

Universidade Federal de Campina Grande
Centro de Ciências e Teconologia
Programa de Pós-Graduação em Matemática
Curso de Mestrado em Matemática

Estudo de Modelos ARIMA com Variáveis Angulares para Utilização na Perfuração de Poços Petrolíferos

por

Areli Mesquita da Silva

sob orientação do

Prof. Dr. Francisco Antônio Morais de Souza

Dissertação apresentada ao Corpo Docente do Programa de Pós-Graduação em Matemática - CCT - UFCG, como requisito parcial para obtenção do título de Mestre em Matemática.

Campina Grande - PB

Julho/2007

Estudo de Modelos ARIMA com Variáveis Angulares para Utilização na Perfuração de Poços Petrolíferos

por

Areli Mesquita da Silva

Dissertação apresentada ao Corpo Docente do Programa de Pós-Graduação em Matemática - CCT - UFCG, como requisito parcial para obtenção do título de Mestre em Matemática.

Área de Concentração: Probabilidade e Estatística

Aprovada por:

Prof. Dr. André Gustavo Campos Pereira

Prof. Dr. Antonio José da Silva

Prof. Dr. Francisco Antônio Morais de Souza
Orientador

Universidade Federal de Campina Grande
Centro de Ciências e Tecnologia
Programa de Pós-Graduação em Matemática
Curso de Mestrado em Matemática

Julho/2007

Agradecimentos

A Deus por mais essa dádiva em minha vida!

A meus pais, Eri e Manoel, pelo investimento e incentivo dados em todos os momentos.

A Fúlvio (Vinho) pelo apoio, companheirismo e por sempre procurar deixar meu ego nas alturas!

Ao meu orientador, professor Francisco Antônio Morais de Souza, por todos os ensinamentos, pacientemente, compartilhados, sem os quais, teria sido inviável desenvolver este trabalho.

À ANP (Agência Nacional do Petróleo, Gás e Biocombustíveis) e aos demais órgãos financiadores pela concessão da bolsa.

Aos professores André Gustavo Campos Pereira e Antonio José da Silva por terem aceito participar da banca.

Ao professor Brandão por suas brilhantes sugestões (veja como o Apêndice B ficou lindo!).

A todos os professores de graduação e pós-graduação da UAME/UFCG que estiveram sempre na torcida!

A todos os funcionários da UAME/UFCG que nunca economizam esforços na hora de ajudar!

A Joelma (Joca) por acreditar que, um dia, daria tudo certo...(nunca esquecerei da pergunta: *E a integral?*).

A Cris, Grayci (minhas irmãs acadêmicas), a Rosângela (Rosinha), Tatiana (Chaty), Juliana, Jacqueline, Hallyson, Jesualdo (Nash), Joseane, Leomaques,..., pelo carinho e convivência.

A todos que, com simples gestos, contribuíram para que este trabalho fosse concluído.

Dedicatória

A minha família.

“A grandeza de um ser humano não está no quanto ele sabe, mas no quanto ele tem consciência que não sabe. O destino não é freqüentemente inevitável, mas uma questão de escolha. Quem faz escolha, escreve sua própria história, constrói seus próprios caminhos.”

Augusto Cury

Resumo

Séries temporais envolvendo dados angulares aparecem nas mais diversas áreas do conhecimento. Por exemplo, na perfuração de um poço petrolífero direcional, o deslocamento da broca de perfuração, ao longo da trajetória do poço, pode ser considerado uma realização de uma série temporal de dados angulares. Um dos interesses, neste contexto, consiste em realizar previsões de posicionamentos futuros da broca de perfuração, as quais darão mais apoio ao engenheiro de petróleo na tomada de decisão de quando e como interferir na trajetória de um poço, de modo que este siga o curso planejado. Neste trabalho, estudamos algumas classes de modelos que podem ser utilizados para a modelagem desse tipo de série.

Abstract

Time series involving angular data appear in many diverse areas of scientific knowledge. For example, in the drilling of a directional oil well, the displacement of the drill, along the path of the well, can be considered as an angular data time series. One of the objectives, in this context, consists in carrying out forecasts of the future positions of the drill, which will give more support to the petroleum engineer in the decision-making of when and how interfere in the path of a well, so that this follows the planned course. In this work, we study some classes of models that can be utilized for the modeling of that kind of series.

Sumário

Introdução	1
1 Quantificação de Incertezas de Subsuperfície	4
1.1 Incertezas de Subsuperfície	5
1.2 Impacto de Incertezas Dinâmicas sobre um Programa de Perfuração . .	6
1.3 Justificativa da Aquisição de Dados Complementares	7
2 Séries Temporais	9
2.1 Modelos Lineares	14
2.1.1 Modelos Auto-Regressivos	16
2.1.2 Modelos de Médias Móveis	18
2.1.3 Modelos Auto-Regressivos e de Médias Móveis	19
2.1.4 Modelos Auto-Regressivos Integrados e de Médias Móveis	20
2.2 A Função de Autocorrelação Parcial	22
2.3 Alguns Casos Particulares de Modelos Lineares	24
2.3.1 Modelo Auto-Regressivo de Ordem 1 - AR(1)	24
2.3.2 Modelo Auto-Regressivo de Ordem 2 - AR(2)	24
2.3.3 Modelo de Médias Móveis de Ordem 1 - MA(1)	26
2.3.4 Modelo de Médias Móveis de Ordem 2 - MA(2)	26
2.3.5 Modelo Auto-Regressivo e de Médias Móveis de Ordem (1,1) - ARMA(1,1)	28
2.4 Identificação de Modelos ARIMA	29
2.4.1 Procedimentos de Identificação	29
2.4.2 Estimativas Preliminares	34

2.5	Estimação de Modelos ARIMA	35
2.5.1	Método dos Momentos	35
2.5.2	Método de Máxima Verossimilhança	37
2.5.3	Variância dos Estimadores	40
2.6	Diagnóstico de Modelos ARIMA	41
2.6.1	Teste de Autocorrelação Residual	41
2.6.2	Teste de Box-Pierce	42
2.6.3	Teste da Autocorrelação Cruzada	42
2.7	Previsão com Modelos ARIMA	43
2.7.1	Previsão de Erro Quadrático Médio (EQM) mínimo	44
2.7.2	Formas Básicas de Previsão	45
2.7.3	Equação de Previsão	46
2.7.4	Atualização das Previsões	47
2.7.5	Intervalos de Confiança	47
3	Séries Temporais Envolvendo Dados Angulares	49
3.1	Modelos	49
3.1.1	Processo Gaussiano Transformado	50
3.1.2	Processo Arqueado	51
3.1.3	Processos Baseados em Funções de Ligação	52
3.2	Seleção do Modelo	53
3.3	Identificação do Modelo	54
3.4	Ajuste do Modelo	55
3.4.1	Modelo Gaussiano Arqueado	55
3.4.2	Modelo Gaussiano Transformado	56
3.4.3	Modelos com Ligação Direta e Inversa	57
	Conclusão	58
	Apêndices	59
A	Demonstração da Desigualdade (3.1)	59
A.1	Resultados Utilizados	59
A.2	Demonstração da Desigualdade (3.1)	59

	iii
B Demonstração do Teorema (3.1)	61
B.1 Definições	61
B.2 Propriedades	62
B.3 Demonstração do Teorema (3.1)	62
Bibliografia	71

Introdução

A perfuração de poços direcionais na indústria de petróleo, é uma técnica utilizada de forma cada vez mais freqüente, tanto para atingir formações produtoras situadas abaixo de locações verticalmente inacessíveis, como também para perfurar vários poços a partir de um mesmo ponto [Thomas, 2001]. Sua utilização se dá, em particular, em poços *offshore*.

A primeira etapa no projeto de um poço direcional é determinar o tipo de trajetória a ser seguida para se atingir o alvo desejado, que pode ser uma formação com acúmulo de hidrocarbonetos. Nessa etapa são levados em consideração os seguintes elementos:

- A profundidade do(s) ponto(s) de mudança de trajetória;
- O afastamento horizontal;
- A direção/locação do objetivo;
- A profundidade vertical final do poço;
- As inclinações dos diversos trechos.

A mudança de orientação da trajetória do poço é uma operação dispendiosa que envolve a retirada da coluna de perfuração e a introdução de uma ferramenta especial contendo um motor de fundo [Lima], que tem a finalidade de iniciar a deflexão do poço e orientá-lo para a direção desejada. Feita a deflexão, a ferramenta com o motor de fundo é retirada e retorna-se com a coluna normal de perfuração, continuando até um próximo desvio ou até atingir o alvo desejado (formação com acúmulo de hidrocarbonetos).

Sob o ponto de vista operacional, em cada mudança da direção do poço, a sua nova orientação é feita a partir de informações obtidas em superfície, sobre a inclinação e direção do poço [Thomas, 2001]. Essas informações podem ser enviadas pelo fluido de perfuração ou através de um cabo elétrico e são registradas de forma contínua e instantânea (no caso do cabo elétrico). É com base nessas informações que o engenheiro de petróleo toma a decisão sobre interferências na trajetória do poço.

O deslocamento da broca de perfuração, ao longo da trajetória do poço, pode ser visto como uma realização de uma série temporal, onde a componente aleatória corresponde à posição real da broca em cada momento. Por mais controle que se tenha do processo, essa posição não é determinística, isto é, pode ser vista como uma variável aleatória seguindo uma determinada distribuição de probabilidade.

Uma série temporal consiste de um conjunto de observações ordenadas no tempo [Morettin e Toloi, 2004]. São exemplos de séries temporais:

- Cotações diárias do barril de petróleo;
- Índice de poluição de uma região produtora de petróleo;
- Registros de marés em um porto marítimo;
- Preços diários das ações de uma empresa de petróleo, por exemplo, a Petrobras.

Em geral, na análise de uma série temporal, estamos interessados em:

- Investigar o mecanismo gerador dessa série;
- Fazer previsões para valores futuros da série;
- Procurar periodicidade relevantes nos dados.

Considerando essa abordagem do poço direcional como uma série temporal, o nosso interesse consiste em estudar modelos adequados para fazer previsões de posicionamentos futuros da broca, ou seja, previsões sobre a inclinação e direção do poço. Essas previsões podem ser feitas a partir dos registros obtidos contínua e instantaneamente, além da litologia da rocha atravessada pelo próprio poço ou das litologias das rochas atravessadas por outros poços do campo em desenvolvimento.

Como proposta de modelo para fazer as previsões, temos o modelo ARIMA (Auto-Regressivo Integrado e de Médias Móveis), dada a sua ampla divulgação e utilização [Morettin e Toloi, 2004].

Estatisticamente falando, estamos tratando de uma modelagem através de um modelo ARIMA, aplicada à inclinação e direção de um poço direcional. O desafio e a contribuição desse trabalho consiste na utilização de uma variável angular como variável resposta, uma vez que na literatura são utilizadas variáveis lineares.

No Capítulo 1, discutimos sobre a quantificação de incertezas de subsuperfície, bem como o impacto dessas incertezas sobre um programa de perfuração de poços petrolíferos.

No Capítulo 2, apresentamos um estudo sobre séries temporais, enfatizando a classe de modelos Auto-Regressivos Integrados e de Médias Móveis (ARIMA), utilizados para descrição, interpretação e previsão de séries temporais.

Dedicamos o Capítulo 3 a análise de séries temporais envolvendo dados circulares, já que a variável de interesse é a posição da broca ao longo da trajetória de um poço petrolífero.

Nos Apêndices, recordamos algumas definições e enunciamos os principais resultados utilizados nas demonstrações.

Capítulo 1

Quantificação de Incertezas de Subsuperfície

A modelagem de reservatórios é uma tarefa bastante árdua, devido à complexidade física envolvida na predição do escoamento e à dificuldade de se obter dados para a modelagem [Charles et al., 2001]. Com isso, as atividades relacionadas à predição dos parâmetros de interesse econômico, tais como: o volume total da rocha, a localização do óleo, perfis de produção e estimativas de reservas, são de difícil realização.

Os responsáveis por tomadas de decisão devem realizar uma quantificação sistemática dos riscos técnicos associados a qualquer desenvolvimento recente. Além disso, uma quantificação do impacto das diversas incertezas de subsuperfície (estruturais, geológicas e dinâmicas) sobre os parâmetros de interesse econômico, pode auxiliar a justificar a aquisição e processamento de mais dados e, com isso, reduzir a incerteza inerente ao processo de tomada de decisão.

A seguir, discutiremos o valor da quantificação de incertezas de subsuperfície no processo de tomada de decisões de investimentos, apresentaremos uma experiência relativa ao impacto das incertezas dinâmicas sobre o programa de perfuração e, por fim, veremos como utilizar a quantificação de incertezas estruturais para justificar a aquisição de dados complementares.

1.1 Incertezas de Subsuperfície

Atualmente, existem ferramentas para a construção de modelos geológicos 3D e para a quantificação de incertezas sobre os parâmetros associados a esses modelos [Charles et al., 2001]. Tais ferramentas têm facilitado a compreensão do impacto de cada uma das incertezas de subsuperfície sobre o campo de produção. Por exemplo, existe uma cadeia de ferramentas desenvolvidas para lidar com a quantificação de incertezas de subsuperfície cujo suporte é constituído por três *softwares* principais, a saber, *ALEA*, *JACTATM* e *EST*.

A partir de mapas de incerteza produzidos por intérpretes sísmicos, o *ALEA* simula diversos modelos estruturais do reservatório, calcula os correspondentes volumes totais de rochas, além de exportar estas superfícies para o *JACTATM*. Com isso, é possível quantificar o impacto das incertezas oriundas, por exemplo, da conversão tempo-profundidade sobre as incertezas associadas ao volume total da rocha.

Após identificar os parâmetros que afetam as propriedades do reservatório, pode-se simular diversos modelos geológicos de reservatórios, bem como calcular seus respectivos volumes.

O *software JACTATM* realiza uma combinação entre ambientes deposicionais, tipos de rochas e simulações de parâmetros petrofísicos e permite que as incertezas que afetam os parâmetros geoestatísticos sejam incorporadas. As realizações resultantes podem ser visualizadas em 3D, analisadas e exportadas para um simulador de fluxo.

O *software EST* possibilita a simulação de fluxo para cada realização geoestatística proveniente do *JACTATM*, no entanto, estas simulações podem ser muito dispendiosas e incompletas, já que o *ALEA* e o *JACTATM* consideram apenas as incertezas estáticas (estruturais e geológicas), não levando em conta as incertezas dinâmicas, tais como: permeabilidade relativa, transmissividade defeituosa, ou qualquer parâmetro de fluxo. Assim, para minimizar a quantidade de operações de simulação de fluxo, é necessário utilizar ferramentas e métodos para incorporar incertezas dinâmicas na quantificação de incertezas associadas a perfis de produção ou a estimativas de reservas.

Após identificar as principais incertezas de subsuperfície, pode-se utilizar os três *softwares ALEA*, *JACTATM* e *EST* para transformar essas incertezas de subsuperfície

em incertezas dos parâmetros de interesse, durante o processo de tomada de decisão. Lembrando que a qualidade do resultado dependerá da confiabilidade das incertezas de subsuperfície.

1.2 Impacto de Incertezas Dinâmicas sobre um Programa de Perfuração

A seguir descreveremos uma experiência, relatada em [Charles et al., 2001], que foi realizada em dois campos petrolíferos, denotados por campo X e campo Y (por motivos relacionados a sigilo).

Um determinado campo X, desenvolvido recentemente, possui um campo satélite Y, ambos de alta pressão e temperatura e separados por uma falha principal e por uma falha tectônica contendo um fluido desconhecido. A produção do campo X pode reduzir o volume de fluidos do campo Y. Tal redução precisa ser quantificada, já que é impossível perfurar qualquer poço após uma depleção de 100 barris. Surge, então, um questionamento: após o início da produção do campo X, por quanto tempo o desenvolvimento do campo Y pode ser adiado?

Há duas alternativas para o desenvolvimento do campo Y: ou a realização de uma perfuração vertical a partir de uma nova plataforma ou uma perfuração direcional a partir de uma plataforma já existente no campo X.

Como o interesse é modelar apenas a depleção, basta considerar somente as incertezas dinâmicas, ou seja, não é necessário fazer uma representação das heterogeneidades ou estruturas do campo Y.

Com a utilização de um modelo de simulação de fluxo, construído pelo operador, foi possível realizar as seguintes atividades:

- 1 - Definição da variável resposta: depleção média em todas as camadas do campo Y;
- 2 - Identificação dos principais parâmetros de incerteza: valor da permeabilidade absoluta da falha tectônica e do aquífero; permeabilidade relativa; tamanho do aquífero a oeste; variação de porosidade dentro do aquífero a norte; principais falhas de transmissividade; faixas de permeabilidade no reservatório; anisotropia

vertical e fluido acumulado na falha tectônica;

- 3 - Identificação da ordem de incerteza dos parâmetros selecionados;
- 4 - Estimação da função densidade de probabilidade (*fdp*) associada a cada parâmetro (tratado como variável aleatória);
- 5 - Utilização da metodologia experimental para identificar os parâmetros de impacto mais significativo na depleção. Os únicos parâmetros que interferiram na depleção foram: a permeabilidade relativa, o fluido acumulado na falha tectônica e a transmissividade defeituosa. A partir da simulação da variável resposta, construiu-se uma superfície resposta como função dos valores assumidos por esses três parâmetros;
- 6 - Realização de uma simulação Monte Carlo utilizando, tanto as *fdp*'s associadas aos parâmetros, como o modelo analítico da superfície de resposta, fornecendo perfis de prováveis depleções.

Para cada uma das hipóteses (distribuição triangular da transmissividade defeituosa e falha de escoamento) foi construído um perfil de depleção. No primeiro caso, ocorreu uma depleção de 100 barris, após 1,7 ano de produção no campo X, enquanto que, no segundo caso, o tempo de depleção foi de 1,2 ano.

Com base no cenário mais pessimista, decidiu-se desenvolver o campo Y, a partir de uma plataforma já existente no campo X. Associado a esta decisão, admitiu-se um risco de 5% (após 1,3 ano de produção do campo X).

1.3 Justificativa da Aquisição de Dados Complementares

A partir da interpretação de dados sísmicos 2D migrados no tempo e de 50 poços disponíveis em um campo maduro, construíram-se mapas de profundidade do nível do reservatório [Charles et al., 2001]. Esses mapas facilitaram a estimativa do possível volume total da rocha e a compreensão dos parâmetros de maior incerteza, tornando possível justificar uma aquisição de dados sísmicos e definir um processamento mais ajustado.

Os parâmetros de incerteza mais significativos foram registrados e ordenados da seguinte maneira:

- Campo de velocidade utilizado para a migração sísmica no tempo;
- Interpolação de dados sísmicos 2D e valores de poços;
- Conversão tempo-profundidade;
- OWC (ponto de contato água-óleo).

A combinação de todas estas incertezas gerou um intervalo de confiança em torno da profundidade do topo do reservatório. Além disso, a simulação de 200 mapas de possíveis profundidades em torno do mapa base, e dentro do intervalo de confiança, resultou numa série de possíveis valores para o volume total da rocha. Porém, para melhor estimar o potencial deste campo, as duas maiores incertezas deveriam ser reduzidas.

Portanto, mesmo quando se trata de um campo maduro com muitos poços, as incertezas geométricas podem influenciar no volume total da rocha. Assim, os parâmetros a ser melhorados precisam ser identificados, a fim de reduzir essas incertezas. Recomenda-se, também, a aquisição de dados sísmicos 3D e que esses tais dados sejam processados utilizando-se uma migração de profundidade melhor do que uma migração de tempo clássica.

Capítulo 2

Séries Temporais

Uma série temporal pode ser vista como um conjunto de observações Z_t , geradas sequencialmente no tempo [Box e Jenkins, 1976]. Fazemos referência ao parâmetro t como sendo o tempo, mas a série Z_t poderá ser função de algum outro parâmetro físico, como espaço, volume, profundidade, etc. Se o conjunto de instantes de tempo for discreto (enumerável) ou não-enumerável, a série será discreta ou contínua, respectivamente. De um modo mais formal, uma série temporal é uma realização ou trajetória de um processo estocástico.

Um processo estocástico é uma família de variáveis aleatórias $\{Z_t; t \in T\}$ definidas num mesmo espaço de probabilidades. Ou seja, para cada $t \in T$, Z_t é uma variável aleatória definida sobre o espaço amostral Ω . Portanto, Z_t é uma função de dois argumentos, $Z(t, w)$, onde $t \in T$ e $w \in \Omega$.

Na Figura 2.1, podemos observar que, para cada $t \in T$, $Z(t, w)$ é uma variável aleatória com uma distribuição de probabilidade. Por outro lado, para cada $w \in \Omega$ fixado, obtemos uma função do tempo, ou seja, uma realização do processo.

São exemplos de séries temporais:

- 1- Valores diários de poluição numa região produtora de petróleo;
- 2- Preços diários das ações de uma empresa de petróleo;
- 3- Cotações diárias do barril de petróleo;
- 4- Rendimento anual per capita;

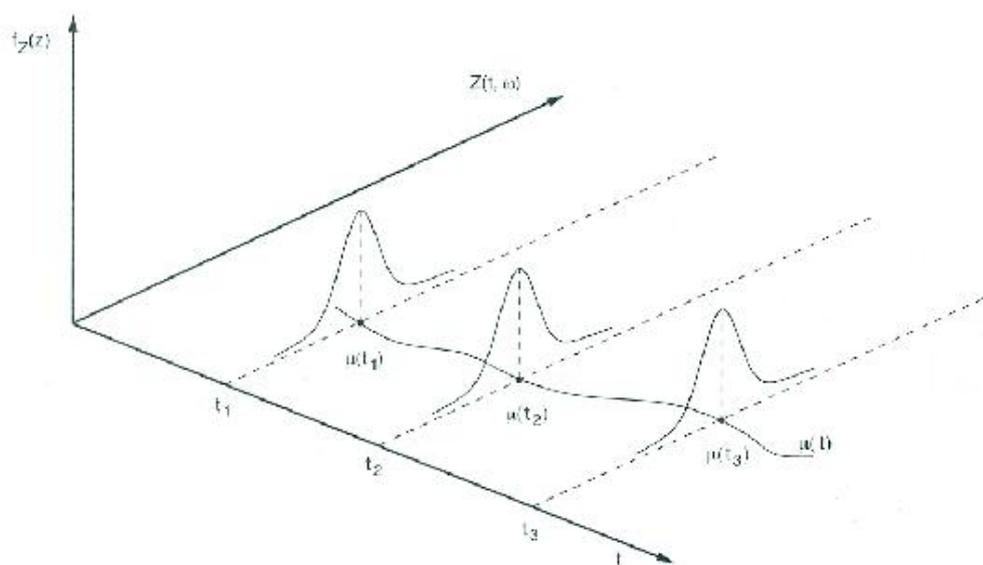


Figura 2.1: Processo estocástico como uma família de variáveis aleatórias.

- 5- Inflação mensal de uma determinada cidade;
- 6- Intensidade da corrente elétrica num dado ponto;
- 7- Intensidade do som num determinado local;
- 8- Registro de marés em um porto marítimo;

As séries 1 a 5 são discretas, enquanto que, as séries 6 a 8 são contínuas.

Os principais objetivos da análise de uma série temporal são

- Investigar o mecanismo gerador dessa série;
- Descrever o comportamento da série;
- Procurar periodicidades relevantes nos dados;
- Realizar previsões de valores futuros da série.

Para atingir esses objetivos, lançamos mão de modelos estocásticos (ou probabilísticos). Uma classe importante de modelos estocásticos para descrição de séries temporais é a dos modelos estacionários, que são baseados na hipótese de que o processo permanece em equilíbrio em torno de um nível médio constante. Em outras

palavras, o processo evolui no tempo de modo que a escolha de uma origem dos tempos não é importante, ou seja, as características de Z_{t+k} , para todo k , são as mesmas de Z_t [Morettin e Toloi, 2004]. Desta forma, a média $\mu(t)$ e a variância $V(t)$ de Z_t são constantes para todo $t \in T$, ou seja,

$$\mu(t) = E[Z_t] = \mu \quad \text{e} \quad V(t) = \text{Var}[Z_t] = E[(Z_t - \mu)^2] = \sigma^2.$$

A covariância entre Z_t e Z_{t+k} , é denominada *função de autocovariância* (facv), e é definida por

$$\gamma_k = \text{Cov}[Z_t, Z_{t+k}] = E[(Z_t - \mu)(Z_{t+k} - \mu)].$$

Pela própria definição de γ_k , temos que $\gamma_0 = \text{Var}[Z_t] = \sigma^2$ e, sendo o processo estacionário, $|\gamma_k| \rightarrow 0$ quando $k \rightarrow \infty$. Este comportamento pode ser observado na Figura 2.2.

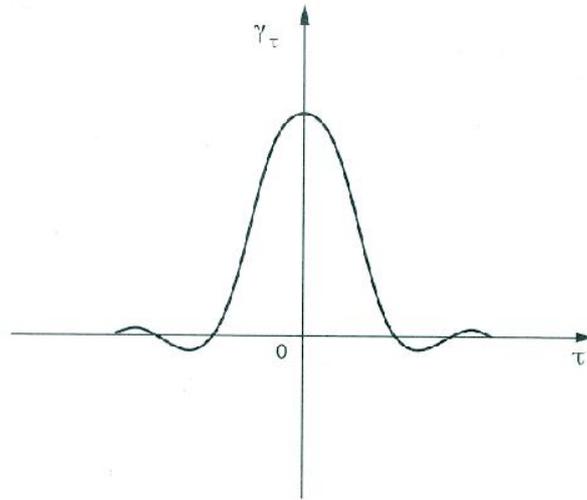


Figura 2.2: Representação da função de autocovariância.

Como a facv pode ser sensível às unidades em que são medidas as observações, é comum utilizarmos a *função de autocorrelação* (fac), dada por

$$\rho_k = \frac{\text{Cov}[Z_t, Z_{t+k}]}{\sqrt{\text{Var}[Z_t]\text{Var}[Z_{t+k}]}} \quad , \quad k \in T.$$

Se o processo for estacionário, então a variância $\sigma^2 = \gamma_0$ é a mesma, tanto no tempo $t+k$ como em t . Assim,

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\gamma_k}{\sigma^2} \quad , \quad k \in T.$$

Observe que a fac é simétrica em torno do zero e $\rho_k = \rho_{-k}$, para todo k . A Figura 2.3 mostra a fac como um gráfico dos valores localizados nas diagonais da matriz de autocorrelação.

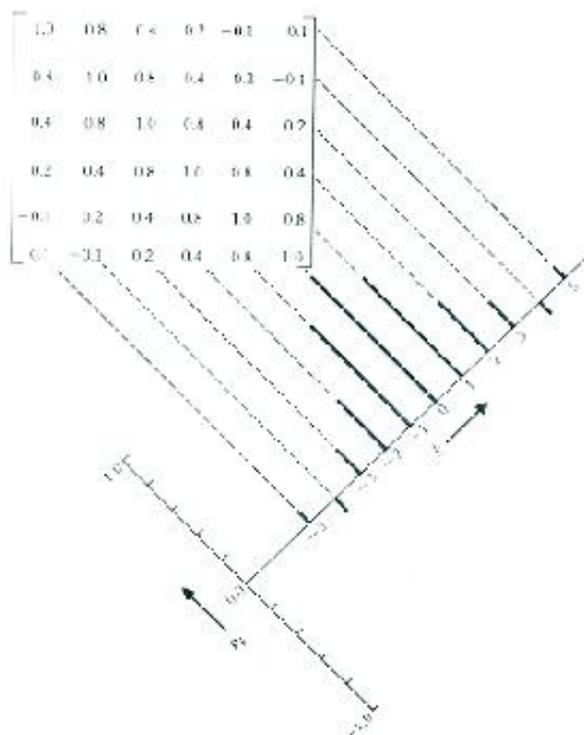


Figura 2.3: Uma matriz de autocorrelação e a fac correspondente.

A fac ρ_k pode ser estimada através da expressão:

$$r_k = \frac{c_k}{c_0},$$

onde

$$c_k = \frac{1}{N} \sum_{t=1}^{N-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z}), k = 0, 1, \dots, N - 1$$

é a estimativa da função de autocovariância γ_k e \bar{Z} é a média amostral da série temporal.

Para que seja viável descrever uma série temporal através de modelos estacionários, devemos supor que tal série é estacionária. No entanto, na prática, a maioria das séries que encontramos apresentam algum tipo de não-estacionariedade, por exemplo, existem séries não-estacionárias quanto ao nível e outras quanto ao nível e à inclinação, como mostram as Figuras 2.4 e 2.5. Outro tipo de não-estacionariedade é a explosiva,

que surge em séries que representam o crescimento de uma colônia de bactérias, por exemplo.

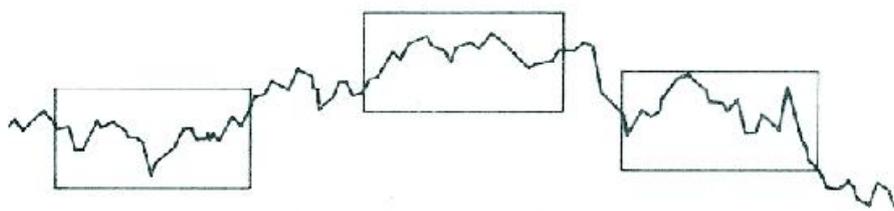


Figura 2.4: Representação de uma série não-estacionária quanto ao nível.

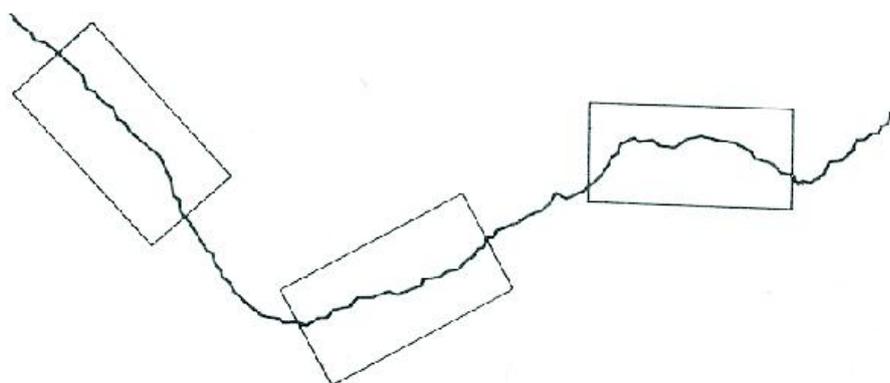


Figura 2.5: Representação de uma série não-estacionária quanto ao nível e à inclinação.

Mais adiante abordaremos a classe de modelos ARIMA, que será útil para descrever de maneira satisfatória séries estacionárias e séries não-estacionárias que não apresentam comportamento explosivo.

A fim de facilitar a manipulação dos modelos abordados mais adiante, utilizaremos o operador translação para o passado, denotado por B e definido por

$$BZ_t = Z_{t-1}$$

$$B^m Z_t = Z_{t-m}, \quad 2 \leq m < t.$$

Mesmo quando uma série é não-estacionária, podemos transformar os dados originais, a fim de tentar obter uma série estacionária. O procedimento mais utilizado consiste em diferenciar sucessivamente a série original, até se obter uma série estacionária. Diferenciar, aqui, significa considerar *diferenças* sucessivas da série original.

A primeira diferença de Z_t é definida por

$$\Delta Z_t = Z_t - Z_{t-1} = Z_t - BZ_t = (1 - B)Z_t,$$

onde B é o operador translação para o passado.

A segunda diferença é

$$\begin{aligned}\Delta^2 Z_t &= \Delta[\Delta Z_t] = \Delta[Z_t - Z_{t-1}] = Z_t - 2Z_{t-1} + Z_{t-2} \\ &= (1 - 2B + B^2)Z_t = (1 - B)^2 Z_t.\end{aligned}$$

A n -ésima diferença de Z_t é definida por

$$\Delta^n Z_t = \Delta[\Delta^{n-1} Z_t].$$

Em geral, pode-se considerar vários modelos diferentes para descrever o comportamento de uma série. No entanto, devemos utilizar critérios de comparação entre eles, a fim de escolher o modelo mais parcimonioso, ou seja, aquele com uma quantidade mínima de parâmetros e que forneça previsões bastante precisas. A escolha do modelo adequado baseia-se num ciclo iterativo do método de Box e Jenkins [Morettin e Toloi, 2004], cujas etapas consistem em:

- 1- Fazer uma descrição da série, através do cálculo de estatísticas resumo e da representação gráfica dos dados e, a partir daí, escolher uma classe de modelos para a análise;
- 2- Identificar um modelo através da análise de autocorrelações, dentre outros critérios;
- 3- Estimar os parâmetros do modelo identificado;
- 4- Realizar uma análise de resíduos, a fim de verificar se o modelo ajustado é adequado para fazer previsões de valores futuros da série.

Se o modelo identificado não for adequado, o ciclo deve ser repetido a partir da etapa 2.

2.1 Modelos Lineares

Os modelos abordados a seguir são casos particulares de um *modelo de filtro linear*. A principal suposição deste modelo é que a série temporal tenha sido gerada a

partir de um filtro linear, ilustrado na Figura 2.6, cuja entrada é um ruído branco a_t , ou seja, para cada $t \in T$, a_t é uma variável aleatória com

$$\begin{aligned} E[a_t] &= 0, \forall t, \\ \text{Var}[a_t] &= \sigma_a^2, \forall t, \\ E[a_t a_s] &= 0, s \neq t. \end{aligned}$$

Assim, a série pode ser expressa da seguinte maneira

$$\begin{aligned} Z_t &= \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots \\ &= \mu + \psi(B)a_t, \end{aligned} \tag{2.1}$$

onde μ , em geral, é o parâmetro que determina o *nível* da série e

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots$$

é o operador linear, cuja finalidade é transformar a_t em Z_t , denominado *função de transferência* do filtro.

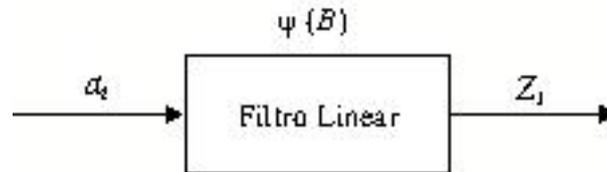


Figura 2.6: Série temporal gerada por um filtro linear.

Quando a série de pesos ψ_1, ψ_2, \dots for finita ou infinita convergente, então Z_t é estacionária com média μ . Caso contrário, Z_t é não-estacionária e μ não tem significado específico [Morettin e Tolo, 2004].

Lembrando que a_t é um ruído branco e supondo que $\sum_{k=0}^{\infty} \psi_k^2 < \infty$, temos que a cov de Z_t pode ser escrita da seguinte maneira

$$\gamma_k = \sigma_a^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+k},$$

com $\psi_0 = 1$. Assim, para $k = 0$, obtemos a variância de Z_t ,

$$\gamma_0 = \text{Var}[Z_t] = \sigma_a^2 \sum_{k=0}^{\infty} \psi_k^2.$$

A série $\tilde{Z}_t = Z_t - \mu$, pode ser escrita como uma soma de valores passados mais um ruído a_t , ou seja,

$$\tilde{Z}_t = \pi_1 \tilde{Z}_{t-1} + \pi_2 \tilde{Z}_{t-2} + \cdots + a_t,$$

ou ainda,

$$\tilde{Z}_t - \pi_1 \tilde{Z}_{t-1} - \pi_2 \tilde{Z}_{t-2} - \cdots = a_t,$$

donde segue que

$$\pi(B) \tilde{Z}_t = a_t, \quad (2.2)$$

onde $\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \cdots$.

Comparando as expressões (2.1) e (2.2), temos que

$$\pi(B) \psi(B) a_t = a_t,$$

daí,

$$\pi(B) = \psi^{-1}(B), \quad (2.3)$$

mostrando que os pesos π_k podem ser obtidos a partir dos pesos ψ_k e vice-versa.

Quanto às condições de estacionariedade e invertibilidade, um processo linear será estacionário se a série $\psi(B)$ convergir para $|B| \leq 1$ e será invertível se $\pi(B)$ convergir para $|B| \leq 1$ [Morettin e Tolo, 2004].

2.1.1 Modelos Auto-Regressivos

Considerando o caso especial de (2.2), em que $\pi_k = 0$, $k > p$, e renomeando os pesos de π_k para ϕ_k , obtemos o *modelo auto-regressivo de ordem p*, denotado por AR(p)

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \cdots + \phi_p \tilde{Z}_{t-p} + a_t, \quad (2.4)$$

ou equivalentemente,

$$\phi(B) \tilde{Z}_t = a_t, \quad (2.5)$$

onde

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$$

é chamado operador auto-regressivo de ordem p .

De (2.5) temos que

$$\tilde{Z}_t = \frac{1}{\phi(B)} a_t = \phi^{-1}(B) a_t,$$

ou seja, o modelo $AR(p)$ pode ser visto como a saída \tilde{Z}_t de um filtro linear, com função de transferência $\phi^{-1}(B)$, desde que a entrada a_t seja um ruído branco.

Para que o processo Z_t seja estacionário, a série $\psi(B) = \phi^{-1}(B)$ deve convergir para $|B| \leq 1$, ou seja, as raízes de $\phi(B) = 0$ devem cair fora do círculo unitário.

Por outro lado, como a série $\pi(B) = \phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ é finita, conseqüentemente, $\pi(B)$ é convergente para $|B| \leq 1$, então não há restrições sobre os parâmetros de um processo auto-regressivo para garantir a invertibilidade de Z_t .

Para encontrar a fac de um processo $AR(p)$, devemos, primeiramente, multiplicar ambos os membros de (2.4) por Z_{t-k} e, em seguida, calcular o valor esperado

$$E[\tilde{Z}_t \tilde{Z}_{t-k}] = \phi_1 E[\tilde{Z}_{t-1} \tilde{Z}_{t-k}] + \phi_2 E[\tilde{Z}_{t-2} \tilde{Z}_{t-k}] + \dots + \phi_p E[\tilde{Z}_{t-p} \tilde{Z}_{t-k}] + E[a_t \tilde{Z}_{t-k}] \quad (2.6)$$

Mas, para $k > 0$, temos $E[a_t \tilde{Z}_{t-k}] = 0$, pois \tilde{Z}_{t-k} envolve ruídos apenas até a_{t-k} , não-correlacionados. Com isso,

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p}, \quad k > 0.$$

Assim, dividindo ambos os membros dessa expressão por $\gamma_0 = \text{Var}[Z_t]$, obtemos a fac

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p}, \quad k > 0. \quad (2.7)$$

Segundo Box e Jenkins, a fac de um processo $AR(p)$, consiste de uma mistura de exponenciais e senóides amortecidas.

A variância do processo pode ser obtida fazendo $k = 0$ na expressão (2.6), obtendo

$$\text{Var}(\tilde{Z}_t) = \text{Var}(Z_t) = \gamma_0 = \phi_1 \gamma_1 + \dots + \phi_p \gamma_p + \sigma_a^2.$$

Dividindo ambos os membros por γ_0 , obtemos

$$1 = \phi_1 \rho_1 + \dots + \phi_p \rho_p + \frac{\sigma_a^2}{\rho_0},$$

donde segue que

$$\gamma_0 = \sigma_a^2 / (1 - \phi_1 \rho_1 - \dots - \phi_p \rho_p). \quad (2.8)$$

Os parâmetros auto-regressivos ϕ_1, \dots, ϕ_p podem ser escritos em termos de $\rho_1, \rho_2, \dots, \rho_p$. Para tanto, basta substituir $k = 1, 2, \dots, p$ em (2.7), obtendo um sistema

A condição de invertibilidade para um processo MA(q) é que $\pi(B) = \theta^{-1}(B)$ convirja para $|B| \leq 1$, isto é, as raízes de $\theta(B) = 0$ devem cair fora do círculo unitário.

A facv de um modelo MA(q) é

$$\begin{aligned} \gamma_k &= E[(a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q})(a_{t-k} - \theta_1 a_{t-k-1} - \dots - \theta_q a_{t-k-q})] \\ &= E \left\{ \left[a_t - \sum_{i=1}^q \theta_i a_{t-i} \right] \left[a_{t-k} - \sum_{j=1}^q \theta_j a_{t-k-j} \right] \right\} \\ &= E[a_t a_{t-k}] - \sum_{i=1}^q \theta_i E[a_{t-k} a_{t-i}] - \sum_{j=1}^q \theta_j E[a_t a_{t-k-j}] + \sum_{i=1}^q \sum_{j=1}^q \theta_i \theta_j E[a_{t-j} a_{t-k-j}]. \end{aligned}$$

Sabendo que

$$E[a_t a_{t-k}] = \begin{cases} \sigma_a^2, & k = 0 \\ 0, & k \neq 0, \end{cases}$$

obtemos

$$\gamma_0 = \text{Var}[Z_t] = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma_a^2 \quad (2.10)$$

e

$$\gamma_k = \begin{cases} (-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q) \sigma_a^2, & k = 1, 2, \dots, q \\ 0, & k > q. \end{cases}$$

Donde segue que a facv de Z_t é

$$\rho_k = \begin{cases} \frac{-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2}, & k = 1, 2, \dots, q \\ 0, & k > q. \end{cases} \quad (2.11)$$

Ao contrário do que ocorre com um modelo AR(p), a facv de um modelo MA(q) se anula para *lags* maiores do que q .

2.1.3 Modelos Auto-Regressivos e de Médias Móveis

Uma das maneiras de tornar um modelo mais parcimonioso, consiste em considerar, simultaneamente, termos auto-regressivos e termos de médias móveis. Com isso, surge uma classe de modelos mistos, denominados *modelos auto-regressivos e de médias móveis de ordem* (p, q), denotados por ARMA(p, q)

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \dots + \phi_p \tilde{Z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q},$$

isto é,

$$\phi(B) \tilde{Z}_t = \theta(B) a_t,$$

onde $\phi(B)$ e $\theta(B)$ são os operadores auto-regressivos e de médias móveis, respectivamente.

As condições de estacionariedade e invertibilidade para um processo ARMA(p, q) é que as raízes de $\phi(B) = 0$ e de $\theta(B) = 0$ caiam fora do círculo unitário.

A facv de um modelo ARMA(p, q) é

$$\gamma_k = E\{(\phi_1 \tilde{Z}_{t-1} + \cdots + \phi_p \tilde{Z}_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}) \tilde{Z}_{t-k}\}.$$

Lembrando que \tilde{Z}_{t-k} depende apenas de choques a_{t-k} , ocorridos até o tempo $t-k$, temos que a covariância cruzada entre \tilde{Z}_t e a_t , definida por

$$\gamma_{za}(k) = E[a_t \tilde{Z}_{t-k}],$$

se anula para valores de $k > 0$ e é diferente de zero para $k \leq 0$. Daí, a facv fica na forma

$$\gamma_k = \phi_1 \gamma_{k-1} + \cdots + \phi_p \gamma_{k-p} + \gamma_{za}(k) - \theta_1 \gamma_{za}(k-1) - \cdots - \theta_q \gamma_{za}(k-q). \quad (2.12)$$

Para $k > q$, obtemos

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \cdots + \phi_p \gamma_{k-p}, \quad k > q.$$

Portanto, a fac do modelo é

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p}, \quad k > q,$$

mostrando que as autocorrelações, para $k > q$, se comportam como nos modelos auto-regressivos.

2.1.4 Modelos Auto-Regressivos Integrados e de Médias Móveis

A seguir, abordaremos uma classe de modelos apropriados para descrever séries temporais não-estacionárias homogêneas, ou seja, séries que, apesar de não evoluírem em torno de uma média constante ao longo do tempo, quando diferenciadas d vezes, tornam-se estacionárias. Por exemplo, se a série for não-estacionária quanto ao nível, então $d = 1$. Isto significa que basta calcular sua primeira diferença para torná-la estacionária. Já séries não-estacionárias quanto à inclinação, devem ser diferenciadas duas vezes ($d = 2$), para obter a estacionariedade [Morettin e Tolo, 2004].

Se $W_t = \Delta^d Z_t$ for estacionária, podemos representá-la através de um modelo ARMA(p, q)

$$\phi(B)W_t = \theta(B)a_t.$$

Neste caso, dizemos que Z_t é uma *integral* de W_t , já que diferenciando Z_t (no sentido de diferença sucessivas) obtemos W_t . Dizemos, ainda, que Z_t segue um *modelo auto-regressivo integrado de médias móveis de ordem* (p, d, q), denotado por ARIMA(p, d, q)

$$\phi(B)\Delta^d Z_t = \theta(B)a_t. \quad (2.13)$$

Sendo W_t estacionária, então todas as raízes de $\phi(B) = 0$ caem fora do círculo unitário.

Uma forma alternativa de escrever a expressão (2.13) é

$$\varphi(B)Z_t = \theta(B)a_t, \quad (2.14)$$

em que

$$\varphi(B) = \phi(B)\Delta^d = \phi(B)(1 - B)^d$$

é um operador auto-regressivo não-estacionário de ordem $p + d$, com d raízes sobre o círculo unitário e as p restantes, fora do círculo unitário.

Com essa notação, o modelo ARIMA pode ser representado pela seguinte expressão

$$Z_t = \varphi_1 Z_{t-1} + \cdots + \varphi_{p+d} Z_{t-p-d} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}, \quad (2.15)$$

que é denominada *equação de diferenças*, bastante útil para o cálculo de previsões.

Quando o interesse é calcular a variância dos erros de previsão, é conveniente expressar o modelo ARIMA na forma de *choques aleatórios*, ou seja, em termos do valor atual e prévios de a_t , ou seja,

$$Z_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots = \psi(B)a_t. \quad (2.16)$$

Outra maneira de representar o modelo ARIMA é a forma *invertida*, que consiste em expressar Z_t em termos de seus valores prévios e do valor atual de a_t , isto é,

$$Z_t = \pi_1 Z_{t-1} + \pi_2 Z_{t-2} + \cdots + a_t. \quad (2.17)$$

Às vezes, é útil considerar uma extensão do modelo ARIMA, acrescentando um termo constante θ_0 na expressão (2.13), obtendo

$$\varphi(B)Z_t = \phi(B)\Delta^d Z_t = \theta_0 + \theta(B)a_t. \quad (2.18)$$

Se $\theta_0 = 0$, o modelo (2.18) pode ser usado para representar séries com tendências estocásticas, ou seja, séries que apresentam mudanças aleatórias no nível e/ou na inclinação. Se $\theta_0 \neq 0$, então o modelo (2.18) é capaz de representar séries com tendência polinomial determinística de grau d . Além disso,

$$E(W_t) = \mu_w = \theta_0 / (1 - \phi_1 - \phi_2 - \dots - \phi_p).$$

O modelo ARIMA é uma generalização dos modelos vistos anteriormente, já que

$$\begin{aligned} \text{ARIMA}(p, 0, 0) &= \text{AR}(p), \\ \text{ARIMA}(0, 0, q) &= \text{MA}(q) \text{ e} \\ \text{ARIMA}(p, 0, q) &= \text{ARMA}(p, q). \end{aligned}$$

2.2 A Função de Autocorrelação Parcial

A função de autocorrelação parcial (facp) é um instrumento bastante útil durante a etapa de identificação do modelo a ser ajustado aos dados observados. Vejamos, a seguir, como essa função é construída.

Denotando por ϕ_{kj} o j -ésimo coeficiente de um modelo $\text{AR}(k)$, temos que ϕ_{kk} é o último coeficiente. Utilizando essa notação, as equações de Yule-Walker podem ser escritas da seguinte maneira:

$$\begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & 1 \end{bmatrix} = \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix}.$$

Resolvendo, sucessivamente, estas equações para $k = 1, 2, \dots$, obtemos

$$\phi_{11} = \rho_1, \quad \phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}, \quad \phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}, \dots$$

De modo geral, para ρ_{kk} , a matriz no numerador é a mesma que a matriz no denominador, exceto pela última coluna, que é substituída pelo vetor de autocorrelação $\boldsymbol{\rho}_k = (\rho_1, \dots, \rho_k)^t$.

A *função de autocorrelação parcial* é definida como sendo a quantidade ϕ_{kk} , encarada como função de k .

Para um processo $\text{AR}(p)$, a facp se anula para todos as defasagens maiores do que p , isto é, o seu gráfico apresenta um "corte" após o defasagem p . Portanto, o gráfico dessa função permite identificar o grau do polinômio auto-regressivo. Já para o processo $\text{MA}(q)$, a facp é dominada por uma mistura de exponenciais e/ou senóides amortecidas. Tal comportamento é semelhante ao da fac de um processo $\text{AR}(p)$. Por fim, a facp de um processo $\text{ARMA}(p,q)$, comporta-se de maneira similar à facp de um processo MA puro [Morettin e Tolo, 2004].

Durante o estágio de identificação do modelo precisaremos calcular estimativas das facp, a fim de compará-las com as respectivas facp teóricas. Por exemplo, no caso dos modelos AR, tais estimativas podem ser feitas, ajustando-se, sucessivamente, processos auto-regressivos de ordem $p = 1, 2, 3, \dots$ por mínimos quadrados e considerando as estimativas $\hat{\phi}_{11}, \hat{\phi}_{22}, \hat{\phi}_{33}, \dots$ do último coeficiente de cada ordem. A facp estimada pode ser obtida, de modo alternativo, substituindo-se, nas equações de Yule-Walker, as fac ρ_j por suas estimativas r_j , isto é,

$$r_j = \hat{\phi}_{k1}r_{j-1} + \hat{\phi}_{k2}r_{j-2} + \dots + \hat{\phi}_{kk}r_{j-k}, j = 1, \dots, k$$

e resolvendo-se essas equações para $k = 1, 2, \dots$

2.3 Alguns Casos Particulares de Modelos Lineares

2.3.1 Modelo Auto-Regressivo de Ordem 1 - AR(1)

O modelo AR(1) é dado por

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + a_t,$$

ou equivalentemente,

$$\phi(B)\tilde{Z}_t = a_t,$$

onde $\phi(B) = 1 - \phi_1 B$.

Para que o processo seja estacionário é necessário que $-1 < \phi_1 < 1$.

Por (2.7), a fac de um processo AR(1) é da forma

$$\rho_k = \phi_1 \rho_{k-1}, \quad k > 0$$

cuja solução é

$$\rho_k = \phi_1^k, \quad k \geq 0.$$

Donde segue que, se $\phi_1 > 0$, a fac decai exponencialmente e, caso $\phi_1 < 0$, ela também decai exponencialmente, alternando valores positivos e negativos. A Figura 2.7 ilustra esse comportamento para $\phi_1 = 0,8$ e $\phi_1 = -0,8$.

Por (2.8) a variância de um processo AR(1) é

$$\gamma_0 = \frac{\sigma_a^2}{1 - \phi_1^2} \Rightarrow \gamma_0 = \frac{\sigma_a^2}{1 - \phi_1^2}.$$

2.3.2 Modelo Auto-Regressivo de Ordem 2 - AR(2)

O modelo AR(2) é dado por

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + a_t,$$

ou ainda,

$$\phi(B)\tilde{Z}_t = a_t,$$

onde $\phi(B) = 1 - \phi_1 B - \phi_2 B^2$.

Para que o processo seja estacionário é preciso que

$$\phi_1 + \phi_2 < 1, \quad \phi_2 - \phi_1 < 1, \quad -1 < \phi_2 < 1.$$

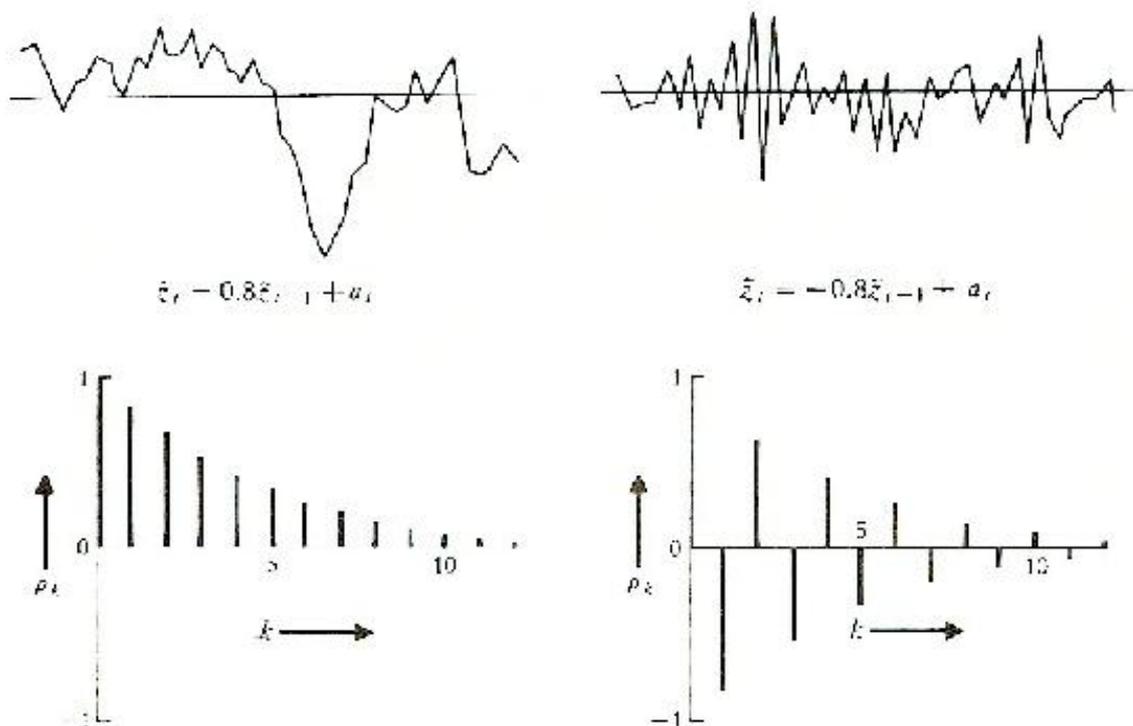


Figura 2.7: Processos AR(1) e suas correspondentes funções de autocorrelação.

A fac de um processo AR(2) é

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}, \quad k > 0.$$

Substituindo $p = 2$ nas equações de Yule-Walker, obtemos

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 \end{aligned}$$

donde segue que

$$\phi_1 = \rho_1(1 - \rho_2)/(1 - \rho_1^2) \quad \text{e} \quad \phi_2 = (\rho_2 - \rho_1^2)/(1 - \rho_1^2).$$

Utilizando as equações de Yule-Walker, podemos também expressar ρ_1 e ρ_2 em função de ϕ_1 e ϕ_2 , da seguinte maneira

$$\rho_1 = \phi_1/(1 - \phi_2) \quad \text{e} \quad \rho_2 = \phi_2 + \phi_1^2/(1 - \phi_2).$$

A Figura 2.8 ilustra a fac de um processo AR(2) para $\phi_1 = 1$, $\phi_2 = 0,89$ e $\phi_1 = -1$, $\phi_2 = -0,89$.

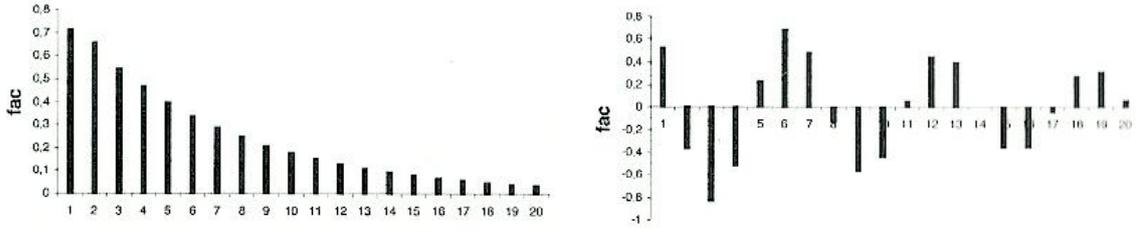


Figura 2.8: Funções de autocorrelação para um processo AR(2).

Para obter a variância de um processo AR(2), basta substituir $p = 2$ em (2.8), obtendo

$$\gamma_0 = \frac{\sigma_a^2}{1 - \phi_1\rho_1 - \phi_2\rho_2}.$$

2.3.3 Modelo de Médias Móveis de Ordem 1 - MA(1)

O modelo MA(1) é representado por

$$\tilde{Z}_t = a_t - \theta_1 a_{t-1} = \theta(B)a_t,$$

onde $\theta(B) = 1 - \theta_1 B$

O processo é invertível se $-1 < \theta_1 < 1$.

Substituindo $q = 1$ na expressão (2.10), obtemos a variância do processo

$$\gamma_0 = (1 + \theta_1^2)\sigma_a^2.$$

Utilizando (2.11), encontramos a função de autocorrelação

$$\rho_k = \begin{cases} \frac{-\theta_1}{1+\theta_1^2}, & k = 1 \\ 0, & k > 1. \end{cases}$$

A Figura 2.9 apresenta a fac de um processo MA(1) para $\theta_1 = 0,8$.

2.3.4 Modelo de Médias Móveis de Ordem 2 - MA(2)

O modelo MA(2) é dado por

$$\tilde{Z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} = \theta(B)a_t,$$

onde $\theta(B) = 1 - \theta_1 B - \theta_2 B^2$.

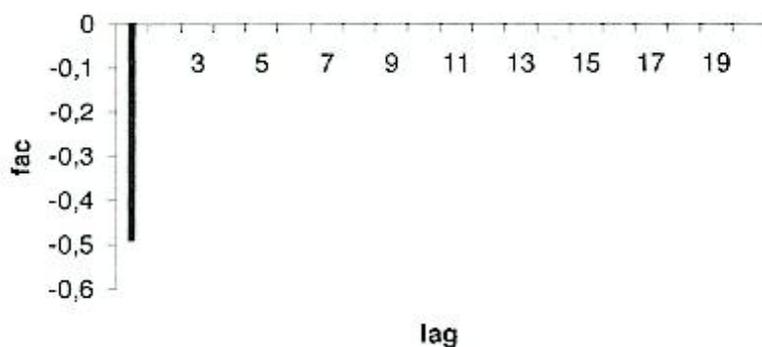


Figura 2.9: Função de autocorrelação de um processo MA(1).

Para que o processo seja invertível é necessário que as raízes de $\theta(B) = 0$ caiam fora do círculo unitário, ou seja, devemos ter

$$\theta_1 + \theta_2 < 1, \quad \theta_2 - \theta_1 < 1, \quad -1 < \theta_2 < 1.$$

Observe que essas condições são equivalentes às condições de estacionariedade para um processo AR(2).

A partir de (2.10) e (2.11) obtemos

$$\begin{aligned} \gamma_0 &= (1 + \theta_1^2 + \theta_2^2)\sigma_a^2, \\ \rho_1 &= \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2}, \\ \rho_2 &= \frac{-\theta_2}{1+\theta_1^2+\theta_2^2}, \\ \rho_k &= 0, \quad k > 2. \end{aligned}$$

A Figura 2.10 apresenta a fac de um processo MA(2) para $\theta_1 = 0,5$, $\theta_2 = -0,3$.

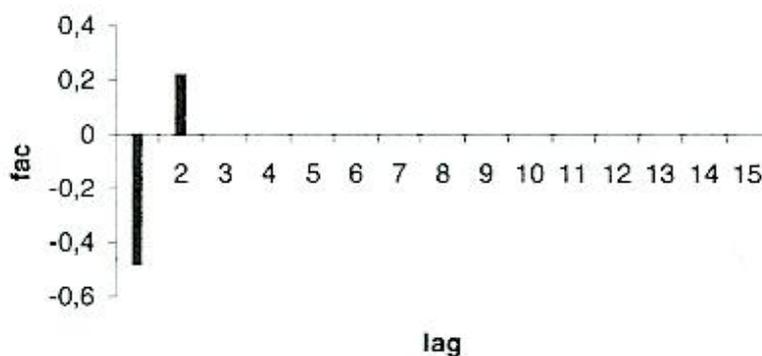


Figura 2.10: Função de autocorrelação de um processo MA(2).

2.3.5 Modelo Auto-Regressivo e de Médias Móveis de Ordem (1,1) - ARMA(1,1)

O modelo ARMA(1,1) é dado por

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + a_t - \theta_1 a_{t-1},$$

Ou equivalentemente,

$$\phi(B)\tilde{Z}_t = \theta(B)a_t,$$

onde $\phi(B) = 1 - \phi_1 B$ e $\theta(B) = 1 - \theta_1 B$.

O processo é estacionário se $-1 < \phi_1 < 1$ e invertível se $-1 < \theta_1 < 1$.

A partir de (2.12) podemos obter

$$\begin{aligned}\gamma_1 &= \phi_1 \gamma_0 + \gamma_{za}(1) - \theta_1 \gamma_{za}(0), \\ \gamma_0 &= \phi_1 \gamma_1 + \gamma_{za}(0) - \theta_1 \gamma_{za}(-1).\end{aligned}$$

Mas,

$$\begin{aligned}\gamma_{za}(1) &= 0, \\ \gamma_{za}(0) &= E[a_t \tilde{Z}_t] = E[a_t(\phi_1 \tilde{Z}_{t-1} + a_t - \theta_1 a_{t-1})] = E[a_t^2] = \sigma_a^2 \\ \gamma_{za}(-1) &= E[a_t \tilde{Z}_{t+1}] = E[a_t(\phi_1 \tilde{Z}_t + a_{t+1} - \theta_1 a_t)] = \phi_1 E[a_t \tilde{Z}_t] + E[a_t a_{t+1}] - \theta_1 E[a_t^2] \\ &= \phi_1 E[a_t^2] - \theta_1 E[a_t^2] = (\phi_1 - \theta_1) \sigma_a^2.\end{aligned}$$

Portanto,

$$\gamma_1 = \phi_1 \gamma_0 - \theta_1 \sigma_a^2 \quad \text{e} \quad \gamma_0 = \theta_1 \gamma_1 + \sigma_a^2 - \theta_1 (\phi_1 - \theta_1) \sigma_a^2,$$

donde segue que

$$\gamma_0 = \frac{(1 + \theta_1^2 - 2\phi_1\theta_1)\sigma_a^2}{1 - \phi_1^2} \quad \text{e} \quad \gamma_1 = \frac{(1 - \phi_1\theta_1)(\phi_1 - \theta_1)\sigma_a^2}{1 - \phi_1^2}.$$

Para valores de $k > 1$, a fac do processo é

$$\rho_k = \phi_1 \rho_{k-1}.$$

A Figura 2.11 ilustra a fac de um processo ARMA(1,1), com $\phi_1 = 0,8$ e $\theta_1 = 0,3$.

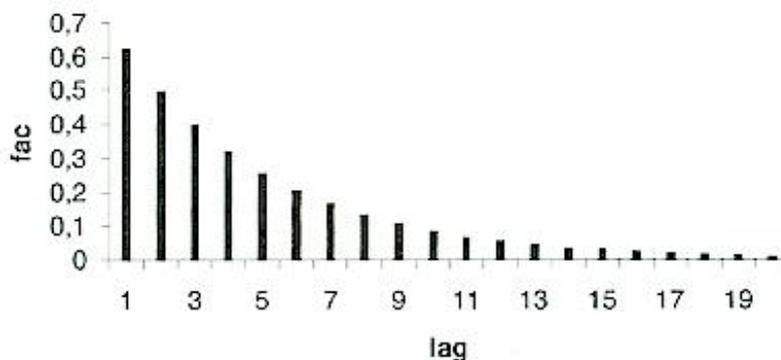


Figura 2.11: Função de autocorrelação de um processo ARMA(1,1).

2.4 Identificação de Modelos ARIMA

A identificação do particular modelo ARIMA a ser ajustado aos dados é uma das etapas mais críticas do ciclo iterativo do método de Box e Jenkins, pois, vários pesquisadores, usando a mesma série, podem identificar modelos diferentes.

O principal objetivo da identificação é encontrar os valores p , d e q do modelo ARIMA(p,d,q), bem como determinar estimativas preliminares dos parâmetros, as quais serão úteis durante o estágio de estimação.

2.4.1 Procedimentos de Identificação

A primeira etapa do processo de identificação consiste em verificar se é necessário transformar a série original, a fim de estabilizar sua variância. Neste sentido, a transformação de Box-Cox é bastante útil

$$Z_t^{(\lambda)} = \begin{cases} \frac{Z_t^\lambda - c}{\lambda} & \lambda \neq 0 \\ \log Z_t & \lambda = 0, \end{cases}$$

onde λ e c são parâmetros a serem estimados.

Para se ter uma noção do tipo de transformação a ser utilizada, pode-se construir um gráfico que traz no eixo das abscissas, médias de subconjuntos de observações da série original e no eixo das ordenadas, a amplitude de cada um desses conjuntos, isto é, se Z_1, Z_2, \dots, Z_k for um tal subconjunto, o gráfico será constituído por pontos da forma (\bar{Z}, w) , onde

$$\bar{Z} = \frac{1}{k} \sum_{i=1}^k Z_i \quad \text{e} \quad w = \max(Z_i) - \min(Z_i).$$

Se w independer de \bar{Z} , os pontos desse gráfico ficarão espalhados em torno de uma reta paralela ao eixo das abscissas; neste caso, não é necessário aplicar nenhuma transformação à série original. Caso w seja diretamente proporcional a \bar{Z} , a transformação logarítmica é adequada. A Figura 2.12 apresenta alguns gráficos que podem ocorrer na prática e os respectivos valores de λ .

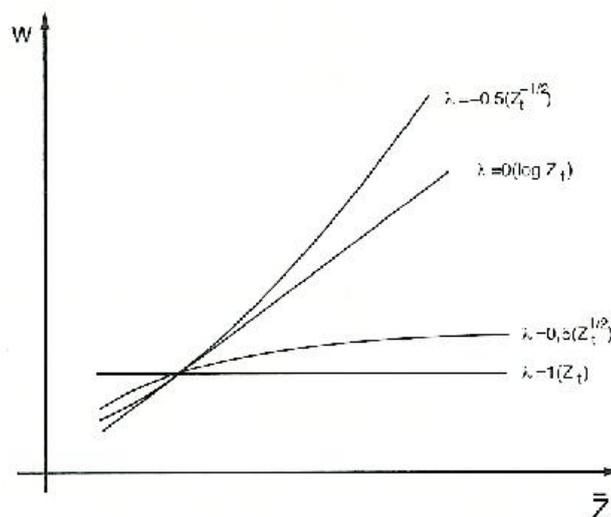


Figura 2.12: Gráficos amplitude \times média, ilustrando alguns valores possíveis de λ .

A segunda etapa do processo de identificação consiste em diferenciar a série, obtida na primeira etapa, até conseguir sua estacionariedade, ou seja, até que o processo $W_t = \Delta^d Z_t$ se reduza a um $ARMA(p, q)$. Uma maneira de saber a quantidade de diferenças, d , necessárias para tornar o processo estacionário consiste em observar quando a fac amostral de W_t decresce rapidamente para zero. Na prática, $d = 0, 1$ ou 2 [Morettin e Toloi, 2004].

A terceira etapa do processo de identificação consiste em analisar o comportamento das autocorrelações e autocorrelações parciais estimadas, as quais devem representar adequadamente as respectivas quantidades teóricas desconhecidas. Através dessa análise, devemos identificar o processo $ARMA(p, q)$. A Tabela 2.1 apresenta um resumo das principais características dos modelos mais usuais.

Na literatura, podemos encontrar outras propostas de identificação de modelos $ARMA(p, q)$. Existem, por exemplo, os métodos baseados em uma função penalizadora, cuja idéia é escolher as ordens k e l que minimizem a seguinte quantidade

$$P(k, l) = \ln \hat{\sigma}_{k,l}^2 + (k + l) \frac{C(N)}{N},$$

Tabela 2.1: Características das fac e facp de um processo ARIMA(p, d, q).

Ordem	(1, d , 0)	(0, d , 1)
comportamento de ρ_k	decai exponencialmente	somente $\rho_1 \neq 0$
comportamento de ϕ_{kk}	somente $\phi_{11} \neq 0$	decaimento exponencial dominante
estimativas iniciais	$\phi_1 = \rho_1$	$\rho_1 = -\theta_1/(1 + \theta_1^2)$
região de admissibilidade	$-1 < \phi_1 < 1$	$-1 < \theta_1 < 1$
Ordem	(2, d , 0)	(0, d , 2)
comportamento de ρ_k	mistura de exponenciais ou ondas senóides amortecidas	somente $\rho_1 \neq 0$ e $\rho_2 \neq 0$
comportamento de ϕ_{kk}	somente $\phi_{11} \neq 0$ e $\phi_{22} \neq 0$	dominada por mistura de exponenciais ou senóides amortecidas
estimativas iniciais	$\begin{cases} \phi_1 = \frac{\rho_1(1-\rho_2)}{1-\rho_1^2}, \\ \phi_2 = \frac{\rho_2-\rho_1^2}{1-\rho_1^2} \end{cases}$	$\begin{cases} \rho_1 = -\frac{\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2}, \\ \rho_2 = -\frac{\theta_2}{1+\theta_1^2+\theta_2^2} \end{cases}$
região de admissibilidade	$\begin{cases} -1 < \phi_2 < 1 \\ \phi_2 - \phi_1 < 1 \\ \phi_2 + \phi_1 < 1 \end{cases}$	$\begin{cases} -1 < \theta_2 < 1 \\ \theta_2 - \theta_1 < 1 \\ \theta_2 + \theta_1 < 1 \end{cases}$
Ordem	(1, d , 1)	
comportamento de ρ_k	decai exponencialmente após o <i>lag</i> 1	
comportamento de ϕ_{kk}	dominada por decaimento exponencial após o <i>lag</i> 1	
estimativas iniciais	$\rho_1 = (1 - \phi_1\theta_1)(\phi_1 - \theta_1)/(1 + \theta_1^2 - 2\phi_1\theta_1)$, $\rho_2 = \rho_1\phi_1$	
região de admissibilidade	$-1 < \phi_1 < 1$, $-1 < \theta_1 < 1$	

onde $\hat{\sigma}_{k,l}^2$ é uma estimativa da variância residual obtida ajustando um modelo ARMA(k, l) às N observações da série e $C(N)$ é uma função do tamanho da série.

Quando o número de parâmetros aumenta, o termo penalizador $(k + l) \frac{C(N)}{N}$ aumenta e a variância diminui. Portanto, minimizar $P(k, l)$ é equivalente a identificar as ordens k e l que equilibrem tal comportamento [Morettin e Tolo, 2004].

A seguir, citaremos alguns procedimentos de identificação baseados em funções penalizadoras particulares.

- Critério de Informação de Akaike

Akaike (1973,1974) propôs que as ordens k e l do modelo deveriam ser escolhidas de modo a minimizar o seguinte critério

$$\text{AIC}(k, d, l) = N \ln \hat{\sigma}_{k,l}^2 + \frac{N}{N-d} 2(k+l+1+\delta_{d0}) + N \ln 2\pi + N, \quad (2.19)$$

onde

$$\delta_{d0} = \begin{cases} 1, & d = 0 \\ 0, & d \neq 0, \end{cases}$$

e $\hat{\sigma}_{k,l}^2$ é o estimador de máxima verossimilhança de σ_a^2 .

Se o interesse for comparar vários modelos, com N fixado, então os dois últimos termos de (2.19) podem ser desconsiderados. Nestes casos, supondo $d = 0$, o critério para determinação das ordens p e q , se reduz a

$$\text{AIC}(k, l) = N \left[\ln \hat{\sigma}_{k,l}^2 + \frac{2}{N} (k+l+2) \right], \quad (2.20)$$

que ainda pode ser reescrito da seguinte maneira

$$\text{AIC}(k, l) = \ln \hat{\sigma}_{k,l}^2 + \frac{2}{N} (k+l), \quad (2.21)$$

já que os valores k e l que minimizam (2.21) são os mesmos que minimizam (2.20), pois

$$\left[\ln \hat{\sigma}_{k,l}^2 + \frac{2}{N} (k+l) \right] < N \left[\ln \hat{\sigma}_{k,l}^2 + \frac{2}{N} (k+l+2) \right].$$

Para os modelos AR(p), o critério AIC se reduz a

$$\text{AIC}(k) = N \ln \hat{\sigma}_k^2 + 2k.$$

Com o intuito de diminuir a probabilidade de selecionar uma ordem maior do que a verdadeira, Hurvich e Tsai (1989) sugeriram uma correção para o AIC, dada por

$$\text{AIC}_c(k) = \text{AIC}(k) + \frac{2(k+1)(k+2)}{N-k+2}.$$

- Critério de Informação Bayesiano

Akaike (1977), Rissanem (1978) e Schwarz (1978), sugerem escolher o modelo cujas ordens k e l minimizam o Critério de Informação Bayesiano, dado por

$$\text{BIC}(k, l) = \ln \hat{\sigma}_{k,l}^2 + (k+l) \frac{\ln N}{N},$$

onde $\hat{\sigma}_{k,l}^2$ é a estimativa de máxima verossimilhança da variância residual do modelo ARMA(k,l).

Para os modelos AR(p), o critério se reduz a

$$\text{BIC}(k) = \ln \hat{\sigma}_k^2 + \frac{k}{N} \ln N.$$

- Critério de Hannan e Quinn

A proposta de Hannan e Quinn (1979) é minimizar a seguinte quantidade

$$\text{HQC}(k, l) = \ln \hat{\sigma}_{k,l}^2 + 2(k+l)c \frac{\ln \ln N}{N}, \quad c > 1.$$

Para modelos AR(p), o critério pode ser escrito da seguinte forma

$$\text{HQC}(k) = \ln \hat{\sigma}_k^2 + 2ck \frac{\ln \ln N}{N}, \quad c > 1.$$

- Critério FPE (*Final Predictor Error*)

Supondo que a série é representada por um modelo AR(p), Akaike (1969) propôs minimizar a seguinte quantidade

$$\text{FPE}(k) = \begin{cases} (1 + \frac{2k}{N}) \hat{\sigma}_k^2, & \mu \text{ conhecido} \\ (1 + \frac{2k+1}{N}) \hat{\sigma}_k^2, & \mu \text{ desconhecido,} \end{cases}$$

onde $\hat{\sigma}_k^2 = c_0 - \sum_{j=1}^k \hat{\phi}_j c_j$.

Pode-se mostrar que o FPE é um estimador assintoticamente não-viciado e consistente para o erro quadrático médio da previsão de Z_{N+1} [Morettin e Tolo, 2004].

- Critério CAT (*Criterion Autoregressive Transfer Function*) - Método de Parzen

Este critério é baseado numa filosofia diferente das anteriores. Primeiramente, deve-se assumir que o verdadeiro modelo é um AR(∞)

$$\pi(B)Z_t = a_t.$$

O próximo passo consiste em estimar a função de transferência $\pi(B)$. Daí, a ordem selecionada \hat{p} é vista como uma aproximação finita ótima para o processo AR(∞).

A seleção de uma função de transferência ótima é feita a partir do valor de k que minimiza a expressão

$$\text{CAT}(k) = \begin{cases} -\left(1 + \frac{1}{N}\right), & k = 0 \\ \frac{1}{N} \sum_{j=1}^k \hat{\sigma}_j^{-2} - \hat{\sigma}_k^{-1}, & k = 1, 2, \dots, \end{cases}$$

onde $\hat{\sigma}_j^2$ é a variância residual estimada para o modelo ajustado de ordem j .

2.4.2 Estimativas Preliminares

A seguir, veremos como obter, a partir das autocorrelações amostrais da série $W_t = \Delta^d Z_t$, estimativas preliminares dos parâmetros do modelo identificado, as quais serão utilizadas como valores iniciais para o processo iterativo de estimação de máxima verossimilhança.

Para processos AR(p) devemos resolver as equações de Yule-Walker, substituindo as autocorrelações teóricas ρ_k por suas estimativas r_j , com isso, obteremos $\hat{\phi}_1, \hat{\phi}_2, \dots, \hat{\phi}_p$.

Uma estimativa inicial da variância residual de um processo AR(p), pode ser obtida substituindo-se, na expressão (2.8), γ_0 por c_0 , os ϕ_j por $\hat{\phi}_j$ e os ρ_j por r_j , obtendo

$$\hat{\sigma}_a^2 = c_0(1 - \hat{\phi}_1 r_1 - \hat{\phi}_2 r_2 - \dots - \hat{\phi}_p r_p).$$

Para processos MA(q), estimativas iniciais para $\theta_1, \theta_2, \dots, \theta_q$, podem ser obtidas, substituindo-se ρ_1, \dots, ρ_q por r_1, \dots, r_q na expressão (2.11) e resolvendo as q equações não-lineares resultantes.

A variância residual pode ser estimada, inicialmente, através da expressão (2.10), substituindo-se γ_0 por c_0 e os θ_j por suas estimativas iniciais, obtendo

$$\hat{\sigma}_a^2 = c_0 / (1 + \hat{\theta}_1^2 + \hat{\theta}_2^2 + \dots + \hat{\theta}_q^2).$$

Para os processos ARMA(p, q), resolvemos as p equações

$$\rho_k = \phi_1 \rho_{k-1} + \dots + \phi_p \rho_{k-p}, \quad k = q + 1, \dots, q + p,$$

substituindo ρ_k por r_k , a fim de obter estimativas preliminares para ϕ_1, \dots, ϕ_p . Em seguida, através da expressão (2.12), obtemos $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_q$ e $\hat{\sigma}_a^2$.

Quando utilizamos o modelo ARIMA, com $\mu_w \neq 0$, isto é,

$$\phi(B)W_t = \theta_0 + \theta(B)a_t,$$

com $\mu_w = \theta_0/(1-\phi_1-\dots-\phi_p)$, podemos obter uma estimativa inicial de θ_0 , substituindo μ_w por \bar{W} e os ϕ_j por $\hat{\phi}_j$, obtendo

$$\hat{\theta}_0 = \bar{W}(1 - \hat{\phi}_1 - \dots - \hat{\phi}_p).$$

2.5 Estimação de Modelos ARIMA

Após identificar um modelo provisório a ser ajustado à série temporal, devemos obter estimativas eficientes para os seus parâmetros.

Vamos denotar por $\boldsymbol{\xi} = (\boldsymbol{\phi}, \boldsymbol{\theta}, \sigma_a^2)$ o vetor com os $p + q + 1$ parâmetros de um modelo ARIMA(p, d, q), onde $\boldsymbol{\phi} = (\phi_1, \dots, \phi_p)$ e $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q)$. A seguinte notação também será útil: $\boldsymbol{\eta} = (\boldsymbol{\phi}, \boldsymbol{\theta})$.

Suponha que a série original $\mathbf{Z} = (Z_1, Z_2, \dots, Z_N)$ tenha sido gerada por um processo ARIMA(p, d, q). A partir daí, considerando d diferenças, podemos gerar uma série W_t estacionária: $\mathbf{W} = (W_1, W_2, \dots, W_n)$, onde $W_t = \Delta^d Z_t$ e $n = N - d$. Com isso, o problema de estimar os parâmetros do modelo ARIMA é equivalente a estimar os parâmetros do modelo ARMA(p, q) estacionário e invertível, representado por

$$a_t = \tilde{W}_t - \phi_1 \tilde{W}_{t-1} - \phi_2 \tilde{W}_{t-2} - \dots - \phi_p \tilde{W}_{t-p} + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q}, \quad (2.22)$$

em que $W_t = \Delta^d Z_t$, $\tilde{W}_t = W_t - \mu_w$ e $\mu_w = E[W_t]$.

Quando $d > 0$, é conveniente considerar $\mu_w = 0$. Caso contrário, μ_w será mais um parâmetro a ser estimado.

A seguir, vamos descrever alguns métodos que possibilitam a obtenção de estimadores para os parâmetros do modelo identificado.

2.5.1 Método dos Momentos

O método dos momentos é um dos métodos de estimação mais simples e antigo. Este método consiste em substituir, nas equações que relacionam as autocorrelações e os

parâmetros do modelo, os momentos teóricos (média, variância e autocorrelação) pelos respectivos momentos amostrais e, em seguida, resolver as equações resultantes.

As estimativas preliminares descritas em 5.4.2 são obtidas através do método dos momentos.

Para o modelo $AR(p)$, o estimador de ϕ , pelo método dos momentos é dado por

$$\hat{\phi}_{MM} = (\hat{\phi}_{1,MM}, \dots, \hat{\phi}_{p,MM})^t = R_p^{-1} \mathbf{r}_p,$$

onde

$$R_p = \begin{bmatrix} 1 & r_1 & r_2 & \dots & r_{p-1} \\ r_1 & 1 & r_1 & \dots & r_{p-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{p-1} & r_{p-2} & r_{p-3} & \dots & 1 \end{bmatrix} \quad \text{e} \quad \mathbf{r}_p = (r_1, r_2, \dots, r_p)^t.$$

Utilizando $\hat{\phi}_{MM}$, podemos também estimar σ_a^2 , através do método dos momentos, obtendo

$$\begin{aligned} \hat{\sigma}_{MM}^2 &= c_0(1 - \hat{\phi}_{1,MM}r_1 - \dots - \hat{\phi}_{p,MM}r_p) \\ &= c_0(1 - \mathbf{r}_p^t \hat{\phi}_{MM}) = c_0(1 - \mathbf{r}_p^t R_p^{-1} \mathbf{r}_p). \end{aligned}$$

Em particular, para $p = 1$, temos que

$$\hat{\phi}_{MM} = \hat{\phi}_{1,MM} = r_1 \quad \text{e} \quad \hat{\sigma}_{MM}^2 = c_0(1 - r_1^2).$$

Para o modelo $MA(q)$, o estimador de $\hat{\theta}$, utilizando o método dos momentos, é obtido resolvendo as equações

$$r_k = \frac{-\hat{\theta}_{k,MM} + \hat{\theta}_{1,MM}\hat{\theta}_{k+1,MM} + \dots + \hat{\theta}_{q-k,MM}\hat{\theta}_{q,MM}}{1 + \hat{\theta}_{1,MM}^2 + \hat{\theta}_{2,MM}^2 + \dots + \hat{\theta}_{q,MM}^2}, \quad k = 1, 2, \dots, q.$$

A variância residual estimada através do método dos momentos é

$$\hat{\sigma}_{MM}^2 = c_0 / (1 + \hat{\theta}_{1,MM}^2 + \hat{\theta}_{2,MM}^2 + \dots + \hat{\theta}_{q,MM}^2).$$

Em particular, para $q = 1$, temos que

$$r_1 = \frac{-\hat{\theta}_{1,MM}}{1 + \hat{\theta}_{1,MM}^2} \quad \text{e} \quad \hat{\sigma}_{MM}^2 = \frac{c_0}{1 + \hat{\theta}_{1,MM}^2}.$$

Para o modelo $ARMA(p,q)$, os parâmetros $\hat{\phi}$ e $\hat{\theta}$, são estimados, através do método dos momentos, em duas etapas:

- (1) estimação de ϕ , através da solução $\hat{\phi}_{MM} = (\phi_{1,MM}, \dots, \phi_{p,MM})$ da seguinte equação

$$r_k = \hat{\phi}_{1,MM}r_{k-1} + \dots + \hat{\phi}_{p,MM}r_{k-p}, \quad k = q + 1, \dots, q + p;$$

- (2) estimação de θ , através da solução $\hat{\theta}_{MM} = (\theta_{1,MM}, \dots, \theta_{q,MM})$ da equação (2.12), utilizando as autocovariâncias amostrais c_k e os estimadores $\hat{\phi}_{MM}$ obtidos na etapa anterior.

Em particular, para $p = q = 1$, obtemos

$$\begin{aligned} r_2 &= \hat{\phi}_{1,MM}r_1 \\ r_1 &= c_1/c_0 = (1 - \hat{\phi}_{1,MM}\hat{\theta}_{1,MM})(\hat{\phi}_{1,MM} - \hat{\theta}_{1,MM})/(1 + \hat{\theta}_{1,MM}^2 - 2\hat{\phi}_{1,MM}\hat{\theta}_{1,MM}). \end{aligned}$$

2.5.2 Método de Máxima Verossimilhança

Vamos denotar por $f(\mathbf{z}|\boldsymbol{\xi})$ a *função densidade (ou de probabilidade) conjunta* de $\mathbf{Z} = (Z_1, Z_2, \dots, Z_N)$. Fixado $\boldsymbol{\xi}$, a função $f(\mathbf{z}|\boldsymbol{\xi})$ associa um determinado valor a cada conjunto de observações \mathbf{z} observado. Agora, quando fixamos \mathbf{z} e variamos $\boldsymbol{\xi}$, obtemos a *função de verossimilhança*, denotada por $L(\boldsymbol{\xi}|\mathbf{z})$. Essa função é de grande importância na teoria de estimação, devido ao "princípio da verossimilhança", que diz o seguinte: dado que o modelo adotado é correto, toda a informação sobre $\boldsymbol{\xi}$ presente na amostra está contida na função de verossimilhança; os outros aspectos dos dados são irrelevantes [Box e Jenkins, 1976]. Em geral, é conveniente trabalharmos com o logaritmo natural de $L(\boldsymbol{\xi}|\mathbf{z})$, denotado por $l(\boldsymbol{\xi}|\mathbf{z})$ e denominado *função de log-verossimilhança*.

Os valores dos parâmetros que maximizam a função de verossimilhança (ou equivalentemente, a função de log-verossimilhança) são chamados *estimadores de máxima verossimilhança* (EMV).

Observe que só é possível calcular os a_t em (2.22) se tivermos valores iniciais para os \tilde{W} 's e para os a 's. Tais valores podem ser obtidos através de dois procedimentos: um condicional e o outro incondicional.

- Procedimento Condicional

O procedimento condicional consiste em substituir os valores iniciais desconhecidos por valores supostamente razoáveis, ou seja, supomos que são dados p valores W_t e q valores a_t , que serão denotados por \mathbf{w}_* e \mathbf{a}_* , respectivamente. A partir daí, os

valores a_1, a_2, \dots, a_n , condicionais à escolha dos valores iniciais \mathbf{w}_* e \mathbf{a}_* , poderão ser calculados através da expressão (2.22).

Supondo que os a_t 's são normalmente distribuídos, a função densidade conjunta de a_1, a_2, \dots, a_n é dada por

$$f(a_1, a_2, \dots, a_n) = \prod_{t=1}^n f(a_t) = (2\pi)^{-n/2} (\sigma_a)^{-n} \exp \left\{ - \sum_{t=1}^n \frac{a_t^2}{2\sigma_a^2} \right\}. \quad (2.23)$$

Dada uma amostra particular \mathbf{w} , a função de verossimilhança associada ao vetor de parâmetros $\boldsymbol{\xi}$ e condicional à escolha de \mathbf{w}_* e \mathbf{a}_* , pode ser obtida a partir das expressões (2.22) e (2.23)

$$L(\boldsymbol{\xi} | \mathbf{w}, \mathbf{w}_*, \mathbf{a}_*) = (2\pi)^{-n/2} (\sigma_a)^{-n} \exp \left\{ - \frac{1}{2\sigma_a^2} \sum_{t=1}^n (\tilde{W}_t - \phi_1 \tilde{W}_{t-1} - \dots - \phi_p \tilde{W}_{t-p} + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q})^2 \right\}.$$

Considerando o logaritmo de L , obtemos

$$l(\boldsymbol{\xi} | \mathbf{w}, \mathbf{w}_*, \mathbf{a}_*) = -\frac{n}{2} \log(2\pi) - n \log(\sigma_a) - \frac{1}{2\sigma_a^2} \sum_{t=1}^n (\tilde{W}_t - \phi_1 \tilde{W}_{t-1} - \dots - \phi_p \tilde{W}_{t-p} + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q})^2.$$

Isto é,

$$l(\boldsymbol{\xi} | \mathbf{w}, \mathbf{w}_*, \mathbf{a}_*) \propto -n \log(\sigma_a) - \frac{1}{2\sigma_a^2} S(\boldsymbol{\eta} | \mathbf{w}, \mathbf{w}_*, \mathbf{a}_*), \quad (2.24)$$

onde

$$\begin{aligned} S(\boldsymbol{\eta} | \mathbf{w}, \mathbf{w}_*, \mathbf{a}_*) &= \sum_{t=1}^n (\tilde{W}_t - \phi_1 \tilde{W}_{t-1} - \dots - \phi_p \tilde{W}_{t-p} + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q})^2 \\ &= \sum_{t=1}^n a_t^2(\boldsymbol{\eta} | \mathbf{w}, \mathbf{w}_*, \mathbf{a}_*). \end{aligned} \quad (2.25)$$

é a soma de quadrados condicional.

Utilizando um asterisco para denotar l e S condicionais a $\mathbf{w}, \mathbf{w}_*, \mathbf{a}_*$, podemos escrever (2.24) e (2.25) da seguinte maneira

$$\begin{aligned} l_*(\boldsymbol{\xi}) &\simeq -n \log(\sigma_a) - \frac{1}{2\sigma_a^2} S_*(\boldsymbol{\eta}), \\ S_*(\boldsymbol{\eta}) &= \sum_{t=1}^n a_t^2(\boldsymbol{\eta} | \mathbf{w}, \mathbf{w}_*, \mathbf{a}_*). \end{aligned}$$

Nosso interesse é maximizar $l_*(\boldsymbol{\xi})$, que é equivalente a minimizar $S_*(\boldsymbol{\eta})$. Portanto, estimadores de máxima verossimilhança serão estimadores de mínimos quadrados e o estudo de $l_*(\boldsymbol{\xi})$ é equivalente ao de $S_*(\boldsymbol{\eta})$.

Os valores iniciais \mathbf{w}_* e \mathbf{a}_* podem ser escolhidos de duas formas:

- (1) um procedimento consiste em substituir os elementos de \mathbf{w}_* e \mathbf{a}_* por suas esperanças. Temos que $E(a_t) = 0$ e, se o modelo não tiver parte determinística, $E(W_t) = 0$. Caso o modelo tenha parte determinística, substituímos cada elemento de \mathbf{w}_* por \bar{w} ;
- (2) se o processo estiver próximo da não-estacionariedade, ou seja, se alguma raiz de $\phi(B)$ estiver próxima do círculo unitário, um procedimento adequado consiste em utilizar a expressão (2.22) para calcular a_{p+1}, a_{p+2}, \dots , colocando os valores anteriores de a_t iguais a zero.

Com isso, teríamos

$$a_{p+1} = \tilde{W}_{p+1} - \phi_1 \tilde{W}_p - \dots - \phi_p \tilde{W}_1 + \theta_1 a_p + \dots + \theta_q a_{p-q+1}$$

e assim por diante.

- Procedimento Não-Condiciona

O procedimento não-condicional consiste em estimar os valores iniciais para os \tilde{W} 's e para os a 's através de um método chamado *backforecasting* ("previsão para o passado"), a fim de gerar valores antes do início da série.

Segundo [Morettin e Toloi, 2004], a função de log-verossimilhança não-condicional pode ser aproximada por

$$l(\boldsymbol{\xi}) \simeq -n \log \sigma_a - \frac{S(\boldsymbol{\eta})}{2\sigma_a^2},$$

onde

$$S(\boldsymbol{\eta}) = S(\boldsymbol{\phi}, \boldsymbol{\theta}) = \sum_{t=-\infty}^n [a_t(\boldsymbol{\eta}, \mathbf{W})]^2 \quad (2.26)$$

é a soma de quadrados não-condicional e

$$[a_t(\boldsymbol{\eta}, \mathbf{W})] = E(a_t | \boldsymbol{\eta}, \mathbf{W}). \quad (2.27)$$

Pode-se obter boas aproximações para os estimadores de máxima verossimilhança através dos estimadores de mínimos quadrados, obtidos minimizando-se a expressão (2.26). Dado $\boldsymbol{\eta}$, o cálculo da soma de quadrados (2.26) é feito através do cálculo das esperanças condicionais (2.27) e através da expressão (2.22). Os valores $[W_{-j}]$ e $[a_{-j}]$, $j = 0, 1, 2, \dots$ são calculados utilizando-se o procedimento *backforecasting*.

Supondo que os W_t 's tenham sido gerados por um processo ARIMA usual

$$\phi(B)W_t = \theta(B)a_t, \quad (2.28)$$

então eles poderiam ter sido, igualmente, gerados pelo processo

$$\phi(F)W_t = \theta(F)e_t, \quad (2.29)$$

onde F é o operador translação para o futuro e e_t é um ruído branco com a mesma variância que a_t [Box e Jenkins, 1976]. A representação (2.28) é chamada *forma forward* do processo e a representação (2.29) é denominada *forma backward*. Assim, fazer previsões antes que a série se inicie é equivalente a prever a série reversa.

2.5.3 Variância dos Estimadores

A precisão dos estimadores encontradas deve ser avaliada através da construção de intervalos de confiança para os parâmetros. Considerando o vetor de parâmetros $\boldsymbol{\eta} = (\boldsymbol{\phi}, \boldsymbol{\theta})$, cuja ordem é $p + q$.

Supondo n suficientemente grande, os estimadores de máxima verossimilhança têm uma distribuição assintótica normal, isto é,

$$\hat{\boldsymbol{\eta}} \xrightarrow{\mathcal{D}} \mathcal{N}_{p+q}(\boldsymbol{\eta}, \mathbf{V}),$$

$$\mathbf{V} = 2\sigma_a^2 \begin{bmatrix} \frac{\partial^2 S(\boldsymbol{\eta})}{\partial \eta_1^2} & \cdots & \frac{\partial^2 S(\boldsymbol{\eta})}{\partial \eta_1 \partial \eta_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 S(\boldsymbol{\eta})}{\partial \eta_k \partial \eta_1} & \cdots & \frac{\partial^2 S(\boldsymbol{\eta})}{\partial \eta_k^2} \end{bmatrix}. \quad (2.30)$$

Além disso, o estimador de máxima verossimilhança de σ_a^2 é dado por

$$\hat{\sigma}_a^2 = \frac{S(\hat{\boldsymbol{\eta}})}{n}$$

e, para n suficientemente grande, $\hat{\sigma}_a^2$ e $\hat{\boldsymbol{\eta}}$ são não-correlacionados [Morettin e Toloi, 2004].

As estimativas das variâncias dos estimadores e covariâncias entre os estimadores são obtidas substituindo-se σ_a^2 em (2.30) por $\hat{\sigma}_a^2$ e calculando-se as derivadas $\frac{\partial^2 S(\boldsymbol{\eta})}{\partial \eta_i \partial \eta_j}$, numericamente. Utilizando as estimativas das variâncias, podemos obter intervalos de confiança para os parâmetros $\eta_i, i = 1, 2, \dots, p + q$.

2.6 Diagnóstico de Modelos ARIMA

Após identificar o modelo e estimar seus parâmetros, devemos verificar se ele representa, satisfatoriamente, os dados observados. Esta verificação pode ser feita através de uma técnica chamada *superajustamento*, a qual consiste em estimar um modelo com parâmetros extras e examinar, primeiramente, se eles são significativos e, em seguida, se a inclusão dos mesmos diminui significativamente a variância residual. Para tanto, precisamos analisar os resíduos do modelo ajustado. Se o modelo ajustado

$$\phi(B)W_t = \theta(B)a_t,$$

com $W_t = \Delta^d Z_t$, for verdadeiro, então os "erros verdadeiros" $a_t = \theta^{-1}(B)\phi(B)W_t$ serão um ruído branco [Morettin e Toloi, 2004].

A seguir, descreveremos alguns testes de diagnósticos de um modelo ajustado a uma série temporal, baseados nas autocorrelações estimadas dos resíduos.

2.6.1 Teste de Autocorrelação Residual

Após estimar ϕ e θ , calculamos os *resíduos estimados* (ou simplesmente *resíduos*) através da seguinte expressão

$$\hat{a}_t = \hat{\theta}^{-1}(B)\hat{\phi}(B)W_t.$$

Se o modelo ajustado for adequado, os resíduos estimados \hat{a}_t deverão estar próximos dos resíduos verdadeiros a_t , conseqüentemente, deverão ser aproximadamente não-correlacionados. Ou seja, denotando por \hat{r}_k as autocorrelações dos resíduos \hat{a}_t , deveríamos ter $\hat{r}_k \simeq 0$. Em particular, supondo que o modelo ajustado é adequado, deveríamos ter, aproximadamente,

$$\hat{r}_k \sim \mathcal{N}(0, 1/n).$$

O cálculo das autocorrelações \hat{r}_k é feito através da expressão

$$\hat{r}_k = \frac{\sum_{t=k+1}^n \hat{a}_t \hat{a}_{t-k}}{\sum_{t=1}^n \hat{a}_t^2}.$$

Para valores "grandes" de k , podemos obter uma indicação de uma possível quebra de comportamento de ruído branco em a_t , comparando \hat{r}_k com os limites $\pm 2/\sqrt{n}$ [Morettin e Toloi, 2004].

2.6.2 Teste de Box-Pierce

Box e Pierce (1970) propuseram um teste bastante útil para indicar se os valores das autocorrelações dos resíduos estimados são muito altos. Se o modelo for apropriado, a estatística

$$Q(K) = n(n+2) \sum_{j=1}^K \frac{\hat{r}_j^2}{n-j}$$

terá, aproximadamente, uma distribuição χ^2 com $K - p - q$ graus de liberdade. Para valores grandes de $Q(K)$ rejeitamos a hipótese de ruído branco para os resíduos.

2.6.3 Teste da Autocorrelação Cruzada

Novos termos de médias móveis podem ser incluídos no modelo, a partir da verificação das autocorrelações \hat{r}_k . Por exemplo, se $|\hat{r}_5| > 2/\sqrt{n}$, então um termo $\theta_5 a_{t-5}$ deve ser inserido no modelo. Uma maneira alternativa, consiste em investigar a *função de correlação cruzada* (fcc), baseada na correlação cruzada entre valores passados da série e o valor presente do ruído, e definida por

$$s_k = \frac{\sum a_t (Z_{t-k} - \bar{Z})}{\sqrt{\sum a_t^2 \sum (Z_t - \bar{Z})^2}}, \quad k = 1, 2, 3, \dots$$

Como os verdadeiros a_t são desconhecidos, utilizamos os resíduos estimados \hat{a}_t e substituímos s_k por

$$\hat{s}_k = \frac{\sum \hat{a}_t (Z_{t-k} - \bar{Z})}{\sqrt{\sum \hat{a}_t^2 \sum (Z_t - \bar{Z})^2}}, \quad k = 1, 2, 3, \dots$$

Se o modelo for apropriado, então a_t e Z_{t-k} devem ser não-correlacionados, para $k \geq 1$, ou seja,

$$\text{Cov}[a_t, Z_{t-k}] = \gamma_{az}(k) = 0, \quad k \geq 1.$$

Daí, se para um certo k_0 , s_{k_0} assumir um valor "grande", o modelo deverá ser considerado inadequado.

Se $|s_k| > 2/\sqrt{n}$, então $\gamma_{az}(k)$ é significativamente diferente de zero. É razoável, portanto, para k suficientemente grande, julgar s_k significativa quando $|\hat{s}_k| > 2/\sqrt{n}$ [Morettin e Toloi, 2004].

Os resíduos podem ser utilizados para modificar o modelo da seguinte maneira: se os resíduos b_t do modelo ajustado

$$\phi_0(B)\Delta^{d_0}Z_t = \theta_0(B)b_t \quad (2.31)$$

não forem aleatórios, podemos utilizar o método de identificação visto na seção 5.4, para descrevê-los através do modelo

$$\bar{\phi}(B)\Delta^{\bar{d}}b_t = \bar{\theta}(B)a_t. \quad (2.32)$$

Daí, substituindo (2.32) em (2.31), obtemos um novo modelo que deverá ser ajustado aos dados

$$\phi_0(B)\bar{\phi}(B)\Delta^{d_0}\Delta^{\bar{d}}Z_t = \theta_0(B)\bar{\theta}(B)a_t,$$

cujos resíduos são aleatórios. Este ciclo de identificação, estimação e verificação deve ser repetido até que um modelo adequado seja encontrado.

2.7 Previsão com Modelos ARIMA

Nas seções 5.4, 5.5 e 5.6 seguimos as etapas do ciclo iterativo de identificação, estimação e diagnóstico, com o objetivo de construir um modelo ARIMA(p, d, q) que representasse adequadamente os dados observados. Agora, vamos utilizar esse modelo para fazer previsões.

Supondo que temos as observações $\dots, Z_{t-2}, Z_{t-1}, Z_t$, até o instante t , nosso interesse é prever um valor Z_{t+h} , $h \geq 1$. Dizemos que t é a *origem* das previsões e h o *horizonte* e denotamos por $\hat{Z}_t(h)$ a previsão de Z_{t+h} (ver Figura 2.13).

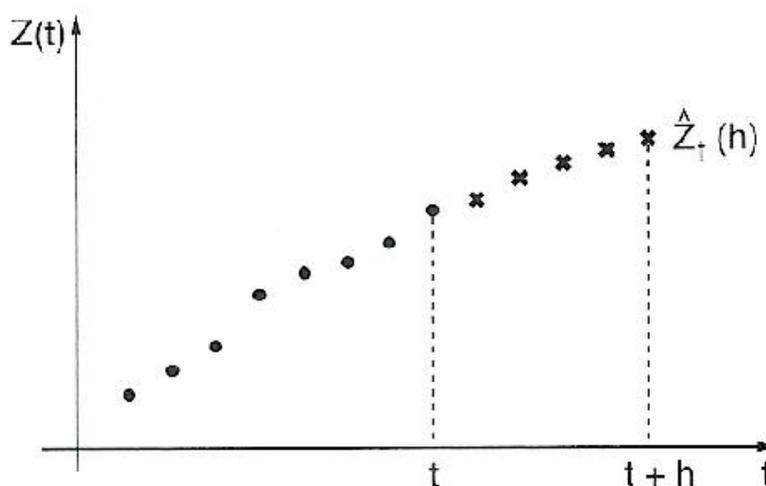


Figura 2.13: Observações de uma série temporal com previsões de origem t e horizonte h .

Primeiramente, vamos assumir que $W_t = (1 - B)^d Z_t$ é estacionário e invertível e que os parâmetros do modelo são conhecidos.

Substituindo t por $t + h$ nas expressões (2.15), (2.16) e (2.17), obtemos o modelo ARIMA(p, d, q) nas três formas básicas:

(i) *forma de equação de diferenças*

$$Z_{t+h} = \varphi_1 Z_{t+h-1} + \cdots + \varphi_{p+d} Z_{t+h-p-d} - \theta_1 a_{t+h-1} - \cdots - \theta_q a_{t+h-q} + a_{t+h}; \quad (2.33)$$

(ii) *forma de choques aleatórios*

$$Z_{t+h} = a_{t+h} + \psi_1 a_{t+h-1} + \psi_2 a_{t+h-2} + \cdots = \sum_{j=0}^{\infty} \psi_j a_{t+h-j} = \sum_{j=-\infty}^{t+h} \psi_{t+h-j} a_j; \quad (2.34)$$

(iii) *forma invertida*

$$Z_{t+h} = \pi_1 Z_{t+h-1} + \pi_2 Z_{t+h-2} + \cdots + a_{t+h} = \sum_{j=1}^{\infty} \pi_j Z_{t+h-j} + a_{t+h}. \quad (2.35)$$

2.7.1 Previsão de Erro Quadrático Médio (EQM) mínimo

Supondo que $\hat{Z}_t(h)$ seja uma função *linear* das observações até o instante t , então, por (2.34), também será uma função de a_t, a_{t-1}, \dots

Indicando a melhor previsão por

$$\hat{Z}_t(h) = \psi_h^* a_t + \psi_{h+1}^* a_{t-1} + \psi_{h+2}^* a_{t-2} + \cdots = \sum_{j=0}^{\infty} \psi_{h+j}^* a_{t-j},$$

nosso objetivo é encontrar os pesos ψ_j^* que minimizem o EQM de previsão, dado por

$$E[Z_{t+h} - \hat{Z}_t(h)]^2 = E \left[\sum_{j=0}^{\infty} \psi_j a_{t+h-j} - \sum_{j=0}^{\infty} \psi_{h+j}^* a_{t-j} \right]^2. \quad (2.36)$$

Observando que $\sum_{j=0}^{\infty} \psi_j a_{t+h-j} = \sum_{j=-h}^{\infty} \psi_{h+j} a_{t-j}$, temos que o *erro de previsão* é

$$e_t(h) = Z_{t+h} - \hat{Z}_t(h) = \psi_0 a_{t+h} + \psi_1 a_{t+h-1} + \cdots + \psi_{h-1} a_{t+1} - \sum_{j=0}^{\infty} (\psi_{h+j} - \psi_{h+j}^*) a_{t-j}.$$

Substituindo essa última expressão em (2.36) e usando o fato de que os a_t são não-correlacionados, podemos reescrever o EQM de previsão da seguinte forma

$$E[e_t(h)]^2 = (1 + \psi_1^2 + \cdots + \psi_{h-1}^2) \sigma_a^2 + \sum_{j=0}^{\infty} (\psi_{h+j} - \psi_{h+j}^*)^2 \sigma_a^2,$$

que é minimizado se $\psi_{h+j}^* = \psi_{h+j}$, $j = 0, 1, 2, \dots$, h fixo. Assim, a previsão de EQM mínimo é dada por

$$\hat{Z}_t(h) = \psi_h a_t + \psi_{h+1} a_{t-1} + \psi_{h+2} a_{t-2} + \dots = \sum_{j=0}^{\infty} \psi_{h+j} a_{t-j}.$$

Conseqüentemente, o erro de previsão é

$$e_t(h) = a_{t+h} + \psi_1 a_{t+h-1} + \dots + \psi_{h-1} a_{t+1}.$$

Logo,

$$Z_{t+h} = e_t(h) + \hat{Z}_t(h), \quad h \geq 1.$$

Utilizando a notação

$$[Z_{t+h}] = E[Z_{t+h} | Z_t, Z_{t-1}, \dots],$$

temos que:

- (a) a previsão de EQM mínimo é a esperança condicional de Z_{t+h} , dadas as observações passadas da série, ou seja, $\hat{Z}_t(h) = [Z_{t+h}]$;
- (b) $[e_t(h)] = 0$ e a variância do erro de previsão é dada por

$$V(h) = (1 + \psi_1^2 + \psi_2^2 + \dots + \psi_{h-1}^2) \sigma_a^2; \quad (2.37)$$

- (c) os erros de previsão a um passo são não-correlacionados, pois

$$e_t(1) = Z_{t+1} - \hat{Z}_t(1) = a_{t+1};$$

- (d) os erros de previsão para intervalos de tempo maiores que um são correlacionados, bem como os erros de previsão para o mesmo horizonte h , de diferentes origens t e $t - j$ [Morettin e Toloi, 2004].

2.7.2 Formas Básicas de Previsão

A previsão $\hat{Z}_t(h)$ pode ser calculada de três formas, utilizando as diversas representações do modelo ARIMA.

(i) *Previsão utilizando a equação de diferenças*

Considerando a esperança condicional em (2.33), temos que

$$\begin{aligned}\hat{Z}_t(h) &= \varphi_1[Z_{t+h-1}] + \cdots + \varphi_{p+d}[Z_{t+h-p-d}] \\ &\quad - \theta_1[a_{t+h-1}] - \cdots - \theta_q[a_{t+h-q}] + [a_{t+h}], \quad h \geq 1,\end{aligned}$$

onde devemos usar os seguintes fatos:

$$[Z_{t+k}] = \hat{Z}_t(k), \quad k > 0,$$

$$[Z_{t+k}] = Z_{t+k}, \quad k \leq 0,$$

$$[a_{t+k}] = 0, \quad k > 0,$$

$$[a_{t+k}] = a_{t+k}, \quad k \leq 0,$$

(ii) *Previsão utilizando a forma de choques aleatórios*

Tomando a esperança condicional em (2.34), obtemos

$$\hat{Z}_t(h) = \psi_1[a_{t+h-1}] + \psi_2[a_{t+h-2}] + \cdots + \psi_{h-1}[a_{t+h}] + \psi_h[a_t] + \cdots + [a_{t+h}].$$

(iii) *Previsão utilizando a forma invertida*

Considerando a esperança condicional em (2.35), temos que

$$\hat{Z}_t(h) = \pi_1[Z_{t+h-1}] + \pi_2[Z_{t+h-2}] + \cdots + [a_{t+h}].$$

2.7.3 Equação de Previsão

Por 2.7.2(i), a equação de previsão, vista como uma função de h , com origem t fixa, satisfaz a equação de diferenças

$$\hat{Z}_t(h) = \sum_{i=1}^{p+d} \varphi_i \hat{Z}_t(h-1), \quad h > q,$$

ou ainda,

$$\varphi(B)\hat{Z}_t(h) = (1-B)^d \phi(B)\hat{Z}_t(h) = 0, \quad h > q,$$

com $\varphi(B)$ operando sobre h .

A função $\hat{Z}_t(h)$, para $h > q - p - d$, consiste de uma mistura de polinômios, exponenciais e senóides amortecidas [Morettin e Tolo, 2004].

2.7.4 Atualização das Previsões

Calculando as previsões de Z_{t+h+1} a partir de duas origens $t+1$ e t , obtemos, respectivamente,

$$\hat{Z}_{t+1}(h) = \psi_h a_{t+1} + \psi_{h+1} a_t + \psi_{h+2} a_{t-1} + \dots \quad (2.38)$$

e

$$\hat{Z}_t(h+1) = \psi_{h+1} a_t + \psi_{h+2} a_{t-1} + \dots \quad (2.39)$$

Subtraindo (2.39) de (2.38), temos que

$$\hat{Z}_{t+1}(h) = \hat{Z}_t(h+1) + \psi_h a_{t+1}.$$

Portanto, quando um novo dado for observado, podemos atualizar a previsão de Z_{t+h+1} , feita no instante t . Essa atualização consiste em prever o valor de Z_{t+h+1} , na origem $t+1$, adicionando à $\hat{Z}_t(h+1)$ um múltiplo do erro de previsão $a_{t+1} = Z_{t+1} - \hat{Z}_t(1)$.

2.7.5 Intervalos de Confiança

Para obtermos um intervalo de confiança para Z_{t+h} , vamos supor que os erros satisfazem as seguintes condições:

$$\begin{aligned} E[a_t] &= 0, \forall t, \\ \text{Var}[a_t] &= \sigma_a^2, \forall t, \\ E[a_t a_s] &= 0, s \neq t, \\ a_t &\sim \mathcal{N}(0, \sigma_a^2), \forall t. \end{aligned}$$

Dado que conhecemos os valores passados e presente da série, $Z_t, Z_{t-1}, Z_{t-2}, \dots$, a distribuição condicional de Z_{t+h} será $\mathcal{N}(\hat{Z}_t(h), V(h))$, onde $V(h)$ é a variância do erro de previsão, calculada através da expressão (2.37).

Assim, temos que

$$U = \frac{Z_{t+h} - \hat{Z}_t(h)}{\sqrt{V(h)}} \sim \mathcal{N}(0, 1)$$

Portanto, fixado o coeficiente de confiança γ , é possível encontrar um valor u_γ na distribuição se U , tal que $P(-u_\gamma < U < u_\gamma) = \gamma$. Em outras palavras,

$$\left[\hat{Z}_t(h) - u_\gamma \sqrt{V(h)}; \hat{Z}_t(h) + u_\gamma \sqrt{V(h)} \right] \quad (2.40)$$

é um intervalo (aleatório) que contém Z_{t+h} com probabilidade γ .

O cálculo de $V(h)$ é feito, substituindo-se σ_a^2 por sua estimativa $\hat{\sigma}_a^2$ (obtida na etapa de estimação dos parâmetros do modelo), ou seja,

$$V(h) = (1 + \psi_1^2 + \psi_2^2 + \cdots + \psi_{h-1}^2) \hat{\sigma}_a^2 = \hat{\sigma}_a^2 \left(1 + \sum_{j=1}^{h-1} \psi_j^2 \right). \quad (2.41)$$

Substituindo (2.41) em (2.40), obtemos

$$\hat{Z}_t(h) - u_\gamma \hat{\sigma}_a \sqrt{1 + \sum_{j=1}^{h-1} \psi_j^2} \leq Z_{t+h} \leq \hat{Z}_t(h) + u_\gamma \hat{\sigma}_a \sqrt{1 + \sum_{j=1}^{h-1} \psi_j^2}.$$

Capítulo 3

Séries Temporais Envolvendo Dados Angulares

Em diversas áreas de conhecimento aparecem dados da forma $(\theta_1, t_1), \dots, (\theta_n, t_n)$, onde $\theta_1, \dots, \theta_n$ consistem de direções em tempos t_1, t_2, \dots, t_n [Mardia e Jupp, 2000]. Em outras palavras, esses dados constituem uma série temporal de dados angulares (circulares ou direcionais).

São exemplos de séries temporais de dados circulares:

- 1- Direção de ventos e correntes marinhas;
- 2- Direção de migrações de animais;
- 3- Posição da broca durante a perfuração de um poço petrolífero.

3.1 Modelos

Existem diversos modelos para descrição e análise de séries temporais de dados angulares, muitos deles construídos a partir de modelos para séries temporais lineares. A escolha do modelo mais adequado é feita em várias etapas: escolha de uma classe geral de modelos; identificação; estimação dos parâmetros do modelo identificado; ajuste e, por fim, segue a etapa de previsão.

A seguir faremos uma descrição de quatro classes de modelos para séries temporais de dados angulares, propostos por [Fisher e Lee, 1994].

3.1.1 Processo Gaussiano Transformado

Seja $\{(X_t, Y_t); t \in T\}$ um processo no plano, onde T é um conjunto de índices, então a projeção radial sobre o círculo unitário gera um processo correspondente Θ_t sobre esse círculo, definido por

$$X_t = R_t \cos \Theta_t, \quad Y_t = R_t \sin \Theta_t.$$

Quando $\{(X_t, Y_t); t \in T\}$ é um processo Gaussiano bivariado estacionário então Θ_t tem uma distribuição Gaussiana transformada. Além disso, se $\{X_t; t \in T\}$ e $\{Y_t; t \in T\}$ são duas realizações independentes de um processo Gaussiano estacionário de média zero e variância unitária então Θ_t tem distribuição uniforme circular [Fisher e Lee, 1994].

O ajuste de tais modelos apresenta um problema de falta de dados, já que a parte radial $\{R_t; t \in T\}$ de um processo Gaussiano transformado não é observada. No entanto, esse problema pode ser contornado através da utilização do algoritmo EM, o qual será abordado na seção 3.2.

A estrutura de correlação de um processo $\{\Theta_t; t \in T\}$ pode ser quantificada através de uma medida de correlação entre duas variáveis circulares Θ_t e Φ_t , denominada *coeficiente de correlação*, introduzido por [Fisher e Lee, 1983] e definido por

$$\rho_\tau = \frac{E[\sin(\Theta_1 - \Theta_2)\sin(\Phi_1 - \Phi_2)]}{\sqrt{E[\sin^2(\Theta_1 - \Theta_2)]E[\sin^2(\Phi_1 - \Phi_2)]}},$$

onde (Θ_1, Φ_1) e (Θ_2, Φ_2) são realizações independentes de (Θ, Φ) .

De modo análogo ao caso linear, pode-se mostrar que (vide Apêndice A)

$$-1 \leq \rho_\tau \leq 1. \quad (3.1)$$

Além disso, se Θ e Φ forem independentes, então $\rho_\tau = 0$. A seguir veremos um resultado importante envolvendo correlação circular.

Teorema 3.1 *Sejam (X_1, Y_1) e (X_2, Y_2) vetores aleatórios independentes com uma distribuição normal bivariada com variâncias iguais a σ^2 e correlação ρ . Sejam Θ_1 e Θ_2 variáveis aleatórias angulares definidas por*

$$(X_i, Y_i) = R_i(\cos \Theta_i, \sin \Theta_i), \quad i = 1, 2.$$

Então a correlação circular entre Θ_1 e Θ_2 é dada por

$$\rho_\tau = \frac{\pi^2}{16} \rho^2 (1 - \rho^2)^2 \left\{ {}_2F_1 \left(\frac{3}{2}, \frac{3}{2}, 2; \rho^2 \right) \right\}^2,$$

onde ${}_2F_1$ é a função hipergeométrica (vide Apêndice B).

A partir desse resultado pode-se definir a função de autocorrelação do processo $\{\Theta_t; t \in T\}$ da seguinte maneira

$$\rho_\tau(k) = \frac{\pi^2}{16} \rho(k)^2 (1 - \rho^2(k))^2 \left\{ {}_2F_1 \left(\frac{3}{2}, \frac{3}{2}, 2; \rho^2(k) \right) \right\}^2, \quad (3.2)$$

onde $\rho(k)$ é a função de autocorrelação comum dos processos X_t e Y_t .

3.1.2 Processo Arqueado

Seja X_t uma série temporal univariada de dados lineares. O arqueamento de X_t em torno do círculo unitário gera uma série temporal arqueada Θ_t definida por

$$\Theta_t = X_t \pmod{2\pi},$$

ou seja, Θ_t é o resto da divisão de X_t por 2π . Assim, um processo linear $\{X_t; t \in T\}$, que dá origem a um processo arqueado $\{\Theta_t; t \in T\}$, pode ser decomposto da seguinte maneira

$$X_t = \Theta_t + 2\pi k_t,$$

onde k_t é um inteiro não observado. Desta forma, o ajuste desse modelo também apresenta um problema de falta de dados, que poderá ser solucionado através do uso do algoritmo EM.

O arqueamento de um processo $\{X_t; t \in T\}$ auto-regressivo AR(p) produz um processo auto-regressivo arqueado (*wrapped*), denotado por WAR(p).

Segundo [Fisher e Lee, 1983], se (X, Y) segue uma distribuição normal bivariada com variâncias σ_X^2 e σ_Y^2 e correlação ρ , então o coeficiente de correlação circular ρ_τ entre $\Theta = X \pmod{2\pi}$ e $\Phi = Y \pmod{2\pi}$ é dado por

$$\rho_\tau = \frac{\sinh(2\rho\sigma_X\sigma_Y)}{\sqrt{\sinh(2\sigma_X^2)\sinh(2\sigma_Y^2)}}.$$

Daí, segue que, se $\{X_t; t \in T\}$ é um processo AR(p), então a função de autocorrelação circular do processo WAR(p) $\{\Theta_t; t \in T\}$ é dada por

$$\rho_\tau = \frac{\sinh[2\rho_k\sigma^2/(1 - \phi_1\rho_1 - \phi_p\rho_p)]}{\sinh[2\sigma^2/(1 - \phi_1\rho_1 - \phi_p\rho_p)]},$$

onde ρ_k é a autocorrelação de defasagem k do processo $\{X_t; t \in T\}$, $\sigma^2/(1 - \phi_1\rho_1 - \phi_p\rho_p)$ é a variância do processo e $\sigma^2, \phi_1, \dots, \phi_p$ são os seus parâmetros.

3.1.3 Processos Baseados em Funções de Ligação

Uma forma alternativa de se construir um processo angular, a partir de um processo linear, consiste em utilizar uma função de ligação, isto é, uma função g bijetiva e monótona que projete a reta real sobre o intervalo $(-\pi, \pi)$ e de modo que $g(0) = 0$. Duas funções de ligação bastante utilizadas são

$$g(x) = 2\text{tg}^{-1}(x)$$

e a função obtida a partir da *probit*, que é

$$g(x) = 2x\{\Phi(x) - 0,5\},$$

onde $\Phi(\cdot)$ é a função de distribuição da normal padrão.

Se g é uma função de ligação e X uma variável linear então $\Theta = g(X)$ é uma variável angular e, de modo recíproco, para uma variável angular Θ , $X = g^{-1}(\Theta)$ é uma variável linear.

A seguir veremos dois modelos circulares construídos a partir de um modelo linear ARMA(p, q): o *modelo com ligação direta* - LARMA, onde os ângulos transformados seguem um modelo linear ARMA e o *modelo com ligação inversa* - IAR, utilizando médias condicionais e a distribuição von Mises [Fisher, 1993].

3.1.3.1 Processo com Ligação Direta - LARMA

Sendo $\{X_t; t \in T\}$ um processo linear, g uma função de ligação e μ um ponto no círculo então o processo angular correspondente $\{\Theta_t; t \in T\}$ é definido por

$$\Theta_t = g(X_t) + \mu.$$

Dizemos que um processo angular estacionário $\{\Theta_t; t \in T\}$ com média direcional μ é um processo auto-regressivo e de médias móveis com ligação (LARMA) quando o seu processo linear com ligação $X_t = g^{-1}(\Theta_t - \mu)$ for um processo ARMA(p, q). Neste caso, a autocorrelação circular de Θ_t é dada por

$$\rho_\tau(k) = \rho_\tau\{g(X_t), g(X_{t+k})\},$$

onde ρ_τ é o coeficiente de correlação circular definido na seção 3.1.1.

As principais vantagens de um modelo linear com ligação são as seguintes:

- Se o processo $\{\Theta_t; t \in T\}$ for estacionário então o processo resultante $\{X_t; t \in T\}$ também é;
- Os parâmetros são facilmente estimados utilizando-se um *software*.

3.1.3.2 Processo com Ligação Inversa - IAR

Para processos AR, pode-se definir um modelo alternativo através do uso de distribuições condicionais.

Se g é uma função de ligação, μ é um ponto do círculo, $\kappa > 0$ e $\omega_1, \dots, \omega_p$ são coeficientes reais, então o processo com ligação inversa - IAR(p) é definido por

$$\Theta_t | (\theta_{t-1}, \dots, \theta_{t-p}) \sim \mathcal{VM}(\mu_t, \kappa)$$

onde

$$\mu_t = \mu + g\{\omega_1 g^{-1}(\theta_{t-1} - \mu) + \dots + \omega_p g^{-1}(\theta_{t-p} - \mu)\},$$

ou seja, Θ_t dado $\theta_{t-1}, \dots, \theta_{t-p}$ segue uma distribuição von Mises com média direcional μ_t e parâmetro de concentração κ constante. Este processo é bastante utilizado para modelar séries dispersas ($\kappa < 2$).

3.2 Seleção do Modelo

Em geral, a escolha de uma classe de modelos deve ser feita levando-se em consideração as características observáveis dos dados.

No contexto de séries temporais envolvendo dados angulares temos quatro classes de modelos à nossa disposição (vide seção 3.1), diferentemente do caso de séries temporais lineares, em que usávamos uma única família de modelos, a saber, os autorregressivos integrados e de médias móveis (ARIMA).

A seguir descreveremos algumas características específicas de cada classe de modelos, as quais auxiliarão na escolha do modelo mais adequado.

A principal característica do modelo Gaussiano transformado é a distribuição uniforme de suas marginais. Caso o processo linear que está sendo transformado tenha correlação alta, então o processo Gaussiano transformado tenderá a ocupar diferentes arcos do círculo correspondendo ao grau de variações observado no processo linear

altamente correlacionado. Por outro lado, para graus baixos de correlação, o processo circular tenderá a se espalhar sobre todo o círculo.

A aparência do processo arqueado é afetada tanto pela correlação quanto pela variância. Se a variância for grande, o processo circular tenderá a se espalhar uniformemente sobre o círculo, enquanto que, para valores pequenos da variância, o processo tenderá a ocupar um arco do círculo.

No caso do processo com ligação direta, valores altos do processo linear tendem a ser transformados em torno do arco $\mu + \pi$, gerando um "vazio" na forma como os dados se distribuem em torno do círculo.

Por fim, para o processo com ligação inversa, quanto menor o valor do parâmetro de concentração κ , os dados estarão espalhados mais uniformemente sobre o círculo e, quanto maior for ω , os dados tenderão a permanecer sobre um pequeno arco em torno da média direcional μ .

Com base nessas características, recomenda-se que, se os dados angulares estiverem distribuídos uniformemente sobre o círculo, então o modelo Gaussiano transformado ou o modelo arqueado são adequados, sendo que para este último, o processo linear arqueado deverá apresentar variância alta. Se, pelo contrário, os dados tendem a se agrupar em torno de um arco, então os modelos com ligação direta ou inversa são mais adequados.

Vale salientar que todos as quatro classes de modelos são bastante flexíveis e capazes de modelar uma diversidade de comportamento de séries temporais de dados angulares.

3.3 Identificação do Modelo

Após selecionar uma classe de modelos, devemos identificar um modelo particular pertencente a essa classe. Uma ferramenta bastante útil nessa etapa é o correlograma amostral da série. No caso dos modelos Gaussiano transformado, arqueado e com ligação inversa, o correlograma amostral pode ser calculado utilizando-se a seguinte expressão para as autocorrelações amostrais

$$\hat{\rho}_\tau(k) = \frac{\det \left(\sum_{t=1}^{\tau-k} X_t X_{t-\tau}' \right)}{\sqrt{\det \left(\sum_{t=1}^{\tau-k} X_t X_t' \right) \det \left(\sum_{t=k+1}^{\tau} X_t X_t' \right)}}, \quad (3.3)$$

onde $X_t = (\cos \theta_t, \sin \theta_t)'$. A idéia é comparar o correlograma amostral com as autocorrelações teóricas de vários modelos, a fim de identificar um modelo experimental que se ajuste satisfatoriamente aos dados.

Para o processo Gaussiano transformado, pode-se estimar a função de autocorrelação $\rho(k)$ do processo linear correspondente, a partir da resolução da equação (3.2) com $\rho_\tau(k)$ substituído por sua estimativa $\hat{\rho}_\tau(k)$, calculada por (3.3).

Para o processo arqueado, a autocorrelação pode ser estimada através da seguinte expressão

$$\rho_\tau(k) = \sinh(2c_0 \hat{\rho}_k) / \sinh(2c_0)$$

onde c_0 é a variância do processo linear, a qual poderá ser estimada através da relação $\bar{R} = \exp(-c_0/2)$, em que \bar{R} é o comprimento médio resultante de θ_t .

No caso do modelo com ligação direta, o cálculo da média direcional da série circular é uma estimativa preliminar $\hat{\mu}$ de μ . Utilizando essa estimativa, a série transformada $X_t = g^{-1}(\Theta_t - \hat{\mu})$ pode ser identificada da maneira usual, isto é, através dos procedimentos de identificação de séries temporais lineares.

3.4 Ajuste do Modelo

3.4.1 Modelo Gaussiano Arqueado

O processo linear arqueado X_t pode ser parametrizado através da média μ , da variância do processo c_0 e das p primeiras covariâncias, c_1, \dots, c_p . O vetor de estatísticas $\mathbf{s} = (\sum X_t, \sum X_t^2, \dots, \sum X_t X_{t-p})'$ é conjuntamente suficiente para estimar o vetor de parâmetros $(\mu, c_0, c_1, \dots, c_p)'$ [Fisher e Lee, 1994].

Considerando uma série temporal, $\theta_1, \dots, \theta_\tau$, o algoritmo EM deve ser utilizado em duas etapas:

▷ Passo E \rightarrow dada uma estimativa γ_N de $(\mu, c_0, c_1, \dots, c_p)'$ calcula-se

$$\mathbf{s}_N = E[\mathbf{s} | \gamma_N, \theta_1, \dots, \theta_\tau];$$

▷ Passo M → calcula-se a estimativa atualizada γ_{N+1} , bem como a solução de

$$\mathbf{s}_N = E[\mathbf{s} | \gamma_{N+1}].$$

A maior dificuldade na aplicação deste algoritmo encontra-se no cálculo da esperança condicional apresentada no passo E. Para se ter uma idéia dessa dificuldade, basta observar a quantidade de cálculos computacionais necessários para determinar $E[\sum X_t X_{t-1} | \theta_1, \dots, \theta_\tau]$. Por exemplo, a esperança condicional de $X_t X_{t+1}$ dados $\theta_1, \dots, \theta_\tau$ é dada por

$$\frac{\sum_{m_1=-\infty}^{\infty} \dots \sum_{m_\tau=-\infty}^{\infty} (\theta_t + 2\pi m_t)(\theta_{t-1} + 2\pi m_{t-1}) f(\theta_1 + 2\pi m_1, \dots, \theta_\tau + 2\pi m_\tau)}{\sum_{m_1=-\infty}^{\infty} \dots \sum_{m_\tau=-\infty}^{\infty} f(\theta_1 + 2\pi m_1, \dots, \theta_\tau + 2\pi m_\tau)},$$

onde f é a densidade do processo AR(p). Como, à medida que j cresce, a influência de θ_{t-j} sobre θ_t decresce rapidamente, então $E[X_t X_{t+1} | \theta_1, \dots, \theta_\tau]$, no caso do processo AR(1), pode ser aproximada por $E[X_t X_{t+1} | \theta_1, \theta_{t-1}]$, cuja expressão é, computacionalmente, mais tratável

$$\frac{\sum_{m_1=-\infty}^{\infty} \sum_{m_{t-1}=-\infty}^{\infty} (\theta_t + 2\pi m_t)(\theta_{t-1} + 2\pi m_{t-1}) f_{t,t-1}(\theta_t + 2\pi m_t, \dots, \theta_{t-1} + 2\pi m_{t-1})}{\sum_{m_1=-\infty}^{\infty} \sum_{m_{t-1}=-\infty}^{\infty} f_{t,t-1}(\theta_t + 2\pi m_t, \dots, \theta_{t-1} + 2\pi m_{t-1})},$$

onde $f_{t,t-1}$ é a densidade conjunta de $(X_t X_{t+1})$. Essa aproximação é aplicável apenas para modelos AR de ordem baixa [Fisher e Lee, 1994].

3.4.2 Modelo Gaussiano Transformado

De modo análogo ao processo arqueado, podemos parametrizar os processos lineares X_t e Y_t em termos da média μ , da variância do processo c_0 e das p primeiras covariâncias, c_1, \dots, c_p . Novamente, na aplicação do algoritmo EM, a maior dificuldade encontra-se no cálculo da esperança condicional.

Considerando a representação polar $(R_t \cos \Theta_t, R_t \sin \Theta_t)$ do vetor (X_t, Y_t) , temos

$$\begin{aligned} E[X_t X_{t-j} | \theta_1, \dots, \theta_\tau] &= E[R_t \cos \Theta_t R_{t-j} \cos \Theta_{t-j} | \theta_1, \dots, \theta_\tau] \\ &= \cos \theta_t \cos \theta_{t-j} E[R_t R_{t-j} | \theta_1, \dots, \theta_\tau] \end{aligned}$$

Se X e Y têm distribuição normal τ -variada com média zero e matriz de covariâncias Σ e se (R_i, θ_i) á representação polar de (X_i, Y_i) , então a esperança condicional de $R_t R_{t-j}$ dados $\theta_1, \dots, \theta_\tau$ é

$$E[R_t R_{t-j} | \theta_1, \dots, \theta_\tau] = \frac{\int_0^\infty \cdots \int_0^\infty r_t r_{t-j} \prod_{s=1}^\tau r_s \exp\left(-\frac{1}{2} r' A r\right) dr_1 \dots dr_\tau}{\int_0^\infty \cdots \int_0^\infty \prod_{s=1}^\tau r_s \exp\left(-\frac{1}{2} r' A r\right) dr_1 \dots dr_\tau},$$

onde $A = (a_{ij})$, $a_{ij} = \sigma^{ij} \cos(\theta_i - \theta_j)$ e $\Sigma^{-1} = (\sigma^{ij})$ [Fisher e Lee, 1994]. Para processos AR de ordem baixa pode-se aproximar $E[R_t R_{t-j} | \theta_1, \dots, \theta_\tau]$ por $E[R_t R_{t-j} | \theta_t, \theta_{t-j}]$. Em particular, para $p = 1$, $E[R_t R_{t-1} | \theta_1, \dots, \theta_\tau]$ pode ser aproximada por

$$\frac{\int_0^\infty \cdots \int_0^\infty r_1^2 r_2^2 \exp\left[-\frac{1}{2\sigma^2}\{r_1^2 + r_2^2 - 2\phi_1 r_1 r_2 \cos(\theta_t - \theta_{t-1})\}\right] dr_1 dr_2}{\int_0^\infty \cdots \int_0^\infty r_1 r_2 \exp\left[-\frac{1}{2\sigma^2}\{r_1^2 + r_2^2 - 2\phi_1 r_1 r_2 \cos(\theta_t - \theta_{t-1})\}\right] dr_1 dr_2}.$$

3.4.3 Modelos com Ligação Direta e Inversa

Após estimar a média através da média direcional amostral da série, pode-se utilizar um *software* para realizar o ajuste do processo AR(p) com ligação direta.

Para ajustar processos AR(p) com ligação inversa, assume-se uma distribuição marginal para $\Theta_1, \dots, \Theta_p$. Vamos supor que esta distribuição seja o produto de distribuições $\mathcal{VM}(\mu, \kappa)$.

A densidade conjunta de $\theta_1, \dots, \theta_\tau$ é dada por

$$\prod_{t=p+1}^\tau f(\theta_t - \mu - g[\omega_1 g^{-1}\{(\theta_{t-1} - \mu)/2\} + \dots + \omega_p g^{-1}\{(\theta_{t-p} - \mu)/2\}]) \prod_{t=1}^p f(\theta_t - \mu),$$

onde f é a densidade $\mathcal{VM}(0, \kappa)$.

Conclusão

A análise de séries temporais envolvendo dados angulares é feita de maneira ligeiramente semelhante à análise de séries temporais envolvendo dados lineares. Vários modelos utilizados para descrever, analisar e interpretar séries temporais angulares são obtidos a partir de modelos para séries temporais envolvendo dados lineares, daí a necessidade de se ter domínio da teoria desenvolvida para a análise de séries temporais lineares. Os modelos estudados têm ampla aplicação na perfuração de poços petrolíferos direcionais e podem ser utilizados para modelar a posição da broca de perfuração ao longo da trajetória de um poço petrolífero, considerada como uma série temporal de dados angulares.

Apêndice A

Demonstração da Desigualdade (3.1)

Para demonstrar essa desigualdade utilizamos dois resultados, enunciados a seguir, cujas demonstrações podem ser encontradas em [James, 2002].

A.1 Resultados Utilizados

Desigualdade de Jensen - *Seja X uma variável aleatória com $E(X) < \infty$ e g uma função convexa. Então*

$$g(E(X)) \leq E(g(X)).$$

Desigualdade de Cauchy-Schwarz - *Sejam X e Y variáveis aleatórias com segundo momento finito. Então*

$$E(|XY|) \leq \sqrt{E(X^2)E(Y^2)}.$$

A.2 Demonstração da Desigualdade (3.1)

Podemos expressar o coeficiente de correlação da seguinte forma

$$\rho_\tau = \frac{E(XY)}{\sqrt{E(X^2)E(Y^2)}},$$

com $X = \text{sen}(\Theta_1 - \Theta_2)$ e $Y = \text{sen}(\Phi_1 - \Phi_2)$.

Pela desigualdade de Jensen temos que

$$|E(XY)| \leq E(|XY|), \tag{A.1}$$

já que $E(XY) < \infty$ e a função modular é convexa.

Por outro lado, usando a desigualdade de Cauchy-Schwarz, obtemos

$$E(|XY|) \leq \sqrt{E(X^2)E(Y^2)}. \quad (\text{A.2})$$

De (A.1) e (A.2), segue que

$$|E(XY)| \leq \sqrt{E(X^2)E(Y^2)},$$

daí,

$$-\sqrt{E(X^2)E(Y^2)} \leq E(XY) \leq \sqrt{E(X^2)E(Y^2)},$$

logo,

$$-1 \leq \frac{E(XY)}{\sqrt{E(X^2)E(Y^2)}} \leq 1,$$

ou seja,

$$-1 \leq \rho_\tau \leq 1.$$

Apêndice B

Demonstração do Teorema (3.1)

Antes de demonstrar esse teorema, vamos apresentar algumas definições e propriedades, que serão úteis durante a demonstração.

B.1 Definições

Definição 1 A função densidade de probabilidade de um vetor aleatório (X, Y) com distribuição normal padrão bivariada é dada por

$$f(x, y) = \frac{1}{2\pi\sigma^2(1-\rho^2)^{1/2}} \exp\left\{-\frac{1}{2(1-\rho^2)}(x^2 + y^2 - 2\rho xy)\right\},$$

onde ρ é a correlação entre X e Y .

Definição 2 A função gama, é denotada por $\Gamma(\cdot)$ e definida por

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} \exp(-t) dt, \alpha > 0.$$

Definição 3 A função hipergeométrica, com argumentos α, β, γ e x , é definida por

$${}_2F_1(\alpha, \beta, \gamma; x) = \sum_{m=0}^{\infty} \frac{\alpha^{[m]} \beta^{[m]} x^m}{\gamma^{[m]} m!} = 1 + \frac{\alpha\beta}{\gamma} \frac{x}{1!} + \frac{\alpha(\alpha+1)\beta(\beta+1)}{\gamma(\gamma+1)} \frac{x^2}{2!} + \dots,$$

onde

$$h^{[m]} = \begin{cases} 1 & \text{se } m = 0, \\ \prod_{i=0}^{m-1} (h+i) & \text{se } m > 0. \end{cases}$$

B.2 Propriedades

1) A função gama satisfaz

$$(i) \Gamma(\alpha + 1) = \alpha\Gamma(\alpha), \alpha > 0;$$

$$(ii) \Gamma(n) = (n - 1)!, \text{ para qualquer inteiro positivo } n;$$

$$(iii) \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

2) Para quaisquer inteiros positivos p e q valem as seguintes fórmulas de integração

$$(i) \int \text{sen}^p x \cos^q x \, dx = -\frac{\text{sen}^{p-1} x \cos^{q+1} x}{p+q} + \frac{p-1}{p+q} \int \text{sen}^{p-2} x \cos^q x \, dx + c;$$

$$(ii) \int \cos^q x \, dx = \frac{\cos^{q-1} x \text{sen} x}{q} + \frac{q-1}{q} \int \cos^{q-2} x \, dx + c.$$

B.3 Demonstração do Teorema (3.1)

Seja (R_i, Θ_i) a representação polar de (X_i, Y_i) , ou seja,

$$(X_i, Y_i) = (R_i \cos \Theta_i, R_i \text{sen} \Theta_i), \quad i = 1, 2.$$

Como a distribuição de Θ_1 e Θ_2 é uniforme, podemos escrever ρ_τ da seguinte maneira

$$\rho_\tau = 4(E[\text{sen}\Theta_1 \text{sen}\Theta_2]E[\cos\Theta_1 \cos\Theta_2] - E[\text{sen}\Theta_1 \cos\Theta_2]E[\text{sen}\Theta_1 \cos\Theta_2]).$$

Para calcular cada valor esperado, precisamos encontrar a função densidade conjunta de $(R_1, \Theta_1, R_2, \Theta_2)$. Para tanto, utilizamos o método do jacobiano.

Sendo $X = R \cos \Theta$ e $Y = R \text{sen} \Theta$, $f_{(X,Y)}$ e $f_{(R,\Theta)}$ as densidades de (X, Y) e (R, Θ) , respectivamente, então

$$f_{(R,\Theta)}(r, \theta) = f_{(X,Y)}(r \cos \theta, r \text{sen} \theta) |J|,$$

onde J é o jacobiano da transformação, calculado da seguinte forma

$$J = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \text{sen} \theta \\ \text{sen} \theta & r \cos \theta \end{vmatrix} = r \cos^2 \theta + r \text{sen}^2 \theta = r,$$

assim,

$$\begin{aligned} f_{(R,\Theta)}(r,\theta) &= \frac{r}{2\pi(1-\rho^2)^{1/2}} \exp\left\{\frac{-1}{2(1-\rho^2)}[r^2 \cos^2 \theta + r^2 \sin^2 \theta - 2\rho r^2 \cos \theta \sin \theta]\right\} \\ &= \frac{r}{2\pi(1-\rho^2)^{1/2}} \exp\left\{\frac{-1}{2(1-\rho^2)}[r^2 - 2\rho r^2 \cos \theta \sin \theta]\right\}. \end{aligned}$$

Por hipótese, (X_1, Y_1) e (X_2, Y_2) são independentes, donde segue que (R_1, Θ_1) e (R_2, Θ_2) também são independentes. Daí, a densidade conjunta de $(R_1, \Theta_1, R_2, \Theta_2)$ pode ser encontrada da seguinte maneira

$$\begin{aligned} f_{(R_1, \Theta_1, R_2, \Theta_2)}(r_1, \theta_1, r_2, \theta_2) &= f_{(R_1, \Theta_1)}(r_1, \theta_1) f_{(R_2, \Theta_2)}(r_2, \theta_2) \\ &= \frac{r_1}{2\pi(1-\rho^2)^{1/2}} \exp\left\{\frac{-1}{2(1-\rho^2)}[r_1^2 - 2\rho r_1^2 \cos \theta_1 \sin \theta_1]\right\} \\ &\quad \times \frac{r_2}{2\pi(1-\rho^2)^{1/2}} \exp\left\{\frac{-1}{2(1-\rho^2)}[r_2^2 - 2\rho r_2^2 \cos \theta_2 \sin \theta_2]\right\}. \end{aligned}$$

Após algumas manipulações algébricas, obtemos

$$\begin{aligned} f_{(R_1, \Theta_1, R_2, \Theta_2)}(r_1, \theta_1, r_2, \theta_2) &= \frac{r_1 r_2}{(2\pi)^2(1-\rho^2)} \exp\left\{\frac{-1}{2(1-\rho^2)}[r_1^2 + r_2^2 \cos(\theta_1 - \theta_2)]\right\} \\ &= \frac{r_1 r_2}{(2\pi)^2(1-\rho^2)} \exp\left\{\frac{-r_1^2}{2(1-\rho^2)}\right\} \exp\left\{\frac{-r_2^2}{2(1-\rho^2)}\right\} \\ &\quad \times \exp\left\{\frac{\rho r_1 r_2 \cos(\theta_1 - \theta_2)}{(1-\rho^2)}\right\} \\ &= \frac{r_1 r_2}{(2\pi)^2(1-\rho^2)} \exp\left\{\frac{-r_1^2}{2(1-\rho^2)}\right\} \exp\left\{\frac{-r_2^2}{2(1-\rho^2)}\right\} \\ &\quad \times \sum_{n=0}^{\infty} \frac{1}{n!} r_1^n r_2^n \left[\frac{\rho}{(1-\rho^2)} \cos(\theta_1 - \theta_2)\right]^n \\ &= \frac{1}{(2\pi)^2(1-\rho^2)} \sum_{n=0}^{\infty} \frac{1}{n!} r_1^{n+1} \exp\left\{\frac{-r_1^2}{2(1-\rho^2)}\right\} \\ &\quad \times \exp\left\{\frac{-r_2^2}{2(1-\rho^2)}\right\} \left[\frac{\rho}{(1-\rho^2)}\right]^n \cos^n(\theta_1 - \theta_2) \end{aligned}$$

• Cálculo de $E(\sin \Theta_1 \cos \Theta_2)$

Por definição de valor esperado, temos que

$$\begin{aligned} E(\sin \Theta_1 \cos \Theta_2) &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \sin \theta_1 \cos \theta_2 f_{(\Theta_1, \Theta_2)}(\theta_1, \theta_2) d\theta_1 d\theta_2 \\ &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \sin \theta_1 \cos \theta_2 \left[\int_0^{\infty} \int_0^{\infty} f_{(R_1, \Theta_1, R_2, \Theta_2)}(r_1, \theta_1, r_2, \theta_2) dr_1 dr_2 \right] d\theta_1 d\theta_2 \end{aligned}$$

$$\begin{aligned}
&= C \sum_{n=0}^{\infty} \frac{1}{n!} \int_0^{\infty} r_1^{n+1} \exp \left\{ \frac{-r_1^2}{2(1-\rho^2)} \right\} dr_1 \int_0^{\infty} r_2^{n+1} \exp \left\{ \frac{-r_2^2}{2(1-\rho^2)} \right\} dr_2 \\
&\quad \times \underbrace{\left[\frac{\rho}{(1-\rho^2)} \right]^n \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \text{sen} \theta_1 \cos \theta_2 \cos^n(\theta_1 - \theta_2) d\theta_1 d\theta_2}_{I_1},
\end{aligned}$$

onde $C = 1/[(2\pi)^2(1-\rho^2)]$.

- Cálculo da integral I_1

Fazendo a seguinte mudança de variáveis

$$\theta_1 = x + y$$

$$\theta_2 = y$$

obtemos

$$\begin{aligned}
I_1 &= \iint_S \text{sen}(x+y) \cos y \cos^n x \, dx \, dy = \iint_S (\text{sen} x \cos y + \text{sen} y \cos x) \cos y \cos^n x \, dx \, dy \\
&= \underbrace{\iint_S \cos^n x \text{sen} x \cos^2 y \, dx \, dy}_{I_{11}} + \underbrace{\iint_S \cos^{n+1} x \text{sen} y \cos y \, dx \, dy}_{I_{12}}
\end{aligned}$$

onde S é a nova região de integração devido à mudança de variáveis.

- Cálculo da integral I_{11}

$$\begin{aligned}
I_{11} &= \int_{-2\pi}^0 \int_{-x-\pi}^{\pi} \cos^n x \text{sen} x \cos^2 y \, dy \, dx + \int_0^{2\pi} \int_{-\pi}^{-x+\pi} \cos^n x \text{sen} x \cos^2 y \, dy \, dx \\
&= \int_{-2\pi}^0 \cos^n x \text{sen} x \left[\frac{y}{2} + \frac{\text{sen} 2y}{4} \right]_{-x-\pi}^{\pi} dx + \int_0^{2\pi} \cos^n x \text{sen} x \left[\frac{y}{2} + \frac{\text{sen} 2y}{4} \right]_{-\pi}^{-x+\pi} dx \\
&= \int_{-2\pi}^0 \cos^n x \text{sen} x \left[\pi + \frac{x}{2} + \frac{\text{sen} 2x}{4} \right] dx + \int_0^{2\pi} \cos^n x \text{sen} x \left[\pi - \frac{x}{2} - \frac{\text{sen} 2x}{4} \right] dx \\
&= \pi \int_{-2\pi}^0 \cos^n x \text{sen} x \, dx + \frac{1}{2} \int_{-2\pi}^0 x \cos^n x \text{sen} x \, dx + \frac{1}{4} \int_{-2\pi}^0 \cos^n x \text{sen} x \text{sen} 2x \, dx \\
&\quad + \pi \int_0^{2\pi} \cos^n x \text{sen} x \, dx - \frac{1}{2} \int_0^{2\pi} x \cos^n x \text{sen} x \, dx - \frac{1}{4} \int_0^{2\pi} \cos^n x \text{sen} x \text{sen} 2x \, dx
\end{aligned}$$

como as funções $(x \cos^n x \operatorname{sen} x)$ e $(\cos^n x \operatorname{sen} x \operatorname{sen} 2x)$ são pares, segue que

$$\begin{aligned} I_{11} &= \pi \int_{-2\pi}^0 \cos^n x \operatorname{sen} x \, dx + \pi \int_0^{2\pi} \cos^n x \operatorname{sen} x \, dx \\ &= \pi \left[\int_{-2\pi}^0 \cos^n x \operatorname{sen} x \, dx + \int_0^{2\pi} \cos^n x \operatorname{sen} x \, dx \right] \\ &= \pi \int_{-2\pi}^{2\pi} \cos^n x \operatorname{sen} x \, dx = \pi \cdot 0 = 0, \end{aligned}$$

já que a função $(\cos^n x \operatorname{sen} x)$ é ímpar.

- Cálculo da integral I_{12}

$$\begin{aligned} I_{12} &= \int_{-2\pi}^0 \int_{-x-\pi}^{\pi} \cos^{n+1} x \operatorname{sen} y \cos y \, dy \, dx + \int_0^{2\pi} \int_{-\pi}^{-x+\pi} \cos^{n+1} x \operatorname{sen} y \cos y \, dy \, dx \\ &= \int_{-2\pi}^0 \cos^{n+1} x \left[\frac{\operatorname{sen}^2 y}{2} \right]_{-x-\pi}^{\pi} dx + \int_0^{2\pi} \cos^{n+1} x \left[\frac{\operatorname{sen}^2 y}{2} \right]_{-\pi}^{-x+\pi} dx \\ &= -\frac{1}{2} \int_{-2\pi}^0 \cos^{n+1} x \operatorname{sen}^2 x \, dx + \frac{1}{2} \int_0^{2\pi} \cos^{n+1} x \operatorname{sen}^2 x \, dx = 0, \end{aligned}$$

pois a função $(\cos^{n+1} x \operatorname{sen}^2 x)$ é uma função par. Portanto,

$$I_1 = 0, \quad \forall n,$$

donde segue que

$$E(\operatorname{sen} \Theta_1 \cos \Theta_2) = 0.$$

• Cálculo de $E(\cos \Theta_1 \cos \Theta_2)$

Por definição de valor esperado, temos que

$$\begin{aligned} E(\cos \Theta_1 \cos \Theta_2) &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \cos \theta_1 \cos \theta_2 f_{(\Theta_1, \Theta_2)}(\theta_1, \theta_2) d\theta_1 d\theta_2 \\ &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \cos \theta_1 \cos \theta_2 \left[\int_0^{\infty} \int_0^{\infty} f_{(R_1, \Theta_1, R_2, \Theta_2)}(r_1, \theta_1, r_2, \theta_2) dr_1 dr_2 \right] d\theta_1 d\theta_2 \\ &= C \sum_{n=0}^{\infty} \frac{1}{n!} \underbrace{\int_0^{\infty} r_1^{n+1} \exp \left\{ \frac{-r_1^2}{2(1-\rho^2)} \right\} dr_1}_{I_2} \underbrace{\int_0^{\infty} r_2^{n+1} \exp \left\{ \frac{-r_2^2}{2(1-\rho^2)} \right\} dr_2}_{I_3} \\ &\quad \times \underbrace{\left[\frac{\rho}{(1-\rho^2)} \right]^n \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \cos \theta_1 \cos \theta_2 \cos^n(\theta_1 - \theta_2) d\theta_1 d\theta_2}_{I_4}. \end{aligned}$$

- Cálculo da integral I_k , $k = 2, 3$

Fazendo a substituição $t = r_k^2/[2(1 - \rho^2)]$, segue que

$$r_k = [2(1 - \rho^2)]^{1/2} t^{1/2} \quad \Rightarrow \quad dr_k = \frac{[2(1 - \rho^2)]^{1/2}}{2t^{1/2}} dt.$$

Daí,

$$\begin{aligned} I_k &= \int_0^\infty \{[2(1 - \rho^2)]^{1/2} t^{1/2}\}^{n+1} \exp(-t) \frac{[2(1 - \rho^2)]^{1/2}}{2t^{1/2}} dt \\ &= \frac{[2(1 - \rho^2)]^{\frac{n+2}{2}}}{2} \int_0^\infty t^{(\frac{n}{2}+1)-1} \exp(-t) dt \\ &= \frac{[2(1 - \rho^2)]^{\frac{n+2}{2}}}{2} \Gamma\left(\frac{n+2}{2}\right), \quad k = 2, 3. \end{aligned}$$

Ou seja,

$$I_2 = I_3 = \frac{[2(1 - \rho^2)]^{\frac{n+2}{2}}}{2} \Gamma\left(\frac{n+2}{2}\right)$$

- Cálculo da integral I_4

Fazendo a seguinte mudança de variáveis

$$\theta_1 = x + y$$

$$\theta_2 = y$$

obtemos

$$\begin{aligned} I_4 &= \iint_S \cos(x + y) \cos y \cos^n x \, dx \, dy = \iint_S (\cos x \cos y + \operatorname{sen} x \operatorname{sen} y) \cos y \cos^n x \, dx \, dy \\ &= \underbrace{\iint_S \cos^{n+1} x \cos^2 y \, dx \, dy}_{I_{41}} - \underbrace{\iint_S \cos^{n+1} x \cos^2 y \, dx \, dy}_{I_{42}}. \end{aligned}$$

onde S é a nova região de integração devido à mudança de variáveis.

- Cálculo da integral I_{42}

$$\begin{aligned} I_{42} &= \int_{-2\pi}^0 \int_{-x-\pi}^{\pi} \cos^n x \operatorname{sen} x \operatorname{sen} y \cos y \, dy \, dx + \int_0^{2\pi} \int_{-\pi}^{-x+\pi} \cos^n x \operatorname{sen} x \operatorname{sen} y \cos y \, dy \, dx \\ &= \int_{-2\pi}^0 \cos^n x \operatorname{sen} x \left[\frac{\operatorname{sen}^2 y}{2} \right]_{-x-\pi}^{\pi} dx + \int_0^{2\pi} \cos^n x \operatorname{sen} x \left[\frac{\operatorname{sen}^2 y}{2} \right]_{-\pi}^{-x+\pi} dx \end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{2} \int_{-2\pi}^0 \cos^n x \operatorname{sen}^3 x dx + \frac{1}{2} \int_0^{2\pi} \cos^n x \operatorname{sen}^3 x dx \\
&= \frac{1}{2} \int_0^{2\pi} \cos^n x \operatorname{sen}^3 x dx + \frac{1}{2} \int_0^{2\pi} \cos^n x \operatorname{sen}^3 x dx = \int_0^{2\pi} \cos^n x \operatorname{sen}^3 x dx \\
&= \left[-\frac{\operatorname{sen}^2 x \cos^{n+1} x}{n+3} \right]_0^{2\pi} + \frac{2}{n+3} \int_0^{2\pi} \cos^n x \operatorname{sen} x dx = \frac{2}{n+3} \left[-\frac{\cos^{n+1} x}{n+1} \right]_0^{2\pi} = 0.
\end{aligned}$$

- Cálculo da integral I_{41}

$$\begin{aligned}
I_{41} &= \int_{-2\pi}^0 \int_{-x-\pi}^{\pi} \cos^{n+1} x \cos^2 y dy dx + \int_0^{2\pi} \int_{-\pi}^{-x+\pi} \cos^{n+1} x \cos^2 y dy dx \\
&= \int_{-2\pi}^0 \cos^{n+1} x \left[\frac{y}{2} + \frac{\operatorname{sen} 2y}{4} \right]_{-x-\pi}^{\pi} dx + \int_0^{2\pi} \cos^{n+1} x \left[\frac{y}{2} + \frac{\operatorname{sen} 2y}{4} \right]_{-\pi}^{-x+\pi} dx \\
&= \int_{-2\pi}^0 \cos^{n+1} x \left[\pi + \frac{x}{2} + \frac{\operatorname{sen} 2x}{4} \right] dx + \int_0^{2\pi} \cos^{n+1} x \left[\pi - \frac{x}{2} - \frac{\operatorname{sen} 2x}{4} \right] dx \\
&= \pi \int_{-2\pi}^0 \cos^{n+1} x dx + \frac{1}{2} \int_{-2\pi}^0 x \cos^{n+1} x dx + \frac{1}{2} \int_{-2\pi}^0 \cos^{n+2} x \operatorname{sen} x dx \\
&\quad + \pi \int_0^{2\pi} \cos^{n+1} x dx - \frac{1}{2} \int_0^{2\pi} x \cos^{n+1} x dx - \frac{1}{2} \int_0^{2\pi} \cos^{n+2} x \operatorname{sen} x dx \\
&= 2\pi \int_{-2\pi}^0 \cos^{n+1} x dx + \int_{-2\pi}^0 x \cos^{n+1} x dx + \int_{-2\pi}^0 \cos^{n+2} x \operatorname{sen} x dx \\
&= \int_{-2\pi}^0 (x+2\pi) \cos^{n+1} x dx + \left[-\frac{\cos^{n+3} x}{n+3} \right]_{-2\pi}^0 = \int_{-2\pi}^0 (x+2\pi) \cos^{n+1} x dx.
\end{aligned}$$

Fazendo

$$u = (x + 2\pi) \Rightarrow du = dx,$$

$$dv = \cos^{n+1} x dx \Rightarrow v = \int \cos^{n+1} x dx,$$

Para calcular v utilizamos, recursivamente, a propriedade 2(ii) enunciada no início deste apêndice. Observe que a aplicação recursiva dessa propriedade, reduz a potência $(n+1)$ do cosseno e, dependendo da paridade de n , essa redução resulta em $\int \cos x dx$ ou $\int dx$.

Se n for par então

$$\begin{aligned}
v &= \int \cos^{n+1} x dx = \frac{\cos^n x \operatorname{sen} x}{(n+1)} + \frac{n}{(n+1)} \int \cos^{n-1} x dx \\
&= \frac{\cos^n x \operatorname{sen} x}{(n+1)} + \frac{n}{(n+1)} \frac{\cos^{n-2} x \operatorname{sen} x}{(n-1)} + \frac{n(n-2)}{(n+1)(n-1)} \int \cos^{n-3} x dx
\end{aligned}$$

$$\begin{aligned}
&= \frac{\cos^n x \operatorname{sen} x}{(n+1)} + \frac{n}{(n+1)} \frac{\cos^{n-2} x \operatorname{sen} x}{(n-1)} + \frac{n(n-2)}{(n+1)(n-1)} \frac{\cos^{n-4} x \operatorname{sen} x}{(n-3)} \\
&\quad + \frac{n(n-2)(n-4)}{(n+1)(n-1)(n-3)} \int \cos^{n-5} x \, dx \\
&= \frac{\cos^n x \operatorname{sen} x}{(n+1)} + \frac{n}{(n+1)} \frac{\cos^{n-2} x \operatorname{sen} x}{(n-1)} + \frac{n(n-2)}{(n+1)(n-1)} \frac{\cos^{n-4} x \operatorname{sen} x}{(n-3)} \\
&\quad + \dots + \frac{n(n-2)(n-4) \dots 8.6.4}{(n+1)(n-1)(n-3) \dots 9.7.5} \frac{\cos^2 x \operatorname{sen} x}{3} \\
&\quad + \frac{n(n-2)(n-4) \dots 8.6.4.2}{(n+1)(n-1)(n-3) \dots 9.7.5.3} \int \cos x \, dx.
\end{aligned}$$

Integrando por partes, temos que

$$\begin{aligned}
I_{41} &= [uv]_{-2\pi}^0 - \int_{-2\pi}^0 v \, du = 0 - \int_{-2\pi}^0 v \, du = \frac{1}{n} \left[\frac{\cos^{n+1} x}{(n+1)} \right] \\
&= + \frac{n}{(n+1)(n-1)} \left[\frac{\cos^{n-1} x}{(n-1)} \right]_{-2\pi}^0 + \dots + \frac{n(n-2)(n-4) \dots 8.6.4}{(n+1)(n-1)(n-3) \dots 9.7.5.3} \left[\frac{\cos^3 x}{3} \right]_{-2\pi}^0 \\
&\quad + \frac{n(n-2)(n-4) \dots 8.6.4.2}{(n+1)(n-1)(n-3) \dots 9.7.5.3} [\cos x]_{-2\pi}^0 = 0.
\end{aligned}$$

Portanto, se n for um número par então $I_{41} = 0$.

Se n for ímpar então o cálculo de v através da aplicação recursiva da propriedade 2(ii) resulta em

$$\begin{aligned}
v &= \int \cos^{n+1} x \, dx = \frac{\cos^n x \operatorname{sen} x}{(n+1)} + \frac{n}{(n+1)} \frac{\cos^{n-2} x \operatorname{sen} x}{(n-1)} \\
&= + \frac{n(n-2)}{(n+1)(n-1)} \frac{\cos^{n-4} x \operatorname{sen} x}{(n-3)} + \dots + \frac{n(n-2)(n-4) \dots 7.5.3}{(n+1)(n-1)(n-3) \dots 8.6.4} \frac{\cos x \operatorname{sen} x}{2} \\
&\quad + \frac{n(n-2)(n-4) \dots 7.5.3.1}{(n+1)(n-1)(n-3) \dots 8.6.4.2} \int dx.
\end{aligned}$$

Daí, para n ímpar, temos que

$$\begin{aligned}
I_{41} &= [uv]_{-2\pi}^0 - \int_{-2\pi}^0 v \, du = 0 - \int_{-2\pi}^0 v \, du \frac{1}{(n+1)} \left[\frac{\cos^{n+1} x}{(n+1)} \right]_{-2\pi}^0 \\
&= + \frac{n}{(n+1)(n-1)} \left[\frac{\cos^{n-1} x}{(n-1)} \right]_{-2\pi}^0 + \dots + \frac{n(n-2)(n-4) \dots 7.5.3}{(n+1)(n-1)(n-3) \dots 8.6.4.2} \left[\frac{\cos^3 x}{3} \right]_{-2\pi}^0 \\
&\quad + \frac{n(n-2)(n-4) \dots 7.5.3.1}{(n+1)(n-1)(n-3) \dots 8.6.4.2} \left[-\frac{x}{2} \right]_{-2\pi}^0 = \frac{n(n-2)(n-4) \dots 7.5.3.1}{(n+1)(n-1)(n-3) \dots 8.6.4.2} (2\pi^2) \\
&= \frac{n(n-1)(n-2)(n-3)(n-4) \dots 8.7.6.5.4.3.2.1}{(n+1)[(n-1)(n-3) \dots 8.6.4.2]^2} (2\pi^2) \\
&= \frac{n!}{(n+1)[(n-1)(n-3) \dots 8.6.4.2]^2} (2\pi^2).
\end{aligned}$$

Sendo n ímpar, podemos escrevê-lo na forma $n = 2m + 1$, assim,

$$\begin{aligned} I_{41} &= \frac{(2m+1)!}{(2m+2)[(2m)(2m-2)\dots 8.6.4.2]^2} (2\pi^2) \\ &= \frac{(2m+1)!}{2(m+1)[2^m m(m-1)\dots 4.3.2.1]^2} (2\pi^2) = \frac{(2m+1)!}{(m+1)2^{2m}[m!]^2} \pi^2 \\ &= \frac{(2m+1)!}{2^{2m}(m+1)m!m!} \pi^2 = \frac{(2m+1)!}{2^{2m}(m+1)m!} \pi^2 = \frac{\pi^2}{2^{2m}} \binom{2m+1}{m}. \end{aligned}$$

Logo,

$$I_{41} = \begin{cases} 0 & \text{se } n \text{ for par;} \\ \frac{\pi^2}{2^{2m}} \binom{2m+1}{m} & \text{se } n \text{ for ímpar, } n = 2m + 1. \end{cases}$$

Portanto,

$$I_4 = I_{41} - I_{42} = I_{41} - 0 = I_{41} = \begin{cases} 0 & \text{se } n \text{ for par;} \\ \frac{\pi^2}{2^{2m}} \binom{2m+1}{m} & \text{se } n \text{ for ímpar, } n = 2m + 1. \end{cases}$$

Substituindo I_2 , I_3 e I_4 na expressão de $E(\cos \Theta_1 \cos \Theta_2)$, obtemos

$$\begin{aligned} E(\cos \Theta_1 \cos \Theta_2) &= \frac{1}{(2\pi)^2(1-\rho^2)} \sum_{m=0}^{\infty} \frac{1}{(2m+1)!} \frac{1}{4} [2(1-\rho^2)]^{(2m+1+2)} \\ &\quad \times \Gamma^2\left(\frac{2m+1+2}{2}\right) \frac{\rho^{2m+1}}{(1-\rho^2)^{2m+1}} \frac{\pi^2}{2^{2m}} \binom{2m+1}{m} \\ &= \frac{1}{2(1-\rho^2)} \sum_{m=0}^{\infty} 2^{2m} (1-\rho^2)^{2m+3} \Gamma^2\left(m + \frac{3}{2}\right) \\ &\quad \times \frac{\rho^{2m} \rho}{(1-\rho^2)^{2m+1}} \frac{1}{2^{2m}} \frac{1}{m!(m+1)!} = \frac{1}{2} \rho (1-\rho^2) \sum_{m=0}^{\infty} \frac{\Gamma^2\left(m + \frac{3}{2}\right)}{\Gamma(m+2)} \frac{(\rho^2)^m}{m!}. \end{aligned}$$

Podemos reescrever $\Gamma(m+2)$ da seguinte maneira

$$\begin{aligned} \Gamma(m+2) &= (m+1)! = (m+1)m(m-1)\dots 4.3.2 = 2.3.4\dots(m-1)m(m+1) \\ &= (2+0)(2+1)(2+2)\dots(2+m-3)(2+m-2)(2+m-1) \\ &= \prod_{i=0}^{m-1} (2+i) = 2^{[m]}. \end{aligned}$$

Além disso, pode-se mostrar, por indução, que

$$\Gamma\left(m + \frac{3}{2}\right) = \left(\frac{3}{2}\right)^{[m]} \frac{\sqrt{\pi}}{2}.$$

Assim,

$$E(\cos \Theta_1 \cos \Theta_2) = \frac{\pi}{8} \rho(1 - \rho^2) \sum_{m=0}^{\infty} \frac{\left(\frac{3}{2}\right)^{[m]} \left(\frac{3}{2}\right)^{[m]} (\rho^2)^m}{2^{[m]} m!} = \frac{\pi}{8} \rho(1 - \rho^2) {}_2F_1\left(\frac{3}{2}, \frac{3}{2}, 2; \rho^2\right),$$

onde ${}_2F_1\left(\frac{3}{2}, \frac{3}{2}, 2; \rho^2\right)$ é a função hipergeométrica definida no início deste apêndice.

• Cálculo de $E(\sin \Theta_1 \sin \Theta_2)$

O cálculo de $E(\sin \Theta_1 \sin \Theta_2)$ é feito de maneira análoga ao anterior, de onde vem

$$E(\sin \Theta_1 \sin \Theta_2) = \frac{\pi}{8} \rho(1 - \rho^2) {}_2F_1\left(\frac{3}{2}, \frac{3}{2}, 2; \rho^2\right).$$

Substituindo, cada valor esperado calculado, na expressão do coeficiente de correlação, obtemos

$$\rho_\tau = \frac{\pi^2}{16} \rho^2 (1 - \rho^2)^2 \left\{ {}_2F_1\left(\frac{3}{2}, \frac{3}{2}, 2; \rho^2\right) \right\}^2,$$

como queríamos mostrar.

Referências Bibliográficas

- [Box e Jenkins, 1976] Box, G. E. P. e Jenkins, G. M (1976). *Time Series Analysis - forecasting and control*. San Francisco: Holden-Day.
- [Charles et al., 2001] Charles, T., Guéméné, J. M., Vicent, G., Dubrule, O. e TotalFinaElf (2001). *Experience with the Quantification of Subsurface Uncertainties*. Society of Petroleum Engineers, nº 68703.
- [Fisher, 1993] Fisher, N. I (1993). *Statistical Analysis of Circular Data*. Cambridge: University Press.
- [Fisher e Lee, 1983] Fisher, N. I. e Lee, A. J (1983). A correlation coefficient for circular data. *Biometrics* 70, 327-332.
- [Fisher e Lee, 1992] Fisher, N. I. e Lee, A. J (1992). Regression models for an angular response. *Biometrics* 48, 665-677.
- [Fisher e Lee, 1994] Fisher, N. I. e Lee, A. J (1994). Time series analysis of circular data. *Journal of the Royal Statistical Society*, B 56, 327-339.
- [James, 2002] James, B. B (2002). *Probabilidade: Um curso em Nível Intermediário*. Rio de Janeiro: IMPA.
- [Johnson e Kotz, 1969] Johnson, L. N. e Kotz, S (1969). *Discrete Distributions*. Boston.
- [Lima] Lima, H. R. P. L. *Fundamentos de Perfuração*. Notas de um curso do Centro de Desenvolvimento de Recursos Humanos da Petrobras.
- [Mardia e Jupp, 2000] Mardia, K. V. e Jupp, P. E (2000). *Directional Statistics*. New York: John Wiley & Sons.

- [Morettin e Toloi, 2004] Morettin, P. A. e Toloi, C. M. C (2004). *Análise de Séries Temporais*. São Paulo: Edgard Blücher.
- [Silva, 2004] Silva, A. M (2004). *Um Modelo Estocástico para Previsão de Desvios da Coluna de Perfuração em Poços Petrolíferos*. Monografia de Graduação - Programa de Recursos Humanos da ANP/MCT - PRH(25). Universidade Federal de Campina Grande - UFCG. Campina Grande - PB.
- [Thomas, 2001] Thomas, J. E (2001). *Fundamentos de Engenharia de Petróleo*. Editora Interciência.
- [Thomas, 2003] Thomas, G. B (2003). *Cálculo*. São Paulo: Addison Wesley.
- [Wilks, 1995] Wilks, D. S (1995). *Statistical Methods in the Atmospheric Sciences*. Academic Press.