma)$$

tem-se

$$L_C(\phi, \theta, \Sigma) \propto |\Sigma|^{-\frac{T-p}{2}} \exp \left\{ -\frac{1}{2} \text{traço } \Sigma^{-1} S_C(\phi, \theta) \right\} \quad (80)$$

onde

$$S_C(\phi, \theta) = \sum_{t=p+1}^T \varepsilon_t \varepsilon_t' \quad (81)$$

i) Função de Verossimilhança Exata

Como no caso univariado, se os zeros de $|\theta(L)|$ estão perto do círculo unitário a aproximação (80) é inadequada. Hilmer e Tiao (1979) e Nicholls e Hall (1979) derivaram a função de verossimilhança exata que pode ser expressa como

$$L(\phi, \theta, \Sigma) = L_C(\phi, \theta, \Sigma) L_1(\phi, \theta, \Sigma). \quad (82)$$

ii) Estimação não Linear

Após obter a função de verossimilhança, como é impossível obter uma solução explícita, métodos numéricos têm que ser utilizados. Hilmer e Tiao (1979) maximizam a função de verossimilhança $L(\phi, \theta, \Sigma)$ iterativamente:

- Determina-se um valor inicial $(\phi, \theta)_0$ usando (60), (61) ou (62) ou (63).
- Este valor inicial é usado para estimar

$$\hat{\Sigma}_0 = \frac{1}{n} \sum \hat{\varepsilon}_t \hat{\varepsilon}_t'$$

- Um procedimento de mínimos quadrados não linear é usado para determinar outro valor $(\phi, \theta)_1$.

- Uma nova estimativa $\hat{\Sigma}_1$ é obtida.
- O procedimento continua obtendo-se estimativas alternadamente até convergir.

iv) Verificação

Como no caso univariado após a estimação, testes de verificação da adequabilidade do modelo são obtidos através de:

- a) Introdução de parâmetros adicionais que são comparados com seus desvio padrão.
- b) Análise das matrizes de correlação cruzada e cruzada parcial dos resíduos. Novamente utilizando a função indicadora com desvio padrão $1/\sqrt{T}$.
- c) Testes de portmanteau propostos por Hosking (1980,1981) e Li e McLeod (1981).

5.7 Previsão

Seja $\hat{Y}_t(\ell)$ a previsão de erro quadrático médio de $Y_{t+\ell}$. Seja também

$$c_t(\ell) = Y_{t+\ell} - \hat{Y}_t(\ell)$$

o correspondente vetor de erros de previsão. Então $\hat{Y}_t(\ell)$ é computado recursivamente de

$$\begin{aligned} \hat{Y}_t(\ell) = & \phi \hat{Y}_t(\ell - 1) + \dots + \phi_p \hat{Y}_t(\ell - p) + \\ & + \theta_1 E[\varepsilon_{t+\ell-1}] + \dots + \theta_q E[\varepsilon_{t+\ell-q}], \end{aligned} \quad (83)$$

onde

$$\hat{Y}_t(j) = Y_{t+j} \quad j \leq 0$$

$$E[\varepsilon_{t+j}] = \begin{cases} 0 & j > 0 \\ \varepsilon_{t+j} & j = 0 \end{cases}$$

Se ε_t tem distribuição normal multivariada o vetor de erros de previsão ℓ passos a frente tem distribuição normal multivariada com vetor de média zero e matriz de covariância

$$V\{e_t(\ell)\} = \sum_{i=0}^{\ell-1} \Psi_i \Sigma \Psi_i' \quad (84)$$

As previsões podem ser atualizadas por

$$\hat{Y}_{t+1}(j) = \hat{Y}_t(j+1) + \Psi_j \varepsilon_{t+1} \quad (85)$$

onde $\varepsilon_{t+1} = Y_{t+1} - \hat{Y}_t(1) = e_t(1)$.

5.8 Análise de Autovalores e Autovetores

Nesta seção apresenta-se alguns tipos de análise de autovalores e autovetores úteis na análise de séries temporais multivariadas. Estas análises têm por objetivo: (Camacho e McLeod, 1983):

- detectar relações lineares exatas entre as séries;
- ajudar a interpretar o modelo ajustado.

Essas análises são baseadas nos autovalores e autovetores das matrizes: $\Gamma(0)$, Σ , Φ_j , Θ_j , $\Gamma^{-1}(0)\Gamma_{\hat{Y}}(0)$, onde $Y_t = \hat{Y}_{t-1}(1) + \varepsilon_t$ e $\Gamma_{\hat{Y}}(0) = E\{\hat{Y}_{t-1}(1)\hat{Y}_{t-1}'(1)\}$ é a matriz de co

variância de $\hat{Y}_{t-1}(1)$.

i) Relações Lineares Contemporâneas

Suponha que tenha-se M relações lineares da forma

$$c_j' Y_t = 0 \quad (86)$$

onde $c_j' = (c_{1j}, \dots, c_{mj})$. Seria conveniente que o analista soubesse da existência de tal relação em geral porém isto não ocorre. Os autovalores de $\Gamma(0)$ podem ser usados para detectar se tal relação existe. Se M relações (86) existem então M autovalores de $\Gamma(0)$ são zero. Em virtude de erros de arredondamento possíveis, autovalores de $\Gamma(0)$ próximos de zero fornecem indicação de relações lineares nos dados. Esta análise ajuda a decidir que variáveis devem ser incluídas na análise.

ii) Relações Lineares Defasadas

Séries temporais multivariadas podem apresentar relações lineares da forma

$$h_1' Y_t + h_2' Y_{t-1} + \dots + h_r' Y_{t-r} = 0 \quad (87)$$

onde h_i são vetores $(k \times 1)$ de coeficientes. Box e Tiao (1977) mostraram que neste caso a matriz de covariância dos resíduos de um modelo vetorial $AR(p)$ é singular para $p \geq r$. Por outro lado se alguns dos autovalores da matriz de covariância Σ são zero, então existem relações da forma (87). Isto significa que o vetor Y_t

é gerado por um vetor de ruídos m_1 -dimensional ($m_1 < m$) onde m_1 é o rank de Σ .

Portanto, relações defasadas podem ser detectadas pelos autovalores de $\hat{\Sigma}$ obtido do ajustamento de um modelo vetorial AR(j).

iii) Análise Canônica

Suponha que o processo é escrito na forma

$$Y_t = \hat{Y}_{t-1}(1) + \varepsilon_t \quad (88)$$

Se o modelo para a série é AR(p)

$$Y_{t-r}(1) = \sum_{j=1}^p \pi_j Y_{t-j} \quad (89)$$

Suponha-se que tem-se somente uma série Y_{1t} . Se Y_{1t} é estacionária, de (88),

$$\sigma_{Y_1}^2 = \sigma_{\hat{Y}_1}^2 + \sigma_{\varepsilon_t}^2 \quad (90)$$

e a quantidade $\lambda = \sigma_{\hat{Y}_1}^2 / \sigma_{Y_1}^2$ é uma medida da preditabilidade da série usando seu passado.

Para o caso de m séries, considera-se combinações lineares da forma

$$U_t = C' Y_t = \hat{U}_t + V_t = C' \hat{Y}_{t-1}(1) + C' \varepsilon_t \quad (91)$$

e deseja-se aquela combinação que fornece a máxima preditabilidade. Neste caso

$$\lambda = \frac{\sigma_U^2}{\sigma_{\hat{U}}^2} = \frac{C' \Gamma_{\hat{Y}}(0) C}{C' \Gamma_Y(0) C} \quad (92)$$

Box e Tiao (1977) mostram que λ é o maior autovalor de $\Gamma_Y^{-1}(0)\Gamma_\varphi(0)$ e C é o autovetor associado. Portanto, os autovalores de $\Gamma_Y^{-1}(0)\Gamma_\varphi(0)$ podem ser usados para classificar relações lineares da forma $C'Y_t$ de mais preditíveis a menos preditíveis, bastando ordenar seus autovalores. O menos preditível corresponderá quase a um ruído branco.

6. Outras Metodologias

Da mesma forma que na Seção 8 do capítulo anterior, nesta seção faz-se uma apresentação de metodologias alternativas de Box-Jenkins-Tiao.

6.1 Modelagem Recursiva de Hannan-Kavalieris

É uma extensão da metodologia de Hannan-Rissanen (1982) para os modelos vetoriais ARMA e ARMAX e apresentada em detalhes na monografia de Hannan e Kavalieris (1983). O critério para determinação de ordens é a extensão de Hannan-Quinn

$$T \ln \det \hat{\Sigma}(n) + d(n)C(T) \quad C(T) > 2 \quad (93)$$

onde $n = (p, q, r)$ é a ordem, $C(T)$ é uma seqüência determinada em cada etapa, $\hat{\Sigma}(n)$ é uma estimativa de Σ quando a ordem é n , e $d(n)$ é o número de parâmetros.

6.2 Modelagem ARMAV(n, n-1) de Pandit-Wu

É uma extensão da metodologia da Seção 8.2 do capítulo II. No caso de m séries o critério F é (ver Pandit, 1973) ou Prins (1980)

$$\Lambda = \frac{|A_0|}{|A_1|} \quad (94)$$

$$F = (1 - \Lambda^{1/\ell})/m\lambda \div \Lambda(k\ell - 2\lambda) \sim F(m, (k\ell - 2\lambda))$$

$$k = T - r - \frac{m - \delta - 1}{2}$$

$$L = \sqrt{(m^2 \delta^2 - 4) / (m^2 + \delta^2 - 5)}$$

$$\lambda = (m\delta - 2) / 4$$

onde

A_0 = matriz de produtos cruzados do modelo ajustado.

r = número de parâmetros no modelo (= $(p+q) \times m^2$ para um modelo ARMAV(p,q)).

δ = número de parâmetros restritos a serem iguais a zero.

A_1 = matriz de produtos cruzados do modelo com menos parâmetros.

T = número de observações.

m = número de séries.

6.3 Modelagem em Espaço de Estados de Akaike

A metodologia de Akaike é facilmente estendida para o caso de modelos vetoriais ARMA (Akaike, 1976) e ARMAX (Cooper e Wood, 1982b). Na realidade a sua importância e relevância nas aplicações é exatamente no caso multivariado. Considerando os análogos vetoriais dos resultados da Seção 8.4 do Capítulo II, como por exemplo matrizes $(m \times m)$ ϕ_j e Σ_j no lugar de ϕ_j e γ_j , respectivamente, as etapas do procedimento são análogas. Apenas indicaremos alguns pontos necessários para uma utilização correta. Etapas:

i) $\hat{\Sigma}_j$ no lugar de Y_j .

ii) Análogo ao caso univariado.

$$Y^- = (Y_{1t}, \dots, Y_{mt}, Y_{1t-1}, \dots, Y_{mt-1}, \dots, Y_{1t-M}, \dots, Y_{mt-M})$$

$$Y_M^+ = (Y_{1t}, \dots, Y_{mt}, Y_{1t+1}, \dots, Y_{mt+1}, \dots, Y_{1t+M}, \dots, Y_{mt+M})$$

iv) Análogo ao caso univariado.

v) Inicia-se o procedimento com $Z = (Y_{1t}, \dots, Y_{mt}, Y_{1t+1})$.

vi) Análoga. Porém note que ao voltar para a etapa iv) vai-se entrar com uma componente $Y_{i,t+j}$ de cada vez e não com o vetor Y_{t+j} inteiro.

ii) Da mesma forma

$$Z^{k+1} = \frac{b_1}{b_{k+1}} Z^1 - \dots - \frac{b_k}{b_{k+1}} Z^k \quad (95)$$

Se $Z^{k+1} = Z_{it+1}$ ($1 \leq i \leq k$) esta equação determina a i -ésima linha da matriz F

$$F(i,j) = \begin{cases} -\frac{b_j}{b_{k+1}} & j = 1, 2, \dots, k \\ 0 & \text{outros valores} \end{cases} \quad (96)$$

Se $Z_{it+1} = Y_{j,t+k_j}$, descarte as variáveis $Y_{j,t+k_j}$, $Y_{j,t+k_j+1}, \dots$ do vetor Y_M^+ da etapa (ii) para construir o vetor Z da etapa (v) e (vi). Se Y_M^+ ainda tem componentes retorne à etapa (iv) entrando com mais uma componente do vetor Y_M^+ .

A obtenção das matrizes Ψ é análoga ao caso univa

riado, bem como a obtenção das previsões.

Para modelos ARMAX, (Cooper e Wood (1982b)) o vetor Y^T é construído incluindo as variáveis exógenas.

6.4 Previsão por Componentes Principais

Este método empírico foi sugerido em Crookes, Croston e Sypas (1980) e em Otter e Shuur (1982) e consiste nas etapas

i) Dadas T observações de m séries temporais suponha que

$Y_t = (Y_{1t}, \dots, Y_{mt})$ foi normalizada, isto é, cada Y_{it} tem média zero e variância um. Efetua-se uma análise de componentes principais nestas m variáveis diagonalizando a matriz de correlação cruzada de lag zero, $R(0)$ obtendo m novas variáveis não correlacionadas

$$Z_t = (Z_{1t}, \dots, Z_{mt}) = (Z_t^1, Z_t^2).$$

ii) Determina-se as r ($r < m$) primeiras componentes responsáveis pela maior parte da variação dos dados, isto é,

$$Z_t^1 = (Z_{1t}, \dots, Z_{rt}) \quad (97)$$

iii) Como

$$Z_t = Y_t C = Y_t (C_1, C_2) \quad (98)$$

onde C uma matriz ($k \times k$) de autovetores de $R(0)$, tem-se que Y_t pode ser decomposta como

$$Y_t = Z_t^1 C_1 + Z_t^2 C_2$$

e considera-se $Z_t^2 C_2$ como ruído branco, isto é, tem-se

a regressão

Y_t = Z_t^1 C_t + e (99)

iii) Como os Z_it são não correlacionados utiliza-se a metodo logia univariada de Box-Jenkins para fazer previsões para Z_1t, ..., Z_rt. Através de (99) obtém-se as previsões para os Y_it de interesse.

Previsões para 14 variáveis da economia da Holanda foram obtidas por Otter e Shuur (1982) com resultados comparáveis às previsões do órgão oficial.

7. Causalidade

7.1 Definições

O conceito de causalidade entre variáveis (por exemplo X e Y) e a inferência sobre a direção da causalidade (isto é, X causa Y ou Y causa X) baseada em dados tem a longo tempo ocupado a atenção de estatísticos e economistas bem como outros cientistas. Um exemplo clássico é a relação entre incidência de câncer e fumo (ver Burch, 1978). A dificuldade que se apresenta é a impossibilidade de inferir a direção da causalidade dada a correlação obtida dos dados.

Em séries temporais, aceitando-se o axioma que o futuro não pode causar o presente e o passado, é possível inferir causalidade. Esta possibilidade levou Granger (1969) a dar uma definição de causalidade possível de ser testada empiricamente.

Granger (1969) define causalidade entre duas séries em termos de predictabilidade. A variável X causa outra variável Y com relação a um conjunto de informações ou universo, que inclui X e Y se o valor atual Y é melhor predito usando valores passados de X do que não os usando, todas as outras informações relevantes sendo usadas em cada caso, inclusive valores passados de Y .

Mais especificamente sejam X_t e Y_t duas séries temporais, seja

$$\{A_t : t = 0, \pm 1, \pm 2, \dots\}$$

um dado conjunto de informações, incluindo pelo menos $\{(X_t, Y_t)\}$ e sejam

$$\bar{A}_t = \{A_s : s < t\} \text{ e } \tilde{A}_t = \{A_s : s \leq t\}$$

De maneira similar defina-se $\bar{X}_t, \bar{Y}_t, \tilde{X}_t$ e \tilde{Y}_t .

Dado um conjunto de informações B, seja

$$P_t(Y/B)$$

o preditor um passo a frente para Y dado o conjunto de informações B e seja,

$$\sigma^2(Y/B)$$

o erro quadrático médio resultante.

A definição de Granger é

a) X causa Y se

$$\sigma^2(Y/\bar{A}) < \sigma^2(Y/A - \bar{X}) \quad (100)$$

b) X causa Y instantaneamente se

$$\sigma^2(Y/\bar{A}, \tilde{X}) < \sigma^2(Y/\bar{A}) \quad (101)$$

ausalidade de Y para X é definida de forma análoga.

Feedback ocorre se X causa Y e Y causa X. Quando X causa Y diz-se também que X é um indicador antecedente de Y ou que Y é endôgena em relação a X.

Esta definição não requer que o sistema seja linear, se este é o caso, comparam-se predições lineares.

A generalização para o caso de X_t e Y_t serem vetores é imediata, bem como sua verificação empírica. Por outro lado, sua generalização para caso de um conjunto mais amplo de informação, por exemplo, causalidade entre X e Y na presença de informação sobre Z necessita da consideração de preditores h passos a frente e sua verificação empírica é uma questão em aberto. (ver Tjostheim (1981), Lutkepohl (1982), Hsiao (1982) e Bagsham (1983)).

7.2 Caracterização

Em vista dos comentários anteriores considera-se apenas o caso bivariado. Um primeiro passo na construção de testes para a direção da causalidade será relacionar a definição de Granger com a estrutura dos modelos paramétricos relacionando as séries temporais.

i) Modelos Autorregressivos - Médias Móveis

Se (X_t, Y_t) admite uma representação vetorial ARMA (p, q)

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} - \sum_{j=1}^p \Phi_j \begin{pmatrix} X_{t-j} \\ Y_{t-j} \end{pmatrix} = \begin{pmatrix} \alpha_t \\ \eta_t \end{pmatrix} - \sum_{j=1}^q \Theta_j \begin{pmatrix} \alpha_t \\ \eta_t \end{pmatrix} \quad (102)$$

ou

$$\begin{bmatrix} \phi_{11}(L) & \phi_{12}(L) \\ \phi_{21}(L) & \phi_{22}(L) \end{bmatrix} \begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \theta_{11}(L) & \theta_{12}(L) \\ \theta_{21}(L) & \theta_{22}(L) \end{bmatrix} \begin{bmatrix} \alpha_t \\ \eta_t \end{bmatrix}$$

onde $\phi_{ii}(0) = 1 = \theta_{ii}(0) \quad (i = 1, 2)$

$$\phi_{ij}(0) = 0 = \theta_{ij}(0) \quad (i \neq j)$$

e $\epsilon_t = (\alpha_t, \eta_t)$ é um vetor de ruídos brancos tal que

$$E(\epsilon_t \epsilon_t') = \Sigma$$

$$E(\epsilon_t \epsilon_{t-j}') = 0 \quad (j \neq 0).$$

Mostra-se então que Y não causa X se e só se

$$\phi_{12}(L) = 0 = \theta_{12}(L) \quad (103)$$

e causalidade instantânea ocorre se e só se Σ não é diagonal.

Se (X_t, Y_t) admite a representação vetorial AR(∞).

$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \sum_{j=1}^{\infty} \Pi_j \begin{bmatrix} X_{t-j} \\ Y_{t-j} \end{bmatrix} + \begin{bmatrix} \alpha_t \\ \eta_t \end{bmatrix} \quad (104)$$

ou

$$\begin{bmatrix} \Pi_{11}(L) & \Pi_{12}(L) \\ \Pi_{21}(L) & \Pi_{22}(L) \end{bmatrix} \begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \alpha_t \\ \eta_t \end{bmatrix}$$

onde $\Pi_{ii}(0) = 1 \quad (i = 1, 2)$

$$\Pi_{ij}(0) = 0 \quad (i \neq j).$$

Mostra-se que Y não causa X se e só se

$$\Pi_{12}(L) = 0 \quad (105)$$

e causalidade instantânea ocorre se e só se Σ não é diagonal.

Se (X_t, Y_t) admite a representação vetorial MA(∞)

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} \alpha X_t \\ \eta Y_t \end{pmatrix} + \sum_{j=1}^{\infty} \psi_j \begin{pmatrix} \alpha_{t-j} \\ \eta_{t-j} \end{pmatrix} \quad (106)$$

ou

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{bmatrix} \psi_{11}(L) & \psi_{12}(L) \\ \psi_{21}(L) & \psi_{22}(L) \end{bmatrix} \begin{pmatrix} \alpha_t \\ \eta_t \end{pmatrix}$$

onde $\psi_{ii}(0) = 1 \quad (i = 1, 2)$

$$\psi_{ij}(0) = 0 \quad i \neq j.$$

Mostra-se que Y não causa X se e só se

$$\psi_{12}(L) = 0 \quad (107)$$

e causalidade instantânea ocorre se e só se Σ não é diagonal.

ii) Filtros Bilaterais

Uma caracterização alternativa de causalidade é dada a Sims (1972) que considerou a relação

$$Y_t = \sum_{j=-\infty}^{\infty} \beta_j X_{t-j} + u_t \quad (108)$$

onde u_t é não correlacionado com X_{t-j} , mas não necessariamente ruído branco. Sims mostrou que X causa Y.

na ausência de feedback, se

$$\beta_j = 0, \quad j = -1, -2, \dots$$

Entretanto, se há feedback, $\beta_0 = 0$ não necessariamente implica na ausência de causalidade instantânea.

iii) Correlações Cruzadas dos Ruídos Univariados

Suponha que as séries X e Y são estacionárias ou foram reduzidas a estacionárias após convenientes transformações (Box-Cox, diferenças etc), e que as mesmas admitem a representação ARMA

$$\phi_X(L)X_t = \theta_X(L)\alpha_t, \tag{109}$$

$$\phi_Y(L)Y_t = \theta_Y(L)\eta_t.$$

Considere-se as correlações cruzadas

$$\rho_{\alpha\eta}(k) = \text{Corr}(\alpha_t, \alpha_{t+k}) \quad k = 0, 1, \dots \tag{110}$$

que além de refletirem o tipo de dependência linear entre α e η e conseqüentemente entre X e Y , fornecem também o tipo de causalidade entre essas variáveis para sistemas lineares. Na Figura III2, são apresentados resultados de Pierce e Haugh (1977) sobre algumas das relações de causalidade caracterizadas por restrições em $\rho_{\alpha\eta}(k)$. Devido aos resultados de Pierce (1979) e Pierce e Haugh (1979) qualquer das relações na Figura III2 envolvendo causalidade instantânea são válidas na ausência de feedback. Em particular Y não causa X , na ausência de feedback se e só se

$$\rho_{\alpha\eta}(k) = 0 \quad k = 0, 1, 2, \dots \quad (111)$$

e se há feedback $\rho_{\alpha\eta}(0) = 0$ não necessariamente implica na ausência de causalidade instantânea.

Figura III2 - Relações de Causalidade entre Duas Variáveis

Relação	Restrição em $\rho_{\alpha\eta}(k)$
X causa Y	$\rho_{\alpha\eta}(k) \neq 0$ algum $k > 0$
Y causa X	$\rho_{\alpha\eta}(k) \neq 0$ algum $k < 0$
Causalidade instantânea	$\rho_{\alpha\eta}(0) \neq 0$
Feedback	$\rho_{\alpha\eta}(k) \neq 0$ algum $k < 0$ e algum $k > 0$
X causa Y mas não instantaneamente	$\rho_{\alpha\eta}(k) \neq 0$ algum $k > 0$ e $\rho_{\alpha\eta}(0) = 0$
Y não causa X	$\rho_{\alpha\eta}(k) = 0 \quad \forall k < 0$
Y não causa X nunca	$\rho_{\alpha\eta}(k) = 0 \quad \forall k \leq 0$
Causalidade unidirecional de X para Y	$\rho_{\alpha\eta}(k) \neq 0$ algum $k > 0$ e $\rho_{\alpha\eta}(k) = 0$ para (a) $\forall k < 0$ ou (b) $\forall k \leq 0$
X e Y são relacionadas instantaneamente	$\rho_{\alpha\eta}(0) \neq 0$ e $\rho_{\alpha\eta}(k) = 0 \quad \forall k \neq 0$
X e Y são independentes	$\rho_{\alpha\eta}(k) = 0 \quad \forall k$

7.3 Testes de Causalidade

i) Teste de Granger

O teste implícito na definição de Granger (1969) consiste em utilizar um modelo bivariado AR(M) de ordem M suficientemente grande para aproximar a representação AR(∞) em (104), isto é,

$$X_t = \sum_{j=1}^M \pi_{11j} X_{t-j} + \sum_{j=1}^M \pi_{12j} Y_{t-j} + \alpha_t, \quad (112)$$

$$Y_t = \sum_{j=1}^M \pi_{21j} X_{t-j} + \sum_{j=1}^M \pi_{22j} Y_{t-j} + \eta_t. \quad (113)$$

Por exemplo, a hipótese nula de que Y não causa X é equivalente a hipótese $\pi_{21j} = 0 \quad j = 1, \dots, M$ na regressão (113). Esta hipótese é testada usando o teste F usual. Analogamente outras hipóteses podem ser testadas.

Este procedimento sofre o defeito de que um grande número de graus de liberdade (observações) são perdidas ao estimar os parâmetros do modelo (112) ou (113) que é usado para aproximar a estrutura do correspondente modelo AR(∞) em (104). Uma alternativa é tentar obter um modelo parcimonioso da forma mista ARMA(p,q) vetorial. A hipótese nula Y não causa X pode ser testada através do teste da razão de verossimilhança da hipótese

$$\phi_{12}(L) = 0 = \theta_{12}(L)$$

em (102), que toma a forma

$$LR = T \log \left\{ \frac{|\hat{\Sigma}_r|}{|\hat{\Sigma}|} \right\} \sim \chi_{p+q}^2 \quad (114)$$

e é assintoticamente distribuída como uma χ^2 com $p+q$ graus de liberdade. Em (114) $\hat{\Sigma}$ e $\hat{\Sigma}_r$ são as matrizes de variância covariância dos resíduos das estimações respectivamente dos modelos completo (102) e restrito (103).

O teste de causalidade instantânea é obtido testando se Σ é diagonal.

ii) Teste de Sims

Este teste é baseado no filtro bilateral (108) e na prática, por exemplo para testar se Y não causa X testa-se $\beta_j = 0$, $j = -1, -2, \dots, -M_1$, através do teste F na regressão

$$Y_t = \sum_{j=-M_1}^{M_2} \beta_j X_{t-j} + u_t, \quad (115)$$

com M_1 e M_2 suficientemente grandes. Entretanto, em geral u_t não é ruído branco e isso distorce os resultados do teste F . Para evitar este efeito Sims (1972) sugeriu utilizar arbitrariamente o filtro $(1 - 0.75L)^2$. Outra alternativa (ver por exemplo Aaker, Carman e Jacobson, 1982) é incluir valores de Y defasados em (109), isto é,

$$Y_t = \sum_{j=-M_1}^{M_2} \beta_j X_{t-j} + \sum_{i=1}^{M_3} \gamma_i Y_{t-i} + \varepsilon_t \quad (116)$$

A escolha dos M_1 em (115) ou (116) é arbitrária e se esses valores são grandes, muitos graus de liberdade se perdem ao estimar os parâmetros.

iii) Teste de Haugh-Pierce

Testes baseados nas correlações cruzadas dos ruídos univariados obtidos de modelos ARMA univariados para as séries X e Y foram desenvolvidos por Haugh (1976) (ver também McLeod, (1979) e Pierce, (1977)).

Seja $r_{\alpha\beta}(k)$ a correlação cruzada amostral, isto é, a estimativa de $\rho_{\alpha\beta}(k)$ em (110). Sob a hipótese de que a série X_t é independente da série Y_t , tem-se que

$$a) r_{\alpha\beta}(k) \sim N\left(0, \frac{1}{T}\right), \quad (117)$$

$$b) Q_M = T^2 \sum_{k=-M_1}^{M_2} \frac{1}{T-k} r_{\alpha\beta}^2(k) \sim \chi^2_{(M_1+M_2+1)} \quad (118)$$

e testes estatísticos sobre causalidade são facilmente construídos. Por exemplo, comparando-se $r_{\alpha\beta}(k)$ com $2/\sqrt{T}$ verifica-se que valor de $r_{\alpha\beta}(k)$ é significativamente diferente de zero. A localização do valor significativo dá uma indicação da causalidade. A estatística Q_M testa $r_{\alpha\beta}(k)$ em grupos e não individualmente.

Comparações de simulações sobre os diversos testes de causalidade são apresentadas em Nelson e Schwert (1982) e Battaglia e Carluci (1982), com resultados conflitantes. Os primeiros autores concluíram que os testes baseados em regressão são preferíveis ao de Haugh-Pierce, entretanto os segundos autores concluem exatamente o contrário.

Finalmente, uma discussão detalhada sobre causalidade pode ser vista em Newbold (1982).

IV - PROGRAMAS DE COMPUTADOR (DO IM E COPPE, UFRJ)

Nome	Origem	Técnica
BAYSEA	Inst. of Statist. Mathematics, Japão	- Decomposição Bayesiana.
DSDM-L1	Université de Montréal, Canadá	- Transformação Geral de Box-Cox em Regressão com heterodasticidade e/ou autocorrelação dos erros.
PACK	Ohio State University; USA	- Box-Jenkins univariado. - Análise de Intervenção. - Função de transferência.
SIBYL/RUNNER	INSEAD - França	- Amortecimentos exponenciais. - Decomposição CENSUS11 - Filtro Adaptativo. - Box-Jenkins univariado
TIMSAC	Inst. of Statist. Mathematics, Japão	- Séries univariadas e multivariadas, diversas análises no domínio do tempo e da frequência para previsão e controle estocástico. - Espaço de estados.
WMTS-1	University of Wisconsin, USA	- Box-Jenkins-Tiao multivariado

- Além destes, a UFRJ tem disponível no NCE/UFRJ os seguintes programas para estatística: SPSS, BMD e IMSL.

- Uma lista sobre programas para previsão e séries temporais é apresentada em

"Rice, G., Mahmoud, E. e Vadivel, S.M.P. - 1984 - *A directory of 132 packages for forecasting and planning*, J. of Business Forecasting, Spring, 1984".

REFERÊNCIAS

- Aaker, D.A., Carman, J.M. e Jacobson, R. - 1982 - *Modeling advertising - sales relationships involving feedback: a time series analysis of six cereals brands*. J. Marketing Res. v. 19 pg. 116-125.
- Abraham, B. - 1980 - *Intervention analysis and multiple time series*. Biometrika v. 67, pg. 73-78.
- Abraham, B. e Ledolter, J. - 1983 - *Statistical Methods for Forecasting*. Wiley.
- Agreen, A. - 1981 - *A survey of some work on ARIMA modelling and intervention analysis*. Research Report 81-7. Department of Statistics University of Uppsala.
- Aigner, D.J. - 1977 - *A compendium on estimation of the autorregressive - moving average model for time series data*. Int. Econ. Rev. 12, pg. 348-371.
- Akaike, H. - 1969 - *Fitting autorregressive models for prediction*. Ann. Inst. Statist. Math. v. 21, pg. 243-247.
- Akaike, H. - 1974 - *Markovian representation of stochastic processes and its application to the analysis of autorregressive moving average processes*. Ann. Inst. Statist. Math. v. 26, pg. 363-387.
- Akaike, H. - 1974 - *Stochastic Theory of minimal realization*. IEEE Transactions on Automatic Control, v. AC-19, pg. 667-674.
- Akaike, H. - 1976 - *Canonical correlation analysis and the use of an information criterion*. Em Advances and Case Studies in System Identification. (Ed. R. Mehra e D.G. Lainiotis), pg. 27-96. Academic Press.

- Akaike, H., - 1978 - *On the identification of state space models and their use in control*. Em *Directions in Time Series* (Ed. D.R. Brillinger e G.C. Tiao). Inst. of Math. Statistics.
- Andel, J., - 1982 - *Fitting models in time series analysis*. Math. Operationsforsch Statist., Ser. Statistics, v. 13, pg. 121-143.
- Anderson, O.D., - 1975 - *Time Series Analysis and Forecasting: The Box-Jenkins Approach*. Butterworths.
- Anderson, O.D., - 1980 - *A new approach to ARMA modelling: some comments*. Em *Analysing Time Series* (Ed. O.D. Anderson), pg. 43-71, North Holland.
- Ansley, G.F. - 1979 - *An algorithm for the exact likelihood of a mixed autorregressive - moving average process*. Biometrika, v. 66, pg. 59-65.
- Bagshaw, M.L. - 1983 - *Extension of Granger causality in multivariate time series models*. Working paper 8303 do Federal Reserve Bank of Cleveland.
- Bard, Y. - 1974 - *Nonlinear Parameter Estimations*. Academic Press.
- Bartlett, M.S., - 1955 - *Stochastic Processes*. Cambridge University Press.
- Basilevsky, A., e Hum, D.P.J., - 1979 - *Karhunen-Loeve analysis of historical time series with an application to plantation birth in Jamaica*. J. Am. Statist. Ass., v. 74, pg. 284-290.
- Battaglia, F., e Carlucci, F., - 1982 - *A simulation study of causality tests*. Em *Time Series Analysis Theory & Practice 1* (Ed. O.D. Anderson), pg. 655-669, North Holland.
- Bell, W.R., e Hillmer, S.C., - 1983 - *Modeling time series with calendar variation*. J. Am. Statist. Ass., v. 78, pg. 526-534.

- Bequin, J.M., Gourieroux, C., e Monfort, A., - 1980 - *Identification of a mixed autorregressive - moving average process: the corner method*. Em Time Series (Ed. O.D. Anderson), pg. 423-435, North Holland.
- Bequin, J.M., Gourieroux, C., e Monfort, A., - 1981 - *The applicability of the corner method: a reply*. J. Opl. Res. Soc., v. 32, pg. 1042-1045.
- Beveridge, S., - 1981 - *Autorregressive predictors: a survey of the evidence and additional results*. Manuscripto - Faculty of Business Administration. University of Alberta.
- Bhansali, R.J., - 1980 - *Autorregressive and window estimates of the inverse correlation*. Biometrika, 67, pg. 551-566.
- Bhansali, R.J., - 1983 - *The inverse partial correlation function of a time series and its application*. J. of Multivariate An. v. 13, pg. 310-327.
- Bhattacharyya, M.N., - 1979 - *Forecasting the demand for telephones in Australia*. Appl. Statist., v. 23, pg. 1-10.
- Bhattacharyya, M.N., e Layton, A.P., - 1979 - *Effectiveness of seat belt legislation on the Queensland road toll - An Australian case study in intervention analysis*. J. Am. Statist. Ass., v. 74, pg. 596-603.
- Bhattacharyya, M.N., - 1982 - *Lydia Pinkham data remodelled*. J. Time Series An., v. 3, pg. 81-102.
- Box, G.E.P., e Jenkins, G.M., - 1976 - *Time Series Analysis, Forecasting and Control* (2ª Ed.). Holden Day.
- Box, G.E.P., e Newbold, P., - 1971 - *Some comments on a paper of Coen, Gomme and Kendall*. J.R. Statist. Soc. A, v. 134, pg. 229-240.
- Box, G.E.P., e Pierce, D.A., - 1970 - *Distribution of residual autocorrelations in autorregressive - integrated moving average time series models*. J. Am. Statist. Ass., v. 65, pg. 1509-1529.

- Box, G.E.P., e Tiao, G.C., - 1975 - *Intervention analysis with applications to economic and environmental problem*. J. Am. Statist. Ass., v. 70, pg. 70-79.
- Box, G.E.P., e Tiao, G.C., - 1976 - *Comparison of forecast and actuality*. Appl. Statist., v. 25, pg. 195-200.
- Box, G.E.P., e Tiao, G.C., - 1977 - *A canonical analysis of multiple time series*. Biometrika, v. 64, pg. 355-365.
- Brewer, K.R.W., - 1973 - *Some consequences of temporal agregation and systematic sampling for ARMA and ARMAX models*. J. of Econometrics, v. 1, pg. 133-154.
- Burch, P.R.J., - 1978 - *Smoking and lung cancer: the problem of inferring cause (with discussion)*. J. Roy. Statist. Soc. A, v. 41, pg. 437-477.
- Camacho, F., e McLeod, A.I., - 1983 - *Multivariate time series modelling using the WMTS - 1 package*. TR-83-02 Department of Statistical and Actuarial Sciences. University of Western Ontario.
- Chatfield, C., - 1975 - *The Analysis of Time Series: Theory and Practice*. Chapman and Hall.
- Chatfield, C., - 1979 - *Inverse autocorrelations*. J. R. Statist. Soc. A, v. 142, pg. 363-377.
- Clarke, B.R., e Godolphin, E.J., - 1982 - *Comparative power studies for goodness of fit tests of time series models*. J. Time Series An., v. 3, pg. 141-151.
- Cleveland, W.P., - 1972 - *The inverse autocorrelation of a time series and their applications*. Technometrics, 14, pg. 277-298.
- Cleveland, W.P., e Devlin, S.J., - 1980 - *Callendar effects in monthly time series: detection by spectrum analysis and graphical methods*. J. Am. Statist. Ass., v. 75, pg. 487-496.

- Cleveland, W.P., e Grupe, M.R., - 1981 - *Modeling time series when calendar effects are present*. Manuscrito Apresentado na ASA-CENSUS-NBER Conference on Applied Time Series Analysis of Economic Data. Outubro.
- Cleveland, W.P., e Tiao, G.C., - 1976 - *Decomposition of seasonal time series: a model for the Census X11 program*. J. Am. Statist. Ass., v. 71, pg. 581-587.
- Conradsen, K., e Spliid, H., - 1981 - *A seasonal adjustment filter for use in Box-Jenkins analysis of seasonal time series*. App. Statist., v. 30, pg. 172-177.
- Cooper, D.M., e Wood, E.F., - 1982a - *Identifying multivariate time series models*. J. Time Series An., v. 3, pg. 153-164.
- Cooper, D.M., e Wood, E.F., - 1982b - *Identification of multivariate time series and multivariate input-output models*. Water Resources Res., v. 18, pg. 937-946.
- Cooper, D.M., e Wood, E.F., - 1982c - *Parameter estimation of multiple input-output time series models: application to rainfall-runoff process*. Water Resources Res., v. 18, pg. 1352-1364.
- Cramer, H., e Leadbetter, M.R., - 1966 - *Stationary and Related Stochastic Processes: Sample Function Properties and their Applications*. John Wiley.
- Crookes, J.G., Croston, K., e Sypsas, P., - 1980 - *Process components for multivariate time series analysis*. J. Op. Res. Soc., v. 31, pg. 325-330.
- Damsleth, E., - 1983 - *Estimation in moving average models, why does it fail?*. J. Statist. Comp. Simul., v. 16, pg. 109-128.
- De Jong, P., - 1976 - *The recursive fitting of autorregressions*. Biometrika, 63, pg. 525.

- Dent, W.T., e Min, A.S., - 1978 - *A monte Carlo Study of autorregressive integrated moving average processes*. J. of Econometrics, v. 7, pg. 23-55.
- Dent, W.T., - 1977 - *Computation of the exact likelihood function of an ARIMA process*. J. Statist. Comp. Simul., v. 5, pg. 193-206.
- Durbin, J., - 1959 - *Efficient estimation of parameters in moving average models*. Biometrika, v. 46, pg. 306-316.
- Durbin, J., - 1965 - *The fitting of time series models*. Rev. Int. Statist. Inst., v. 28, pg. 233-244.
- Fukunaga, K., - 1972 - *Introduction to Statistical Pattern Recognition*. Academic Press.
- Gardner, G., Harvey, A.C., e Phillips, G.B.A., - 1980 - *An algorithm for the exact maximum likelihood estimation by means of Kalman filtering*. App. Statist., vol. 29, pg. 311-322.
- Glasbey, C.A., - 1982 - *A generalization of partial autocorrelations useful in identifying ARMA models*. Technometrics, v. 24, pg. 223-228.
- Gooijer, J.G., e Heuts, R.M.J., - 1981 - *The corner method: an investigation of an order discrimination procedure for general ARMA processes*. J. Opl. Res. Soc., v. 32, pg. 1039-1042.
- Granger, C.W.J., - 1969 - *Investigating causal relations by econometric models and cross-spectral methods*. Econometrica, v. 17, pg. 424-438.
- Granger, C.W.J., - 1982 - *Acronyms in time series analysis (ATSA)*. - J. Time Series An., v.3, pg. 103-107.

- Granger, C.W.J., e Morris, M.J., - 1976 - *Time series modelling and interpretation*. J.R. Statist. Soc. A, v. 139, pg. 246-257.
- Granger, C.W.J., e Newbold, P., - 1974 - *Spurious regressions in econometrics*. J. of Econometrics, v. 2, pg. 111-120.
- Granger, C.W.J., e Newbold, P., - 1977 - *Forecasting Economic Time Series*. Academic Press.
- Gray, H.L., Kelley, G.D., e McIntire, D.D., - 1978 - *A new approach to ARMA modelling*. (with discussion). Comm. Statist. - Simula. Computa. B 7, pg. 1-96.
- Gray, H.L., e Woodward, M.A., - 1981 - *Application of S arrays to seasonal data*. Am Applied Time Series II (Ed. D.F. Findley). Academic Press, pg. 379-413.
- Hamilton, D.C., e Watts, D.G., - 1978 - *Interpreting partial autocorrelation functions of seasonal time series models*. Biometrika, v. 65, pg. 135-140.
- Hannan, E.J., - 1969 - *The identification of vector mixed autorregressive moving average systems*. Biometrika, v.56, pg. 751-765.
- Hannan, E.J., - 1970 - *Multiple Time Series*. Wiley.
- Hannan, E.J., - 1971 - *The identification problem for multiple equation systems with moving average erros*. Econometrica, v. 39, pg. 751-765.
- Hannan, E.J., - 1980 - *The estimation of the order of an ARMA process*. Ann. Statist., v. 8, pg. 1071-1081.
- Hannan, E.J., e Kavalieris, L., - 1983 - *Multivariate linear time series models*. Manuscrito. Australian National University (107 páginas). A ser publicado em Adv. App. Probab.
- Hannan, E.J., e Rissanen, J., - 1982 - *Recursive estimation of ARMA order*. Biometrika, v. 69, pg. 81-94.

Harvey, A.C., - 1981 - *Time Series Models*. Philip Alan.

Harvey, A.C., - 1982a - *Estimation procedure for a class of univariate time series models*. Discussion Paper n° 28. London School of Economics Econometric Programme.

Harvey, A.C., - 1982b - *An alternative framework for time series model building and its implication for econometrics*. Discussion Paper n° 32. London School of Economics Econometric Programme.

Harvey, E.C., - 1983 - *A unified view of statistical forecasting procedure*. A ser publicado no J. of Bus. Econ. Statist.

Harvey, A.C., e Phillips, G.D.A., - 1976 - *The maximum likelihood estimation of autorregressive - moving average models by Kalman filtering*. Manuscrito. University of Kent.

Harvey, A.C., e Phillips, G.D.A., - 1979 - *Maximum likelihood estimation of regression models with autorregressive - moving average disturbances*. Biometrika, v. 66, pg. 49-58.

Harvey, A.C., e Pierse, R.G., - 1982 - *Estimating missing observations in economic time series*. Discussion Paper n° A 33. London School of Economics Econometric Programme.

Haugh, L.D., - 1976 - *Checking the independence of two covarian stationary time series: a univariate residual cross-correlation approach*. J. Am. Statist. Ass., v. 71, pg. 378-385.

Haugh, L.D., e Box, G.E.P., - 1977 - *Identification of dynamic regression (distributed lag) models connecting two time series*. J. Am. Statist. Ass., v. 72, pg. 121-130.

- Heuts, R.M.J., e Rens, P.J., - 1980 - *Theory of multivariate transfer function noise models and its application on a simulated structure (multiple input - multiple output models, with a couple noise structure)*. Research Memorandum FEW 95. Department of Economics, Subfaculteit der Econometrie, Tilburg University.
- Heuts, R.M.J., e Willemse, R., - 1981 - *Impulse response patterns for various dinamic time series models*. Reek Ter Discussie n. 81.0.8. Subfaculteit der Econometrie. Katholieke Hogeschool Tilburg.
- Heyse, J.F., e Wey, W.W.S., - 1983 - *Modeling the advertising sales relationship through use of multiple time series techniques*. Presented at 3th International Forecasting Symposium on Forecasting, Philadelphia, June.
- Hillmer, S.C., - 1982 - *Forecasting time series with trading day variation*. J. of Forecasting, v. 1, pg. 385-395.
- Hillmer, S.C., e Tiao, G.C., - 1979 - *Likelihood function of stationary multiple autorregressive - moving average models*. J. Am. Statist. Ass., v. 74, pg. 652-660.
- Hipel, K.W., - 1981 - *Geophysical model discrimination using the Akaike information criterion*. IEEE Trans. Automat. Contr., v. AC 26, pg. 358-377.
- Hipel, K.W., Lennox, W.C., Unny, T.E., e McLeod, A.I., - 1975 - *Intervention analysis in water resources*. Water Resources Res., v. 11, pg. 855-861.
- Hipel, K.W., Li, W.K., e McLeod, A.I., - 1983 - *Causal and dynamic relationships between natural phenomena*. Manuscripto. Department of System Design Engineering, University of Waterloo.
- Hipel, K.W., McLeod, A.I., e Lennox, W.C., - 1977 - *Advances in Box-Jenkins modeling. 1 Model construction*. Water Resources Res., v. 13, pg. 567-575.

- Hipel, K.W., McLeod, A.I., e Noakes, D.J., - 1982 - *Fitting dynamic models to hydrological time series*. Em *Time Series Methods in Hydrosociences* (Ed. A.H. El-Shaarawi e S.R. Esterby), pg. 110-129. Elsevier.
- Hosking, J.R.M., - 1980 - *The multivariate portmanteau statistic*. *J. Am. Statist. Ass.*, v. 75, pg. 602-607.
- Hosking, J.R.M., - 1980 - *The asymptotic distribution of the sample inverse autocorrelations of an autorregressive - moving average process*. *Biometrika*, v. 67, pg. 223-226.
- Hosking, J.R.M., - 1981 - *Lagrange - multiplier tests of multivariate time series models*. *J. Roy. Statist. Soc. B*, v. 43, pg. 219-230.
- Hoyos, A. de, - 1980 - *Processos Estocásticos e Previsão*. Curso do 4º Simpósio nacional de Probabilidade e Estatística, Julho, RJ.
- Hsiao, C., - 1982 - *Time series modelling and causal ordering of Canadian money, income and interest rate*. Em *Time Series Analysis Theory and Practice 1*. (Ed. O.D. Anderson), pg. 671-699. North Holland.
- Hui, B., e O Reilly, D., - 1982 - *State vector analysis*. *The DRI Statistician*, v. 1, pg. 11-23. Data Resources Inc.
- Indejehagopian, J.P., - 1979 - *Les processus stochastiques vectoriels ARMA: une procédure d'identification*. *Revue de statistique Appliquée*, v. 27, pg. 33-45.
- Izenman, A.J., - 1983 - *J. R. Wolf e H.A. Wolfer: On historical note on the Zurich sunspot relative numbers*. *J. R. Statist. Soc. A*, v. 146, pg. 311-315.
- Jenkins, G.M., - 1979 - *Practical Experience with Modelling and Forecasting Time Series*. Gwilym Jenkins e Partners.

- Jenkins, G.M., e Alavi, A.S., - 1981 - *Some aspects of modelling and forecasting multivariate time series*. J. Time Series An., v. 2, pg. 1-47.
- Jenkins, G.M., e McLeod, C., - 1982 - *Case Studies in Time Series Analysis*. v. 1. Gwilym Jenkins e Partners.
- Kendall, M.G., - 1973 - *Time Series*. Griffin.
- Kendall, M., Stuart, A., e Ord, J.K., - 1983 - *The Advanced Theory of Statistics*. (4ª Edição). Griffin.
- Kenny, P.B., e Durbin, J., - 1982 - *Local trend and seasonal adjustments of economic and social time series*. (with discussion). J. R. Statist. Soc. A, v. 145, pg. 1-41.
- Koroliuk, V.S., - 1981 - *Manual de La Teoría de Probabilidades y Estadística Matemática*. Mir.
- Lanna, A.E., e Indursky, A.B., - 1983 - *Predição de Vazões médias mensais com modelos empíricos*. Rev. Bras. Eng.. Caderno de Recursos Hídricos, v. 1, pg. 59-90.
- Ledolter, J., - 1978 - *The analysis of multivariate time series applied to problems in hydrology*. J. of Hydrology, v. 36, pg. 327-352.
- Ledolter, J., e Abraham, B., - 1981 - *Parsimony and its importance in time series forecasting*. Technometrics, v. 23, pg. 411-414.
- Lettenmaier, D.P., - 1980 - *Intervention analysis with missing data*. Water Resources Res., v. 16, pg. 159-171.
- Li, W.K., e McLeod, A.I., - 1981 - *Distribution of the residual autocorrelation in multivariate ARMA time series models*. J. Roy. Statist. Soc. B, v. 43, pg. 231-239.
- Liu, L.M., e Hanssens, D.M. - 1982 - *Identification of multiple-input transfer function models*. Comm. Statist. - Theor. Meth., v. 11, pg. 297-314.

- Ljung, G.M., e Box, G.E.P., - 1978 - *On a measure of lack of fit in time series models*. *Biometrika*, v. 65, pg. 297-303.
- Ljung, G.M., e Box, G.E.P., - 1979 - *The likelihood function of stationary autorregressive - moving average models*. *Biometrika*, v. 66, pg. 265-270.
- Lutkepohl, H., - 1982 - *Non-causality due to omitted variables*. *J. of Econometrics*, v. 19, pg. 367-378.
- Maravall, A., - 1981 - *A note on identification of multivariate time series models*. *J. of Econometrics*, v. 16, pg. 237-247.
- McGregor, J.F., - 1972 - *Topics in control I: State variable approach to time series representation and forecasting*. Technical Report 306. Department of Statistics University of Wisconsin - Madison.
- McLeod, A.I., - 1977 - *Improved Box-Jenkins estimators*. *Biometrika*, v. 64, pg. 531-534.
- McLeod, A.I., - 1979 - *Distribution of the residual cross-correlation in univariate ARMA time series models*. *J. Am. Statist. Assoc.*, v. 74, pg. 849-855.
- McLeod, A.I., Hipel, K.W., e Lennox, W.C., - 1977 - *Advances in Box-Jenkins modeling*. 2 Applications *Water Resources Res.*, v. 13, pg. 577-586.
- McLeod, A.I., e Li, W.K., - 1983 - *Diagnostic checking ARMA time series models using squared - residual autocorrelations*. *J. Time Series An.*, v. 4, pg. 269-273.
- Morettin, P.A., - 1979 - *Análise Harmônica de Processos Estocásticos*. Curso do 12º Colóquio Brasileiro de Matemática, julho, Poços de Caldas, MG.
- Morettin, P.A. - 1981 - *Modelos para Previsão de Séries Temporais*. v. 1 e 2. Curso do 13º Colóquio Brasileiro de Matemática, julho, Poços de Caldas, MG.

- Moriarty, M., e Salamon, G., - 1980 - *Estimation and forecast performance of a multivariate time series model of sales*. J. of Marketing Res., v. 17, pg. 558-564.
- Nelson, C.R., - 1973 - *Applied Time Series Analysis for Managerial Forecasting*. Holding-Day.
- Nelson, C.R., e Schwartz, G.W., - 1982 - *Tests for predicture relationship between time series variables: a Monte Carlo investigation*. J. Am. Statist. Assoc., v. 77, pg. 11-18.
- Nelson, H.L., e Granger, C.W.J., - 1979 - *Experience with using the Box-Cox transformation when forecasting economic time series*. J. of Econometrics, 10, pg. 57-69.
- Nerlove, M., Grether, D.M., e Carvalho, J.L. - 1979 - *Analysis of Economic Time Series: A Synthesis*. Academic Press.
- Newbold, P., - 1974 - *The exact likelihood function for a mixed autorregressive - moving average process*. Biometrika, v. 61, pg. 423-426.
- Newbold, P., - 1982 - *Model checking in time series analysis*. Proceedings of ASA-CENSUS-NBER Conference on Applied Time Series Analysis of Economic Data. (Ed. A. Zellner).
- Newbold, P., - 1982 - *Causality testing in economics*. Em Time Series Analysis, Theory and Practice. (Ed. O.D. Anderson), pg. 701-716. North Holland.
- Newbold, P., e Bos, T., - 1983 - *On q-conditioned partial correlations*. J. Time Series An., v. 4, pg. 53-55.
- Nicholls, D.F., e Hall, A.D. - 1979 - *The exact likelihood function of multivariate autorregressive - moving average models*. Biometrika, v. 66, pg. 259-264.
- Ord, J.K., - 1982 - *An alternative approach to the specification of multiple time series models*. Em Time Series Analysis Theory and Practice 1. (Ed. O.D. Anderson), pg. 85-103. North Holland.

- Otter, P.W., e Schuur, J.F., - 1982 - *Principal component analysis in multivariate forecasting of economic time series*. Em *Time Series Analysis Theory and Practice*. (Ed. O.D. Anderson), pg. 323-333. North Holland.
- Pandit, S.M., 1973 - *Data Dependent Systems: Modeling, Analysis and Optimal Control via Time Series*. Ph.D. Thesis, University of Wisconsin, University Microfilms International.
- Pandit, S.M., e Wu, S.M., - 1983 - *Time Series and System Analysis with Applications*. Wiley.
- Parzen, E., - 1979 - *Time series and whitening filter estimation*. Em *TIMS studies in Management Sciences*. (Ed. Makridakis), v. 12, pg. 149-165.
- Parzen, E., - 1980a - *Time series modelling, spectral analysis and forecasting*. Em *Directions in Time Series* (Ed. D. R. Brillinger e G.C. Tiao), pg. 80-111. Inst. of Math. Statist.
- Parzen, E., - 1980b - *Time series forecasting by ARARMA models*. Manuscrito. Technical Report, nº 14. Inst. of Statist. Texas A. M. University.
- Parzen, E., - 1982 - *ARARMA models for time series analysis and forecasting*. *J. of Forecasting*, v. 1, pg. 67-82.
- Pearlman, J.G., - 1980 - *An algorithm for the exact likelihood of a high-order autorregressive - moving average process*. *Biometrika*, v. 67, pg. 232-233.
- Pereira, B. de B., - 1980 - *Tópicos em Séries Temporais: Métodos Automáticos de Previsão*. Publicação PDD 07/80, COPPE - UFRJ.
- Pereira, B. de B., Coqueiro, R.C.O., e Perrota, A.H.V., - 1983 - *Combinação de informação subjetivas e métodos quantitativos, para previsão de taxas no open-market*. *Pesquisa Operacional*, v. 3, pg. 25-40

Pereira, B. de B., Zanusso Pais, M.B., e Salles, P.R.H., -
1984 - *Introdução à Análise Espectral de Séries Temporais*.
IM-UFRJ. (a ser publicado).

Petto, A.C., - 1983 - *Forecasting accuracy of state-space, Box-Jenkins and deterministic methods for Lydia Pinkham company sales*. Presented at 3th International Symposium on Forecasting, Philadelphia, June.

Pierce, D.A., - 1975 - *Forecasting in dynamic models with stochastic regressions*. J. of Econometrics, v. 3 , pg. 349-374.

Pierce, D.A., - 1977 - *Relationship and the lack thereof - between economic time series, with special reference to money and interest rates*. J. Am. Statist. Ass., v. 72, pg. 11-26.

Pierce, D.A. - 1980 - *Multivariate Time Series*. Escola Latino Americana de Matemática - julho, Mar del Plata.

Pierce, D.A., e Haugh, L.D., - 1977 - *Causality in temporal systems: characterization and a survey*. J. of Econometrics, v. 5, pg. 265-293.

Pierce, D.A., e Haugh, L.D., - 1979 - *The characterization of instantaneous causality: a comment*. J. of Econometrics, v. 10, pg. 257-259.

Pino, F.A., e Morettin, P.A. - 1980 - *Análise de intervenção em séries temporais - aplicações à economia agrícola*. Apresentado eno 4º Simpósio Nacional de Probabilidade e Estatística. Julho, RJ.

Poskitt, D.S., e Tremayne, A.R., - 1980 - *Testing the specification of fitted autorregressive - moving average model*. Biometrika, v. 67, pg. 359-363.

Poskitt, D.S., e Tremayne, A.R., - 1983a - *Model selection and diagnostic checking in univariate time series analysis*. Manuscrito. Department of Economic and Related Studies, University of York.

- Poskitt, D.S., e Tremayne, A.R., - 1983b - *On the posterior odds of time series models*. *Biometrika*, v. 76, pg. 157-162.
- Price, J.M., - 1979 - *The characterization of instantaneous causality: a correction*. *J. of Econometrics*, v. 10, pg. 253-256.
- Priestley, M.B., - 1980 - *System identification, Kalman filtering and stochastic control*. Em *Directions in Time Series*. (Ed. D.R. Brillinger e G.T. Tiao), pg. 188-218. *Inst. of Math. Statistics*.
- Priestley, M.B., - 1981 - *Spectral Analysis and Time Series*. v. 1 e 2, Academic Press.
- Prins, J., - 1980 - *Interactive multivariate time series*. Em *Analysis Time Series*. (Ed. O.D. Anderson), pg. 269-289, North Holland.
- Pukkila, T., - 1978 - *A sample estimation of a regression model with autorregressive moving average errors*. Report A 30, Department of Mathematical Sciences, University of Tampere.
- Quenouille, M.H., - 1957 - *The Analysis of Multiple Time Series*. Griffin.
- Rissanen, J., - 1984 - *Information and accumulated prediction error criteria for order estimation*. Manuscrito. IBM, California.
- Röse, D.A., - 1977 - *Forecasting aggregates of independent ARIMA process*. *J. of Econometrics*, v. 5, pg. 323-345.
- Saboia, J.L.M., - 1976 - *Mortalidade infantil e salário mínimo - uma análise de intervenção para o município de São Paulo*. *Revista Brasileira de Administração*, v. 16, pg. 47-50.
- Salas, J.D., e Obeysekera, J.T.B., - 1982 - *ARMA model identification of hydrologic time series*. *Water Resour Res.*, v. 18, pg. 1011-1021.

- Sims, C.A., - 1972 - *Money, income and causality*. American Economic Review, v. 62, pg. 540-552.
- Solo, V., - 1982 - *Topics in Advanced Time Series*. Curso da 1ª Escola de Inverno de Probabilidade e Estatística de Santiago, julho, Santiago. (A ser publicado pela Springer Verlag, Lectures Notes).
- Solo, V., - 1982 - *ARMA models without MA parameters*. Manuscrito. Harvard University.
- Solo, V., - 1983 - *The exact likelihood for a multivariate ARMA model*. Manuscrito. Harvard University.
- Terasvirta, T., - 1982 - *Mink and muskrat interaction: a structural analysis*. Research Report nº 34. Department of Statistics. University of Helsinki.
- Thury, G., - 1980 - *Time series analysis of the relationship between private consumer expenditure and disposable income for the case of Austria*. Empirica, v. 2,80, pg. 169-198.
- Thury, G., - 1982 - *Modelling private consumer expenditure in Austria by intervention analysis*. Em Time Series Analysis: Theory and Practice 1. (Ed. O.D. Anderson),
- Thury, G., - 1983 - *Applied multiple time series: a case study for Austria*. Presented at 9th International Time Series Meeting, Nottingham, April.
- Tiao, G.C., e Box, G.E.P., - 1981 - *Modelling multiple time series with applications*. J. Am. Statist. Assoc., v. 76, pg. 802-816.
- Tiao, G.C., e Guttman, I., - 1980 - *Forecasting contemporaneous aggregates of multiple time series*. J. of Econometrics, v. 12, pg. 219-230.

- Tiao, G.C., e Tsay, R.S., - 1981 - *Identification of nonstationary and stationary ARMA models*. Technical Report 647. Department of Statistics, University of Wisconsin, Madison.
- Tiao, G.C., e Tsay, R.S., - 1982 - *Consistency properties of least squares estimates of autorregressive parameters in ARMA models*. Technical Report 658. Department of Statistics. University of Wisconsin, Madison.
- Tiao, G.C., e Tsay, R.S., - 1983 - *Multiple time series modelling and extended sample cross-correlations*. J. Bus. Econ. Statist., v. 1, pg. 43-56.
- Tjostheim, D., - 1981 - *Granger causality in multiple time series*. J. of Econometrics, v. 17, pg. 157-176.
- Tsay, R.S., e Tiao, G.C., - 1982 - *Consistent estimates of autorregressive parameters and extended sample autocorrelation function for stationary and nonstationary ARMA models*. Technical Report 683. Department of Statistics, University of Wisconsin, Madison.
- Tucker, W.T., - 1982 - *On the Padé table and its relationship to the R and S arrays and ARMA modelling*. Comm. Statist. - Theor Meth., v. 11, pg. 1335-1379.
- Umashankar, S., e Ledolter, J., - 1983 - *Forecasting with diagonal multiple time series models: an extension of univariate models*. J. of Marketing Res., v. 20, pg. 58-63.
- Vishwarkarma, K.P., - 1983 - *The Box-Jenkins (ARIMA) method is a classic example of curing the disease by killing the patient*. Apresentado no Third International Symposium of Forecasting, Philadelphia.
- Wold, H., - 1965 - *A Graphic introduction to stochastic processes*. Em Bibliography on Time Series and Stochastic Processes. Oliver and Boyd.

- Woodward, W.A., e Gray, H.L., - 1978 - *New ARMA models for Wolfer's sunspot data*. Comm. Statist. Simula - Computa. B7, pg. 97-115.
- Woodward, W.A., e Gray, H.L., - 1981 - *On the relationship between the S array: and the Box-Jenkins method of ARMA model identification*. J. Am. Statist. Ass., v. 76, pg. 579-587.
- Yaglom, A.M., - 1962 - *An introduction to the theory of Stationary Random Functions*. Prentice Hall.
- Yajima, Y., - 1982 - *Estimation of the number of differencing of an ARIMA process by an AR model fitting*. Research Report (B-108) on Information Sciences, Tokyo Institute of Technology.