3
Metodologia

Uma série temporal pode ser caracterizada como uma coleção de observações de uma variável aleatória, dispostas de maneira seqüencial e ordenada em uma determinada unidade de tempo (ano, mês, semana etc). Tais observações, geralmente, estão em intervalos temporais de mesmo tamanho.

Podemos imaginar Z_t como sendo uma observação da variável aleatória Z no instante t. Temos então que $Z_1, Z_2, \cdots, Z_n$ é uma série temporal com N observações (Souza e Camargo, 2004). Logo, o objetivo principal da análise de uma série temporal é fazer inferência a respeito da mesma baseado em suas características principais.

A análise de uma série temporal, em geral, segue dois tipos de abordagens:

- Domínio do tempo
- Domínio da frequência

Analisar a série via domínio do tempo significa dizer que a própria relação das observações dispostas ao longo do tempo e a evolução das mesmas são investigadas levando em consideração a unidade temporal adotada. Segundo Morettin e Toloi (2006), modelos paramétricos são utilizados para fazer análises no domínio do tempo. Entre esses, os mais comuns são os modelos de regressão, modelos auto-regressivos e de médias móveis (ARMA), modelos auto-regressivos integrados de médias móveis (ARIMA), modelos estruturais, modelos não lineares e modelos de memória longa (ARFIMA).

Podemos medir a covariância da série com ela mesma em diferentes períodos de tempo. A isso é atribuído o nome de autocovariância. Como exemplo, podemos medir a relação de covariância entre a velocidade de vento em uma data no instante t com uma data um passo atrás, $t-1$, ou ainda dois passos atrás $t-2$. Podemos então generalizar a função de autocovariância de Z_t com Z_{t+k} como:

$$\begin{aligned} \gamma_k &= COV[Z_t, Z_{t+k}] \\ &= E[(Z_t - \mu)(Z_{t+k} - \mu)] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(Z_t - \mu)(Z_{t+k} - \mu)P(Z_t, Z_{t+k})dZ_t dZ_{t+k} \end{aligned} \tag{3-1}$$

Onde:
μ é a média do processo
$P(Z_t, Z_{t+k})$ é a função densidade de probabilidade conjunta das variáveis Z_t e Z_{t+k}

Logo, podemos estimar a autocovariância de uma série usando:

$$\hat{\gamma}_k = \frac{1}{N}\sum_{t=1}^{N-k}(Z_t - \bar{Z})(Z_{t+k} - \bar{Z}) \tag{3-2}$$

Onde:
$\bar{Z}$ é a média da série temporal Z definida como:

$$\bar{Z} = \frac{1}{N}\sum_{t=1}^{N} Z_t \tag{3-3}$$

Podemos ainda padronizar a função autocovariância, resultando então na função de autocorrelação. A mesma mede a dependência entre os termos levando em conta todos os termos intermediários. Logo, temos como função de autocorrelação:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{COV[Z_t, Z_{t+k}]}{\sqrt{VAR[Z_t]VAR[Z_{t+k}]}} \tag{3-4}$$

Onde γ_0 é a variância de Z_t e Z_{t+k}. Além disso, $\rho_0 = 1$ e $\rho_k = \rho_{-k}$

A estimação da função de autocorrelação é dada por:

$$\hat{\rho}_k = \frac{\gamma_k}{\gamma_0} = \frac{\sum_{t=1}^{N-k}(Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=1}^{N-k}(Z_t - \bar{Z})^2} \tag{3-5}$$

Podemos ainda utilizar uma variação da função de autocorrelação chamada função de autocorrelação parcial que mede a dependência entre os termos de interesse, por exemplo, Z_t com Z_{t+k} não considerando a dependência entre os termos intermediários, $Z_{t+1}, Z_t, \ldots, Z_{t+k-1}$. A mesma pode ser representada por:

$$COR[Z_t, Z_{t+k}|Z_{t+1}, Z_t, \ldots, Z_{t+k-1}] \tag{3-6}$$

O domínio da freqüência leva em consideração as freqüências de ocorrência de determinados eventos dentro de um determinado período de tempo. Segundo Morettin e Toloi (2006), a vantagem de descrever a série temporal no domínio da freqüência está no fato de eliminar problemas de correlação serial já que os componentes são ortogonais nesse tipo de abordagem. Como será visto mais a frente, a ortogonalidade assumirá papel fundamental na elaboração do modelo proposto. A análise espectral é a ferramenta utilizada, em geral, para fazer as análises. Neste tipo de abordagem são utilizadas periodogramas de janelas espectrais.

Ainda segundo Morettin e Toloi (2006), se Z é um processo estacionário discreto, é possível definir o espectro de Z como:

$$f(\lambda) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \gamma(\tau) e^{-i\lambda\tau}, \quad -\infty \leq \lambda \leq \infty \tag{3-7}$$

e

$$\gamma(\tau) = \int_{-\pi}^{\pi} e^{i\lambda\tau} f(\lambda) d\lambda, \quad \tau = 0, \pm 1, \cdots \tag{3-8}$$

Onde:
$e^{i\lambda} = cos(\lambda) + isen(\lambda)$
$i = \sqrt{-1}$

3.1 Processos Estocásticos

Um processo estocástico pode ser definido como uma família de variáveis aleatórias supostamente definidas em um mesmo espaço de probabilidades (Morettin e Toloi, 2006). Portanto, pode-se pensar uma série temporal como uma realização de um processo estocástico. Logo, temos que se existe uma série temporal Z_t, a mesma é definida como função Z da variável independente t, que é gerada por um processo estocástico desconhecido (Souza e Camargo,2004).

O comportamento futuro de um processo estocástico deve ser descrito via funções probabilísticas já que o mecanismo gerador de dados do processo não é conhecido além de uma T-ésima ordem e, portanto, seu comportamento além dessa ordem não pode ser descrito via funções determinísticas. Para contornar o não conhecimento do processo até a T-ésima ordem, assumimos duas condições: ergodicidade e estacionariedade.

De maneira resumida, podemos dizer que a condição de ergodicidade acontece quando uma realização do processo estocástico é capaz de representar todas as características do processo estocástico. Logo, esta é uma condição importantíssima para analisar uma série temporal.

Um processo é chamado de estacionário se suas características são mantidas ao longo do tempo. Em outras palavras se o processo evolui no tempo da mesma maneira independente da escolha de que tempo será analisado, temos um processo estacionário. Existem ainda dois tipos de estacionariedade: fraca (segunda ordem) ou forte (estrita).

Segundo Morettin e Toloi (2006), um processo estocástico é chamado de estritamente estacionário (estacionariedade forte) se todas as suas distribuições finito-dimensionais permanecem as mesmas sob translações no tempo. Ainda, de maneira particular, podemos afirmar que nesse tipo de processo a média e a variância são constantes ao longo tempo. Como em geral o interesse está em caracterizar o processo estocástico utilizando apenas um pequeno número de momentos, a estacionariedade fraca é a suposição que mais se utiliza quando fazemos análise de séries temporais.

Ainda segundo Morettin e Toloi (2006), um processo $Z = \{Z_t, t \in T\}$ é chamado de fracamente estacionário ou estacionário de segunda ordem se, e somente se:

- $E[Z_t] = \mu_t = \mu$, constante para todo $t \in T$
- $E[Z_t^2] < \infty$, para todo $t \in T$
- $\gamma_k = E[(Z_t - \mu)(Z_{t-k} - \mu)]$, para todo $t \in T$

Portanto, podemos pensar a série temporal como sendo uma realização de um processo estocástico ergódico. Uma maneira análoga de imaginar a relação entre uma série temporal e um processo estocástico é pensar como uma relação entre amostra e população, onde a amostra seria a série temporal e a população seria o processo estocástico. Logo, quando modelamos uma série temporal, estamos tentando via uma "amostra" inferir sobre todo um processo estocástico.

3.2 Previsão de Séries Temporais

Um dos objetivos da análise de séries temporais é a sua utilização para fazer previsões sobre valores futuros a respeito da mesma. Para tal, em geral, são utilizados valores passados e atuais de modo a estimar o determinado comportamento da série alguns passos à frente. Estes irão representar o horizonte ou intervalo de previsão a partir da origem (último valor observado).

Pode-se denotar $\hat{Z}_t(k)$ como a esperança condicional de Z_{t+k} dado os eventos ocorridos até o tempo t. Logo:

$$\hat{Z}_t(k) = E[Z_{t+k}|Z_t, Z_{t-1}, \dots] \tag{3-9}$$

Onde: Z_{t+k} são os valores desconhecidos que serão previstos para $k = 1, 2, 3, \dots$

Para fazer previsões, em geral, se utilizam modelos que se baseiam nas informações que se tem sobre a série em estudo. O procedimento de elaboração de um modelo que venha a prever deve se basear em alguma função perda. Por exemplo, o erro quadrático médio é utilizado em larga escala.

Existem diversos métodos de previsão, sejam esses baseados desde simples intuição até modelos quantitativos e complexos. Nessa dissertação, os modelos a serem utilizados são quantitativos, onde as características passadas são observadas e servem de guia para fazer extrapolações além do período de conhecimento do fenômeno observado.

3.3 Método Ingênuo

O método ingênuo, também conhecido como modelo Persistence ou ainda passeio aleatório pode ser considerado como um dos modelos mais simples e, como o próprio nome diz, ingênuos de se fazer previsões. Apesar disso, o mesmo continua a ser bastante utilizado com o objetivo de comparar a qualidade de modelos propostos para séries temporais e sendo portanto este o modelo benchmark em muitos estudos, dado a sua simplicidade. Tal modelo assume que a previsão de um evento que venha a ocorrer no instante $t+1$ será exatamente igual ao evento no instante t. Logo, podemos escrever o modelo da seguinte maneira:

$$Z_{t+1} = Z_t + \epsilon_t \tag{3-10}$$

Por consequência, a previsão k passos à frente $\hat{Z}_t(k)$ pode ser escrita da seguinte maneira:

$$\hat{Z}_t(k) = E[(Z_{t-1+k}|Z_t)] = E(Z_{t-1+k}) \tag{3-11}$$

Uma métrica muitas vezes utilizada que leva em consideração o método ingênuo é o coeficiente U de Theil. Este analisa a qualidade da previsão levando em consideração o erro apresentado pelo modelo proposto e o modelo ingênuo. Podemos escrever o mesmo da seguinte maneira:

$$U = \frac{\sqrt{\sum_{i=1}^{N} (z_t - \hat{z}_t)^2}}{\sqrt{\sum_{i=1}^{N} (z_t - z_{t-1})^2}} \tag{3-12}$$

O coeficiente pode ser interpretado da seguinte forma:

- $U > 1$ O erro do modelo estimado é maior que o do modelo ingênuo
- $U < 1$ O erro do modelo estimado é menor que o do modelo estimado

É importante ressaltar que em geral esperamos que o modelo apresente um resultado com coeficiente U de Theil menor que 1 e tenha assim erro menor que o da previsão apresentado por um modelo ingênuo (Persistence).

3.4 Metodologia Box e Jenkins

Os métodos de Box & Jenkins surgiram na década de 70 e desde então têm sido largamente usados na prática, consolidando-se como um dos métodos mais utilizados para se fazer previsões de séries temporais.

O objetivo nessa metodologia é através das informações contidas na série, detectar o mecanismo gerador da mesma de maneira que esse possa ser representado. Para tal, devemos fazer suposições de estacionariedade e ergodicidade de maneira que uma única realização do processo, série temporal disponível, seja capaz de descrever todo o processo.

Este tipo de modelagem é baseado na teoria geral de sistemas lineares. Segundo Souza e Camargo (2004), essa metodologia supõe que a passagem de um ruído branco por um filtro linear de memória infinita gera um processo estacionário de segunda ordem. Em outras palavras, o filtro linear é aquele que transforma o ruído branco (entrada) em série temporal.

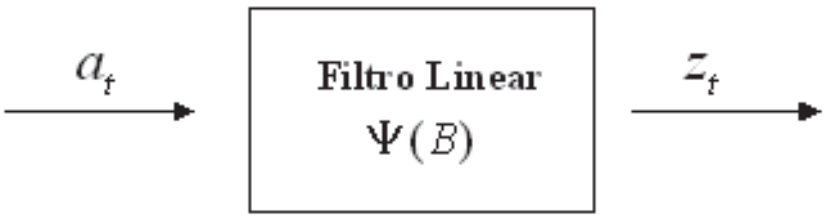

Figura 3.1: Entrada de ruído branco e geração de série temporal via filtro linear (geração de série temporal)

Portanto, ao modelar uma série via modelos de Box & Jenkins, estamos tentando descobrir um filtro (sistema) inverso que possa gerar um processo de ruído branco, mostrando assim que toda as características e estruturas da série foram obtidas.

O processo assume que a série temporal foi gerada através de um processo de segunda ordem. Isto implica que para utilizar séries não estacionárias é

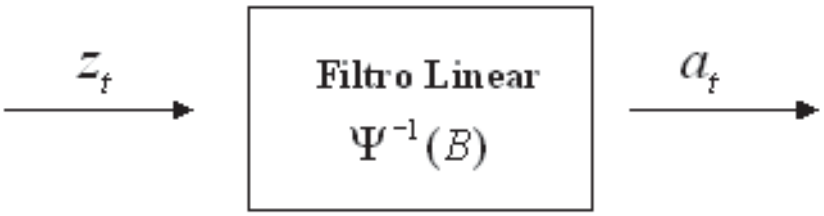

Figura 3.2: Entrada da série temporal e saída de ruído branco via filtro linear (análise de série temporal)

necessário aplicar diferenciações e/ou transformações na série de maneira a torná-la estacionária.

Os modelos dessa metodologia podem ser divididos da seguinte maneira:

- Modelos Autoregressivos (AR): as próprias observações passadas da série são utilizadas para explicar o comportamento da mesma
- Modelos de médias móveis (MA): os resíduos passados são utilizados de maneira a explicar os valores futuros
- Modelos Autoregressivos de Médias Móveis (ARMA): utilizam tanto os valores passados da série quanto os resíduos passados para explicitar os valores futuros
- Modelos Autoregressivos Integrados de Médias Móveis (ARIMA): Semelhantes aos modelos ARMA com a ressalva de aplicar sucessivas diferenças à série com o objetivo de torná-la estacionária
- Modelos Sazonais Autoregressivos Integrados de Médias Móveis (SARIMA): Utilizados quando temos séries com presença de padrões sazonais

O modelo de Box & Jenkins segue, então, a seguinte formulação:

$$w_t = \Psi(B)a_t = \sum_{k=0}^{\infty} \Psi_k a_{t-k} \tag{3-13}$$

Onde a_t é um ruído branco e B é o operador retardo que pode ser definido como:

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

$$B^k Z_t = Z_{t-k} \tag{3-14}$$

$$\nabla = (1 - B)$$

O polinômio $\Psi(B)$ é de ordem infinita. Entretanto, segundo certas condições, temos que todo polinômio infinito pode ser escrito como o quociente de dois polinômios finitos. Logo, podemos expressar o polinômio $\Psi(B)$ da seguinte maneira:

$$\Psi(B) = \frac{\theta(B)}{\phi(B)} \tag{3-15}$$

Onde:
$\theta(B)$, $1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$, representa o polinômio de médias móveis $MA(q)$
$\phi(B)$, $1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$, representa o polinômio de autoregressivo AR(p)

Generalizando, temos o modelo ARIMA(p,d,q):

$$\phi(B)\nabla^d Z_t = \theta(B)a_t \tag{3-16}$$

Onde d, que ainda não havia sido definido, é um inteiro maior ou igual a zero que representa a diferenciação a ser aplicada à série.

O processo de utilização dos modelos ARIMA segue uma ordem natural que pode ser descrita da seguinte maneira:

Identificação

O objetivo nessa fase é identificar as ordens (p,d,q) quando o modelo não tem presença de sazonalidade, além das ordens (P,D,Q) quando o modelo apresenta componente sazonal (SARIMA). Em geral são utilizadas para tal as funções de autocorrelação (FAC), autocorrelação parcial (FACP) e os respectivos correlogramas.

Estimação

Após a fase de identificação, é necessário estimar os parâmetros autoregressivos e de médias móveis. Além disso, caso o modelo tenha presença de componente sazonal, devemos estimar os parâmetros autoregressivos e de médias móveis sazonais. Em geral, para essas estimações utilizamos a função de máxima verossimilhança condicional. Esta é equivalente a estimar via mínimos quadrados quando a_t é proveniente de uma distribuição normal.

A idéia ao estimar os parâmetros desconhecidos via mínimos quadrados é minimizar a soma de quadrado dos erros. Logo, nos modelos ARIMA(p,d,q), temos que minimizar a seguinte soma:

$$S(\xi) = \sum_{t=1}^{T} a_t^2 \tag{3-17}$$

Onde:

$$a_t = \theta^{-1}(B)\phi(B)\nabla^d Z_t \tag{3-18}$$

Assim como já foi dito, sob a suposição de normalidade, podemos estimar os parâmetros de interesse via máxima verossimilhança.

Verificação

Após as fases de identificação e estimação, o modelo deve ser submetido a verificação que atestem que o mesmo se comporta de maneira adequada aos dados utilizados. Para tal, em geral são realizados testes que garantem a qualidade do modelo proposto. A análise dos resíduos produzidos pelo modelo tem caráter fundamental nessa etapa da modelagem. Esses devem apresentar um comportamento característico de ruído branco, atestando que as características da série foram captadas utilizando o modelo escolhido.

Devemos ressaltar que se a conclusão obtida na etapa de verificação for de que o modelo não está adequado, o processo volta a primeira etapa de identificação e outro modelo deve ser escolhido, estimado e verificado. Outra ressalva decorre do fato de que podem existir diversos modelos adequados cabendo ao pesquisador a tarefa de escolher qual se adapta melhor ao seu objetivo (previsão, decomposição etc).

Previsão

O último passo a ser executado na modelagem é a fase de previsão. O modelo aprovado nas fases de identificação, estimação e verificação está, então, apto a fazer previsões.

A idéia aqui é prever um valor Z_{t+k}, onde $k \geq 1$ e representa o horizonte de previsão. Em geral, todos os valores até o instante t, origem da previsão, são previamente conhecidos. Podemos, então, denotar a previsão no instante t de Z no horizonte de previsão k como $\hat{Z}_t(k)$. De maneira ilustrativa, podemos representar a equação de previsão de um modelo ARMA(p,q) com horizonte k e $k \geq 1$, via equação de diferenças como:

$$\begin{aligned} \hat{Z}_t(k) &= E[Z_{t+k}|Z_t] \\ &= \phi_1[Z_{t+k-1}] + \cdots + \phi_p[Z_{t+k-p}] + \\ &+ \theta_1[a_{t+k-1}] + \cdots + \theta_q[a_{t+k-q}] + [a_{t+k}] \end{aligned} \tag{3-19}$$

Devemos assumir para tal o seguinte:

$$\begin{aligned} &[Z_{t+k}] = \hat{Z}_t(k) \quad \text{onde } k > 0 \\ &[Z_{t+k}] = Z_{t+k} \quad \text{onde } k \leq 0 \\ &[a_{t+k}] = 0 \quad \text{onde } k > 0 \\ &[a_{t+k}] = a_{t+k} \quad \text{onde } k \leq 0 \end{aligned}$$

3.5 Modelo ARFIMA

Os modelos Autoregressivos Fracionários Integrados de Médias Móveis (ARFIMA) podem ser entendidos como uma extensão natural dos modelos ARIMA onde o parâmetro d pode agora assumir valores reais e não mais apenas inteiros maiores ou iguais a zero como havia sido previamente definido. O relaxamento da obrigatoriedade, antes definida para o parâmetro d, permite que seja possível a modelagem do comportamento de memória longa de maneira natural, sem a inclusão de um número excessivo de parâmetros. O parâmetro d passa agora a captar o comportamento de longo prazo (memória longa) deixando a tarefa de captar o comportamento de curto prazo para os parâmetros p e q.

Segundo Morettin e Toloi (2006), um processo de memória longa é um

processo estacionário em que a função de autocorrelação decresce hiperbolicamente (suavemente) para zero. Os autores afirmam ainda que em processos de memória longa, existe uma significativa dependência entre distantes observações amostrais da função de autocorrelação. Além disso, neste tipo de processo, temos que a função densidade espectral é não limitada na frequência zero, sendo equivalente a dizer que temos uma função de autocorrelação que não é absolutamente somável.

Podemos, então, definir um processo ARFIMA da seguinte maneira: Seja o processo Z_t um processo estacionário, da forma:

$$\phi(B)\nabla^d Z_t = \theta(B)a_t \tag{3-20}$$

Onde:

$$\frac{-1}{2} \leq d \leq \frac{1}{2}$$

Cabe ressaltar que para $\nabla^d = (1-B)^d$, onde d é um número real, podemos reescrever ∇^d da seguinte forma:

$$\nabla^d = \sum_{k=0}^{\infty} \binom{d}{k} (-1)^k B^k \tag{3-21}$$

Sendo:

$$\binom{d}{k} = \frac{d!}{k!(d-k)!} = \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)} \tag{3-22}$$

É importante ressaltar que os modelos ARFIMA (p,d,q) são estacionários e invertíveis quando atendem as seguintes condições:

- Estacionariedade: Se todas as raízes de $\phi(B) = 0$ estão fora do círculo unitário e $d < \frac{1}{2}$.
- Invertibilidade: Se todas as raízes de $\theta(B) = 0$ estão fora do círculo unitário e $d > -\frac{1}{2}$.

Como já foi dito antes, um processo de memória longa, nesse caso ARFIMA, apresenta declínio até o valor zero na sua função de autocorrelação. A título de ilustração, podemos escrever a função de autocovariância para um processo ARFIMA (0,d,0), onde assumimos que $\mu = E(Z_t) = 0$, μ é um ruído branco e o processo é estacionário e invertível, da seguinte maneira:

$$\gamma(k) = \sigma^2 \frac{(-1)^k \Gamma(1-2d)}{\Gamma(k-d+1)\Gamma(1-k-d)} \tag{3-23}$$

Seguindo as mesmas condições podemos escrever a função de autocorrelação da seguinte forma:

$$\rho(k) = \frac{\Gamma(1-d)\Gamma(k+d)}{\Gamma(d)\Gamma(k+1-d)}$$

$$\rho(k) \sim \frac{\Gamma(1-d)}{\Gamma(d)} |k|^{2d-1} \quad |k| \to \infty \tag{3-24}$$

É fundamental estimar bem o parâmetro d de longa dependência. Em geral, para tal fim, podemos utilizar uma extensa gama de métodos. Não é incomum que tais métodos sejam encontrados divididos em três grupos: Métodos Heurísticos, Métodos de Máxima Verossimilhança e Métodos de Estimação Robusta. Utilizamos na maioria da vezes métodos de máxima verossimilhança ou métodos semi-paramétricos. A seguir, serão apresentados dois métodos de estimação bastante utilizados.

3.5.1
Estimação por Máxima Verossimilhança

Dado um processo ARFIMA (p,d,q), temos que a função de Máxima Verossimilhança de $Z = (Z_1, Z_2, \ldots, Z_n)$ pode ser escrita da seguinte forma:

$$L(\eta, \sigma_a^2) = (2\pi\sigma_a^2)^{-\frac{n}{2}} (r_0 \ldots r_{\eta-1})^{-\frac{1}{2}} exp\left[-\frac{1}{2\sigma_a^2} \sum_{j=1}^{n} \frac{(Z_j - \hat{Z}_j)^2}{r_{j-1}} \right] \tag{3-25}$$

Onde:
$\eta = (d, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)$
$\hat{Z}_j = 1, 2, \ldots, n$ são as previsões um passo à frente e
$r_{j-1} = (\sigma_a^2)^{-1} E(Z_j - \hat{Z}_j)^2$

Podemos, então, usar os seguintes estimadores de máxima verossimilhança para os parâmetros de interesse:

$$\hat{\sigma}_{MV}^2 = n^{-1} S(\hat{\eta}_{MV}) \tag{3-26}$$

Sendo:

$$S(\hat{\eta}_{MV}) = \sum_{j=1}^{n} \frac{(Z_j - \hat{Z}_j)^2}{r_{j-1}} \tag{3-27}$$

Devemos, então, minimizar $l(\eta)$ via $\hat{\eta}_{MV}$, entretanto esse é um processo computacionalmente custoso. Para contornar esse problema, em geral, se utiliza uma aproximação de $l(\eta)$ que considera o periodograma dos dados e a função densidade espectral do processo Z_t. Logo, temos que:

$$l(\eta) \simeq l_*(\eta) = ln\frac{1}{n}\sum_{j} \frac{I_n(\omega_j)}{2\pi f(\omega_j;\eta)} \tag{3-28}$$

onde $I_n(\omega_j)$ é o periodograma dos dados e pode ser escrito da seguinte forma:

$$l_\eta(\omega_j) = \frac{1}{\eta}\left|\sum_{i=1}^{\eta} Z_t e^{-it\omega_j}\right|^2 \tag{3-29}$$

e ω_j é a frequência de Fourier, tal que $\omega_j = \dfrac{2\pi j}{n} \in (-\pi, \pi]$. Além disso, temos ainda a função densidade espectral $f(\omega_j;\eta)$ que pode ser escrita da seguinte forma:

$$f(\omega_j;\eta) = \frac{\sigma_a^2|1 - \theta_1 e^{-i\omega_j} - \cdots - \theta_q e^{-qi\omega_j}|^2}{2\pi|1 - \phi_1 e^{-i\omega_j} - \cdots - \phi_p e^{-pi\omega_j}|^2}|1 - e^{-i\omega_j}|^{-2d} \tag{3-30}$$

3.5.2
Estimação por Método de Regressão Utilizando o Periodograma

Segundo Morettin e Toloi (2006), o método de regressão utilizando periodograma, proposto por Geweke e Porter-Hudak (1983) é baseado na equação que exibe a relação entre a função densidade espectral de um processo ARFIMA (p,d,q) e de um processo ARMA (p,q).

Supondo um processo ARFIMA (p,d,q), podemos escrever sua função de densidade espectral como:

$$f_z(\lambda) = |1 - e^{-i\lambda}|^{-2d} f_u(\lambda) \tag{3-31}$$

onde $f_u(\lambda)$ é a função de densidade espectral do processo de um processo U_t que é um ARMA (p,q) e pode ser escrita da seguinte forma:

$$f_u(\lambda) = \frac{\sigma^2|\theta(e^{-i\lambda})|^2}{2\pi|\phi(e^{-i\lambda})|^2} \tag{3-32}$$

Fazendo algumas operações algébricas, a substituição de λ por $\lambda_j = \frac{2\pi j}{n}$ e utilizando o fato de que $I_z(\lambda_j)$ é um estimador de $f_z(\lambda_j)$, podemos escrever um modelo de regressão linear da forma:

$$Y_j = a - dX_j + \epsilon_j, \quad j = 1, 2, \ldots, m \tag{3-33}$$

onde:
$Y_j = lnI_z(\lambda_j)$

$X_j = ln\left(4sen^2\left(\frac{\lambda_j}{2}\right)\right)$

$\epsilon_j = ln\left(\frac{I_z(\lambda_j)}{f_z(\lambda_j)}\right)$,

$a = lnf_u(0)$ e $m = n^\alpha, \quad 0 < \alpha < 1$

Podemos agora proceder com um método de estimação de parâmetros de regressão como o de mínimos quadrados para o parâmetro d de interesse. Logo, temos que:

$$\hat{d}_{MQ} = -\frac{\sum_{i=1}^{m}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{m}(X_i - \bar{X})^2} \tag{3-34}$$

Após a estimação de d via mínimos quadrados devemos calcular a transformada discreta de Fourier da serie Z_t

$$d_Z(\lambda_i) = \sum_{t=1}^{n} Z_t e^{-\lambda_i t} \tag{3-35}$$

Em seguida deve ser calculado:

$$d_U(\lambda_i) = (1 - e^{-\lambda_i})^{\hat{d}_{MQ}} d_Z(\lambda_i), \quad i = 0, \ldots, n-1 \tag{3-36}$$

Deve ser calculado ainda a transformada inversa de Fourier já que a mesma produz uma estimativa da série sem a presença do componente de memória longa. Então, temos que:

$$\tilde{U}_t = \frac{1}{n}\sum_{j}^{n-1} e^{i\omega_j t} d_U(\lambda_j) \tag{3-37}$$

Após os cálculos das equações acima, podemos proceder para a fase de identificação dos parâmetros p e q. Tal identificação é feita da mesma maneira que fazemos com um modelo ARMA (p,q), já que temos agora a série livre do componente de memória longa. Logo, utilizamos as funções de autocorrelação e autocorrelação parcial de $\tilde{U}_t$. Por fim, devemos calcular todos os parâmetros do modelo ARFIMA (p,d,q) de maneira conjunta, inclusive o parâmetro d, via máxima verossimilhança.

3.6 Análise Harmônica

A idéia principal ao utilizar a Análise Harmônica, ou Análise de Fourier, é determinar que ciclos de determinada frequência são importantes para descrever o comportamento da série. Adicionado a isso, quando estamos nesse tipo de análise (no domínio da frequência), podemos fazer aproximações de um conjunto de dados utilizando uma combinação de funções senoidais, onde os coeficientes são transformadas de Fourier discretas da série.

Em geral, ao utilizar Análise Harmônica, estamos interessados nas periodicidades apresentadas pela série em estudo. Logo, podemos estar em busca da estimação de amplitudes e fases quando as frequências são conhecidas. Ou ainda, podemos não conhecer a frequência nos dados e, portanto, devemos estimá-la e por consequência devemos estimar as amplitudes e as fases.

Segundo Morettin e Toloi (2006), podemos escrever um modelo com uma única periodicidade da seguinte maneira:

$$Z_t = \mu + Acos(\omega t) + Bsen(\omega t) + \epsilon_t \tag{3-38}$$

Onde:

$A = Rcos\phi$
$B = -Rsen\phi$

Sendo ω a frequência, R a amplitude, ϕ o ângulo da fase e ϵ_t o componente aleatório. Temos ainda que:

$$R^2 = A^2 + B^2 \tag{3-39}$$

e portanto:

$$\phi = \begin{cases} arctg(-\frac{B}{A}) & , A > 0, \\ arctg(-\frac{B}{A}) - \pi & , A < 0, B > 0, \\ arctg(-\frac{B}{A}) + \pi & , A < 0, B < 0, \\ -\frac{\pi}{2} & , A = 0, B > 0, \\ \frac{\pi}{2} & , A = 0, B < 0, \\ \text{arbitrário} & , A = 0, B = 0 \end{cases} \tag{3-40}$$

Devemos, então, estimar os parâmetros desconhecidos μ, A e B dados um w que podemos conhecer ou não.

3.6.1 Estimação via mínimos quadrados quando a frequência é conhecida

Utilizamos os estimadores $\hat{\mu}$, $\hat{A}$ e $\hat{B}$ como estimadores dos parâmetros μ, A e B de interesse. Podemos obter tais estimadores minimizando a soma de quadrados dos erros utilizando a equação:

$$SQR(\mu, A, B) = \sum_{t=1}^{N} (Z_t - \mu - Acos(\omega t) - Bsen(\omega t))^2 \tag{3-41}$$

Podemos escrever de forma matricial o modelo harmônico e dessa maneira encontrar os estimadores de mínimos quadrados de maneira mais trivial já que minimizar a soma de quadrados dos erros nos leva a um conjunto de equações cujo a solução se torna complicada. Logo de maneira matricial temos que:

$$Z = W\theta + \epsilon \tag{3-42}$$

onde:

$$Z = (Z_1, Z_2, \ldots, Z_N)'$$

$$\theta = (\mu, A, B),$$

$$W = \begin{bmatrix} 1 & cos(\omega) & sen(\omega) \\ 1 & cos(2\omega) & sen(2\omega) \\ \vdots & \vdots & \vdots \\ 1 & cos(N\omega) & sen(N\omega) \end{bmatrix}$$

de onde podemos obter o estimador de θ:

$$\theta = (W'W)^{-1}W',$$

com

$$W'W = \begin{bmatrix} N & \sum_{t=1}^{N} cos(\omega t) & \sum_{t=1}^{N} sen(\omega t) \\ \sum_{t=1}^{N} cos(\omega t) & \sum_{t=1}^{N} cos^2(\omega t) & \sum_{t=1}^{N} cos(\omega t)sen(\omega t) \\ \sum_{t=1}^{N} sen(\omega t) & \sum_{t=1}^{N} cos(\omega t)sen(\omega t) & \sum_{t=1}^{N} sen^2(\omega t) \end{bmatrix}$$

Logo, os estimadores de mínimos quadrados para os parâmetros de interesse, resultantes das equações acima, quando conhecemos as frequências são:

$$\begin{aligned} \hat{\mu} &= \frac{1}{N}\sum_{t=1}^{N} Z_t = \bar{Z} \\ \hat{A} &= \frac{2}{N}\sum_{t=1}^{N} Z_t cos(\omega t), \quad \omega \neq \pi; \hat{B} = 0 \\ \hat{B} &= \frac{2}{N}\sum_{t=1}^{N} Z_t sen(\omega t), \quad \omega = \pi; \hat{B} = 0; \\ \hat{A} &= \frac{1}{N}\sum_{t=1}^{N} Z_t(-1)^t, \qquad \omega = \pi \\ \hat{R} &= \hat{A}^2 + \hat{B}^2 \\ \hat{\phi} &= arctg\left(-\frac{\hat{B}}{\hat{A}}\right) \end{aligned} \tag{3-43}$$

Segundo Bloomfield (2000), teremos as seguintes aproximações, quando $\omega \neq \frac{2\pi k}{N}$ e não muito próxima de zero:

$$\tilde{\mu} = \hat{\mu} = \bar{Z},$$

$$\hat{A} = \frac{2}{N}\sum_{t=1}^{N}(Z_t - \bar{Z})cos(\omega t),$$

$$\hat{B} = \frac{2}{N}\sum_{t=1}^{N}(Z_t - \bar{Z})sen(\omega t), \tag{3-44}$$

$$\tilde{R} = \tilde{A}^2 + \tilde{B}^2,$$

$$\tilde{\phi} = arctg\left(-\frac{\tilde{B}}{\tilde{A}}\right)$$

3.6.2 Modelo com Periodicidades Múltiplas

Existem diversos casos de séries que apresentam mais de um período e portanto o caso acima, modelo com periodicidade única, pode ser estendido para que permita periodicidade múltipla. Então, podemos escrever um modelo com k períodos da seguinte forma:

$$Z_t = \mu + A_1 cos(\omega_1 t) + B_1 sen(\omega_1 t) + \ldots + A_k cos(\omega_k t) + B_k sen(\omega_k t) + \epsilon_t \tag{3-45}$$

onde $\omega_1, \omega_2, \ldots, \omega_k$ são as frequências que podem ser conhecidas ou não.

De maneira similar ao modelo com um período único, podemos estimar os parâmetros desconhecidos via mínimos quadrados. Então temos as seguintes matrizes:

$$Z = (Z_1, Z_2, \ldots, Z_N)'$$

$$\theta = (\mu, A_1, \ldots, A_k, B_1, \ldots, B_k)'$$

$$W = \begin{bmatrix} 1 & cos(\omega_1) & sen(\omega_1) & \ldots & cos(\omega_k) & sen(\omega_k) \\ 1 & cos(2\omega_1) & sen(2\omega_1) & \ldots & cos(2\omega_k) & sen(2\omega_k) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & cos(N\omega_1) & sen(2\omega_1) & \ldots & cos(N\omega_k) & sen(N\omega_k) \end{bmatrix}$$

Para as frequências de Fourier, $\omega_1 = \frac{2\pi i}{N}, \ldots, \omega_k = \frac{2\pi m}{N}$, temos as seguintes soluções exatas para os parâmetros de interesse:

$$\begin{aligned} \hat{\mu} &= \bar{Z}, \\ \hat{A}_i &= \frac{2}{N}\sum_{t=1}^{N} Z_t cos(\omega_i t), \quad i = 1, 2, \ldots, k \\ \hat{B}_i &= \frac{2}{N}\sum_{t=1}^{N} Z_t sen(\omega_i t), \quad i = 1, 2, \ldots, k \end{aligned} \tag{3-46}$$

Para $\omega = \pi$ temos ainda que:

$$\hat{A}_i = \frac{2}{N}\sum_{t=1}^{N} Z_t(-1)^t$$

$$\hat{B}_i = 0$$

3.6.3
Estimação via mínimos quadrados quando a frequência é desconhecida

A diferença nesse caso está no fato de não conhecermos o parâmetro ω sendo necessária a inclusão do mesmo entre os parâmetros a serem estimados.

Segundo Morettin e Toloi (2006), o melhor valor para ω que minimiza a soma de quadrados, é o valor que minimiza a soma de quadrados residual, o que é equivalente a maximizar:

$$\tilde{R}^2(\omega) = \tilde{A}^2(\omega) + \tilde{B}^2(\omega) \tag{3-47}$$

Que é equivalente a maximizar o periodograma:

$$\begin{aligned} I(\omega) &= \frac{N}{8\pi}\tilde{R}^2(\omega) \\ &= \frac{1}{2\pi N}\left[\left(\sum_{t=1}^{N}(Z_t-\bar{Z})cos(\omega t)\right)^2+\left(\sum_{t=1}^{N}(Z_t-\bar{Z})sen(\omega t)\right)\right] \end{aligned} \tag{3-48}$$

O periodograma pode ser entendido como uma ferramenta apropriada para descobrir frequências, devendo apresentar picos na função quando estamos passa por uma frequência verdadeira dos dados.

Os autores Morettin e Toloi (2006) afirmam ainda que é possível estimar os parâmetros utilizando técnicas de minimização de funções não lineares.

3.7 Métricas de Comparação

Com o intuito de comparar os resultados obtidos, serão utilizadas as métricas de comparação correspondentes a raiz do erro quadrático médio (RMSE), ao erro percentual médio absoluto (MAPE) e ao erro absoluto médio (MAE).

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(Z_t-\hat{Z}_t)^2} \tag{3-49}$$

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{(Z_t-\hat{Z}_t)}{Z_t}\right|\cdot 100 \tag{3-50}$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|(Z_t-\hat{Z}_t)| \tag{3-51}$$