

Estimação da densidade espectral

Paulo Lopes dos Santos

Processamento Estatístico de Sinal, 2024

DEEC, FEUP

Resumo I

1 Introdução

2 Estimação paramétrica

● Processos Ar

- Estimador de Yule-Walker
- Algoritmo de Levinson-Durbin
- Estimador do erro de previsão
- Relação entre os estimadores de Yule-Walker e de erro de previsão

● Determinação da ordem do modelo - Correlação parcial

● Processos MA

- Método da Função Geradora da Autocovariância
- Método da resposta impulsional truncada

Resumo II

- Método dos Mínimos Quadrados para Modelos MA
- Método de Minimização do Erro de Previsão
- Determinação da ordem de modelos MA - Sequência de Autocorrelação
- Processos ARMA
 - Método dos Esfasamentos Superiores para Estimativa da Parte Auto-Regressiva de um Modelo ARMA
 - Método de Ajuste da Resposta Impulsional
 - Método de Mínimos Quadrados para Modelos ARMA
 - Método de Minimização do Erro de Previsão

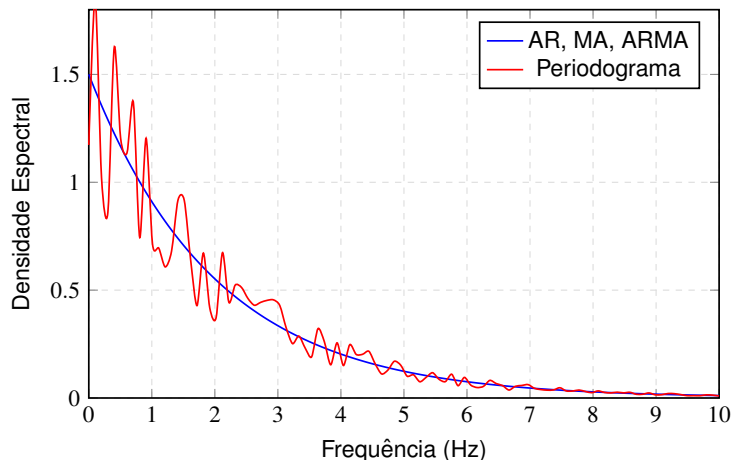
Estimação da densidade espectral

- Essencial no processamento estatístico de sinal.
- Aplicações em diversas áreas:
 - Telecomunicações
 - Processamento de áudio
 - Geofísica
 - etc.
- A densidade espectral revela como a energia de um sinal se distribui pelas frequências facilitando
 - Detecção de fenómenos periódicos
 - Filtragem eficaz (remoção de ruído, compressão de dados)
 - Modelação e previsão de séries temporais (economia e finanças, controlo)

Métodos de estimação de espectro

- Métodos paramétricos:
 - Baseiam-se em modelos como AR, MA, ARMA.
 - Úteis para conjuntos de dados limitados ou com conhecimento a priori.
- Métodos não-paramétricos:
 - Não fazem suposições sobre o processo estocástico.
 - Exemplos: Periodograma, Blackman-Tukey.
 - Vantagens: maior flexibilidade.
 - Desvantagens: podem apresentar alta variância e baixa resolução espectral.

Exemplo Método Paramétrico (AR,MA,ARMA)/ Método não paramétrico (Periodograma)



Estimação Paramétrica - Processos AR

- Processos AR são descritos pela equação:

$$A(q^{-1})y(t) = e(t),$$

- $A(q^{-1}) = 1 + a_1q^{-1} + \dots + a_nq^{-n}$
- $e(t)$: Ruído branco com média nula e variância σ^2 .
- A representação do processo:

$$y(t) = \frac{1}{A(q^{-1})}$$

Densidade Espectral de Potência dum Processo AR

- Os processos AR são estacionários
- A densidade espectral de potência é dada por:

$$\Phi_{yy}(\omega) = \frac{\sigma^2}{|A(e^{-j\omega})|^2}$$

- Relacionada diretamente com o modelo
- Estimação espectral depende da determinação dos coeficientes $a_1, \dots, a_n$.

Determinação dos Parâmetros de Modelos AR

- Os parâmetros de um modelo AR podem ser determinados a partir da sequência de autocovariância.
- Para um processo AR(n), os primeiros $n + 1$ elementos da sequência de autocovariância são obtidos pelo sistema de equações:

$$\begin{cases} \lambda_{yy}(0)a_1 + \lambda_{yy}(1)a_2 + \cdots + \lambda_{yy}(n-1)a_n &= -\lambda_{yy}(1) \\ \lambda_{yy}(1)a_1 + \lambda_{yy}(0)a_2 + \cdots + \lambda_{yy}(n-2)a_n &= -\lambda_{yy}(2) \\ &\vdots \\ \lambda_{yy}(n-1)a_1 + \lambda_{yy}(n-2)a_2 + \cdots + \lambda_{yy}(0)a_n &= -\lambda_{yy}(n), \end{cases}$$

- Estas são as Equações de Yule-Walker, usadas para estimar os parâmetros $a_1, \dots, a_n$.

Resolução Matricial das Equações de Yule-Walker

- As equações podem ser rescritas na forma matricial:

$$\mathbf{\Lambda}_n \boldsymbol{\theta}^{(n)} = -\boldsymbol{\lambda}_{1:n},$$

$$\bullet \mathbf{\Lambda}_n = \begin{bmatrix} \lambda_{yy}(0) & \lambda_{yy}(1) & \cdots & \lambda_{yy}(n-1) \\ \lambda_{yy}(1) & \lambda_{yy}(0) & \cdots & \lambda_{yy}(n-2) \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{yy}(n-1) & \lambda_{yy}(n-2) & \cdots & \lambda_{yy}(0) \end{bmatrix}$$

$$\bullet \boldsymbol{\theta}^{(n)} = [a_1 \ a_2 \ \cdots \ a_n]^T$$

$$\bullet \boldsymbol{\lambda}_{1:n} = [\lambda_{yy}(1) \ \lambda_{yy}(2) \ \cdots \ \lambda_{yy}(n)]^T$$

- Os parâmetros são estimados por:

$$\boldsymbol{\theta}^{(n)} = -\mathbf{\Lambda}_n^{-1} \boldsymbol{\lambda}_{1:n}.$$

Cálculo da Variância do Ruído Branco

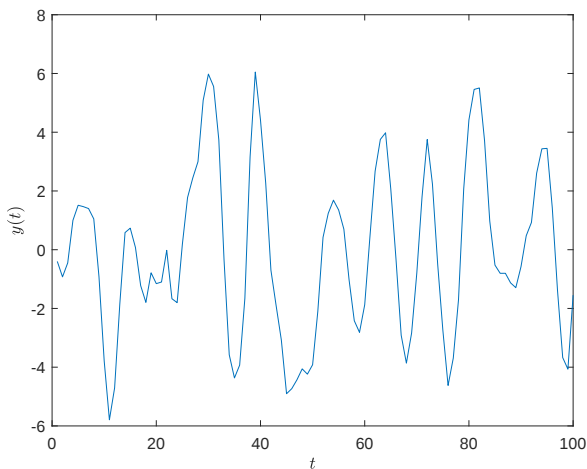
- σ^2 , variância do ruído branco é calculada por

$$\sigma^2 = \lambda_{yy}(0) + \lambda_{yy}(1)a_1 + \cdots + \lambda_{yy}(n)a_n.$$

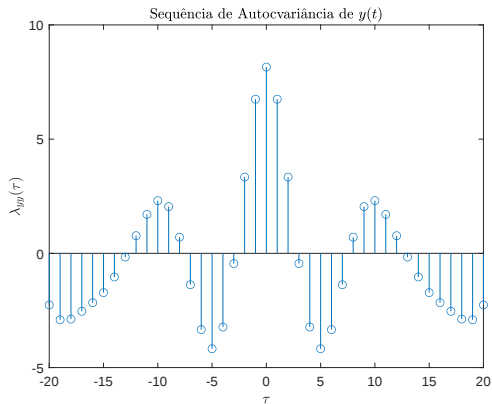
- Em casos reais, usam-se estimativas da sequência de autocovariância para calcular os parâmetros do modelo.

Exemplo - Identificação dum modelo AR(3)

Realização dum processo AR(3)



Exemplo - Identificação dum modelo AR(3)



- Sequência de autocovariância tem duração infinita

Exemplo - Identificação dum modelo AR(3)

τ	0	1	2	3
$\hat{\lambda}_{yy}(\tau)$	8.1539	6.7480	3.3389	-0.4452

$$\hat{\mathbf{\Lambda}}_3 = \begin{bmatrix} \hat{\lambda}_{yy}(0) & \hat{\lambda}_{yy}(1) & \hat{\lambda}_{yy}(2) \\ \hat{\lambda}_{yy}(1) & \hat{\lambda}_{yy}(0) & \hat{\lambda}_{yy}(1) \\ \hat{\lambda}_{yy}(2) & \hat{\lambda}_{yy}(1) & \hat{\lambda}_{yy}(0) \end{bmatrix} = \begin{bmatrix} 8.1539 & 6.7480 & 3.3389 \\ 6.7480 & 8.1539 & 6.7480 \\ 3.3389 & 6.7480 & 8.1539 \end{bmatrix}$$

$$\hat{\mathbf{\lambda}}_{1:3} = [\hat{\lambda}_{yy}(1) \quad \hat{\lambda}_{yy}(2) \quad \hat{\lambda}_{yy}(3)]^T = [6.7480 \quad 3.3389 \quad -0.4452]^T$$

$$\begin{aligned} \hat{\boldsymbol{\theta}} = \hat{\mathbf{\Lambda}}_3^{-1} \hat{\mathbf{\lambda}}_{1:3} &= \begin{bmatrix} 8.1539 & 6.7480 & 3.3389 \\ 6.7480 & 8.1539 & 6.7480 \\ 3.3389 & 6.7480 & 8.1539 \end{bmatrix}^{-1} \begin{bmatrix} 6.7480 \\ 3.3389 \\ -0.4452 \end{bmatrix}^T = \\ &= \begin{bmatrix} -1.9462 \\ 1.5755 \\ -0.4523 \end{bmatrix} \end{aligned}$$

Exemplo - Identificação dum modelo AR(3)

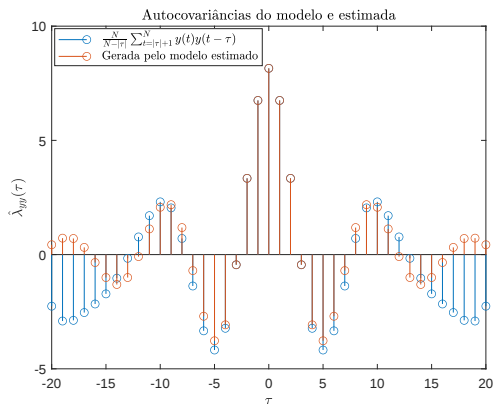
- Modelo estimado:

$$(1 - 1.9462q^{-1} + 1.5755q^{-2} - 0.4523q^{-3}) y(t) = e(t)$$

- Variância de $e(t)$:

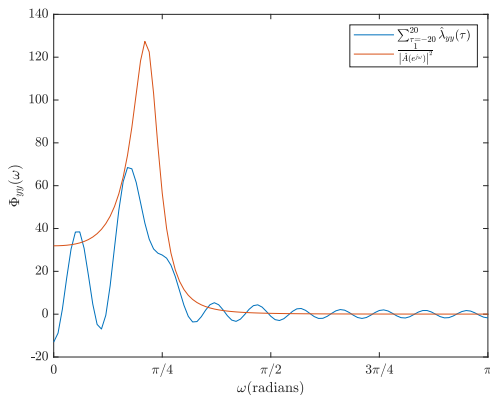
$$\begin{aligned}\hat{\sigma}^2 &= \lambda_{yy}(0) + \lambda_{yy}(1)a_1 + \cdots + \lambda_{yy}(n)a_n = \\ &= 8.1539 + 6.7480 \times (-1.9462) + 3.3389 \times 1.5755 + \\ &\quad (-0.4452) \times (-0.4523) = 0.4825\end{aligned}$$

Exemplo - Identificação dum modelo AR(3)



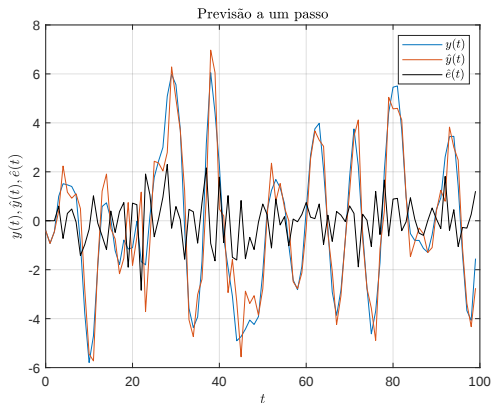
A sequência de autocorrelação estimada coincide com a do modelo para $\tau = -3, \dots, 3$, pois o modelo foi ajustado igualando os primeiros quatro pontos da sequência de autocovariância aos valores estimados.

Exemplo - Identificação dum modelo AR(3)



O espectro estimado a partir da autocovariância é pouco preciso porque o número de amostras de $y(t)$ é pequeno ($N = 100$).

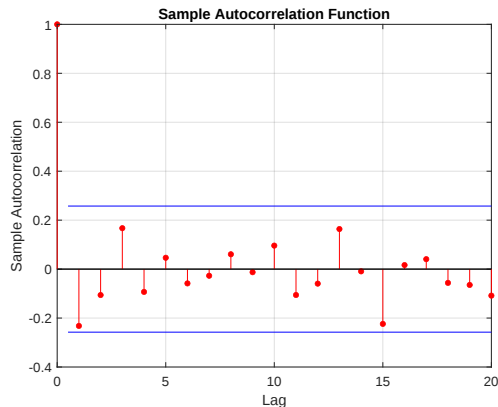
Exemplo - Identificação dum modelo AR(3)



Variância do erro de previsão: $E\{\hat{e}^2(t)\} = 0.8759$

Significativamente diferente de 0.4825 (valor indicado pelo modelo estimado).

Exemplo - Identificação dum modelo AR(3)



Autocovariância de $\hat{e}(t)$ evidencia que é ruído branco.

Algoritmo de Levinson-Durbin

- O estimador de Yule-Walker resolve a equação

$$\Lambda_n \theta^{(n)} = -\lambda_{1:n},$$

em que Λ_n é a matriz das covariâncias:

$$\Lambda_n = \begin{bmatrix} \lambda_{yy}(0) & \lambda_{yy}(1) & \cdots & \lambda_{yy}(n-1) \\ \lambda_{yy}(1) & \lambda_{yy}(0) & \cdots & \lambda_{yy}(n-2) \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{yy}(n-1) & \lambda_{yy}(n-2) & \cdots & \lambda_{yy}(0) \end{bmatrix}$$

$\theta^{(n)}$ o vetor de parâmetros:

$$\theta^{(n)} = [a_1 \quad a_2 \quad \cdots \quad a_n]^T$$

e:

$$\lambda_{1:n} = [\lambda_{yy}(1) \quad \lambda_{yy}(2) \quad \cdots \quad \lambda_{yy}(n)]^T$$

Matriz das covariâncias

$$\begin{bmatrix} \lambda_{yy}(0) & \lambda_{yy}(1) & \lambda_{yy}(2) & \cdots & \lambda_{yy}(n-2) & \lambda_{yy}(n-1) \\ \lambda_{yy}(1) & \lambda_{yy}(0) & \lambda_{yy}(1) & \cdots & \lambda_{yy}(n-3) & \lambda_{yy}(n-2) \\ \lambda_{yy}(2) & \lambda_{yy}(1) & \lambda_{yy}(0) & \ddots & \lambda_{yy}(n-4) & \lambda_{yy}(n-3) \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \lambda_{yy}(n-2) & \lambda_{yy}(n-3) & \lambda_{yy}(n-4) & \ddots & \lambda_{yy}(0) & \lambda_{yy}(1) \\ \lambda_{yy}(n-1) & \lambda_{yy}(n-2) & \lambda_{yy}(n-3) & \cdots & \lambda_{yy}(1) & \lambda_{yy}(0) \end{bmatrix}$$

Matriz das covariâncias

$$\begin{bmatrix}
 \lambda_{yy}(0) & \lambda_{yy}(1) & \lambda_{yy}(2) & \cdots & \lambda_{yy}(n-2) & \lambda_{yy}(n-1) \\
 \lambda_{yy}(1) & \lambda_{yy}(0) & \lambda_{yy}(1) & \cdots & \lambda_{yy}(n-3) & \lambda_{yy}(n-2) \\
 \lambda_{yy}(2) & \lambda_{yy}(1) & \lambda_{yy}(0) & \ddots & \lambda_{yy}(n-4) & \lambda_{yy}(n-3) \\
 \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
 \lambda_{yy}(n-2) & \lambda_{yy}(n-3) & \lambda_{yy}(n-4) & \ddots & \lambda_{yy}(0) & \lambda_{yy}(1) \\
 \lambda_{yy}(n-1) & \lambda_{yy}(n-2) & \lambda_{yy}(n-3) & \cdots & \lambda_{yy}(1) & \lambda_{yy}(0)
 \end{bmatrix}$$

- Todos os elementos da diagonal principal são iguais

Matriz das covariâncias

$$\begin{bmatrix} \lambda_{yy}(0) & \lambda_{yy}(1) & \lambda_{yy}(2) & \cdots & \lambda_{yy}(n-2) & \lambda_{yy}(n-1) \\ \lambda_{yy}(1) & \lambda_{yy}(0) & \lambda_{yy}(1) & \cdots & \lambda_{yy}(n-3) & \lambda_{yy}(n-2) \\ \lambda_{yy}(2) & \lambda_{yy}(1) & \lambda_{yy}(0) & \ddots & \lambda_{yy}(n-4) & \lambda_{yy}(n-3) \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \lambda_{yy}(n-2) & \lambda_{yy}(n-3) & \lambda_{yy}(n-4) & \ddots & \lambda_{yy}(0) & \lambda_{yy}(1) \\ \lambda_{yy}(n-1) & \lambda_{yy}(n-2) & \lambda_{yy}(n-3) & \cdots & \lambda_{yy}(1) & \lambda_{yy}(0) \end{bmatrix}$$

- Todos os elementos da diagonal principal são iguais
- Todos os elementos das sub-diagonais principais são iguais

Matriz das covariâncias

$$\begin{bmatrix} \lambda_{yy}(0) & \lambda_{yy}(1) & \lambda_{yy}(2) & \cdots & \lambda_{yy}(n-2) & \lambda_{yy}(n-1) \\ \lambda_{yy}(1) & \lambda_{yy}(0) & \lambda_{yy}(1) & \cdots & \lambda_{yy}(n-3) & \lambda_{yy}(n-2) \\ \lambda_{yy}(2) & \lambda_{yy}(1) & \lambda_{yy}(0) & \ddots & \lambda_{yy}(n-4) & \lambda_{yy}(n-3) \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \lambda_{yy}(n-2) & \lambda_{yy}(n-3) & \lambda_{yy}(n-4) & \ddots & \lambda_{yy}(0) & \lambda_{yy}(1) \\ \lambda_{yy}(n-1) & \lambda_{yy}(n-2) & \lambda_{yy}(n-3) & \cdots & \lambda_{yy}(1) & \lambda_{yy}(0) \end{bmatrix}$$

- Todos os elementos da diagonal principal são iguais
- Todos os elementos das sub-diagonais principais são iguais
- Todos os elementos em qualquer diagonal paralela à diagonal principal são iguais

Matriz das covariâncias

$$\begin{bmatrix} \lambda_{yy}(0) & \lambda_{yy}(1) & \lambda_{yy}(2) & \cdots & \lambda_{yy}(n-2) & \lambda_{yy}(n-1) \\ \lambda_{yy}(1) & \lambda_{yy}(0) & \lambda_{yy}(1) & \cdots & \lambda_{yy}(n-3) & \lambda_{yy}(n-2) \\ \lambda_{yy}(2) & \lambda_{yy}(1) & \lambda_{yy}(0) & \ddots & \lambda_{yy}(n-4) & \lambda_{yy}(n-3) \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \lambda_{yy}(n-2) & \lambda_{yy}(n-3) & \lambda_{yy}(n-4) & \ddots & \lambda_{yy}(0) & \lambda_{yy}(1) \\ \lambda_{yy}(n-1) & \lambda_{yy}(n-2) & \lambda_{yy}(n-3) & \cdots & \lambda_{yy}(1) & \lambda_{yy}(0) \end{bmatrix}$$

- Todos os elementos da diagonal principal são iguais
- Todos os elementos das sub-diagonais principais são iguais
- Todos os elementos em qualquer diagonal paralela à diagonal principal são iguais

Matriz das covariâncias

$$\begin{bmatrix} \lambda_{yy}(0) & \lambda_{yy}(1) & \lambda_{yy}(2) & \cdots & \lambda_{yy}(n-2) & \lambda_{yy}(n-1) \\ \lambda_{yy}(1) & \lambda_{yy}(0) & \lambda_{yy}(1) & \cdots & \lambda_{yy}(n-3) & \lambda_{yy}(n-2) \\ \lambda_{yy}(2) & \lambda_{yy}(1) & \lambda_{yy}(0) & \ddots & \lambda_{yy}(n-4) & \lambda_{yy}(n-3) \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \lambda_{yy}(n-2) & \lambda_{yy}(n-3) & \lambda_{yy}(n-4) & \ddots & \lambda_{yy}(0) & \lambda_{yy}(1) \\ \lambda_{yy}(n-1) & \lambda_{yy}(n-2) & \lambda_{yy}(n-3) & \cdots & \lambda_{yy}(1) & \lambda_{yy}(0) \end{bmatrix}$$

- Todos os elementos da diagonal principal são iguais
- Todos os elementos das sub-diagonais principais são iguais
- Todos os elementos em qualquer diagonal paralela à diagonal principal são iguais

Λ_n é uma **matriz de Toeplitz**

Matriz das covariâncias

$$\begin{bmatrix} \lambda_{yy}(0) & \lambda_{yy}(1) & \lambda_{yy}(2) & \cdots & \lambda_{yy}(n-2) & \lambda_{yy}(n-1) \\ \lambda_{yy}(1) & \lambda_{yy}(0) & \lambda_{yy}(1) & \cdots & \lambda_{yy}(n-3) & \lambda_{yy}(n-2) \\ \lambda_{yy}(2) & \lambda_{yy}(1) & \lambda_{yy}(0) & \ddots & \lambda_{yy}(n-4) & \lambda_{yy}(n-3) \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \lambda_{yy}(n-2) & \lambda_{yy}(n-3) & \lambda_{yy}(n-4) & \ddots & \lambda_{yy}(0) & \lambda_{yy}(1) \\ \lambda_{yy}(n-1) & \lambda_{yy}(n-2) & \lambda_{yy}(n-3) & \cdots & \lambda_{yy}(1) & \lambda_{yy}(0) \end{bmatrix}$$

Λ_n é uma **matriz simétrica**

Algoritmo de Levinson-Durbin

- Explora a estrutura da matriz das covariâncias (matriz de Toeplitz simétrica).
- Complexidade de tempo: n^2 (mais eficiente que Gauss-Jordan: n^3).
- Usado para calcular recursivamente os parâmetros de modelos auto-regressivos (AR) de ordem $k = 1$ a n .

Representação Recursiva dos Polinómios AR

- Polinómios auto-regressivos:

$$A^{(k)}(q^{-1}) = 1 + a_1^{(k)}q^{-1} + \dots + a_k^{(k)}q^{-k}$$

- A recursão calcula

$$\boldsymbol{\theta}^{(k+1)} = \begin{bmatrix} a_1^{(k+1)} & \dots & a_k^{(k+1)} & a_{k+1}^{(k+1)} \end{bmatrix}^T$$

a partir de

$$\boldsymbol{\theta}^{(k)} = \begin{bmatrix} a_1^{(k)} & \dots & a_k^{(k)} \end{bmatrix}^T$$

.

Recursão Inicial

- Para $k = 1$, o sistema reduz-se a uma única equação:

$$\lambda_{yy}(0)a_1^{(1)} = -\lambda_{yy}(1) \Rightarrow \boldsymbol{\theta}^{(1)} = a_1^{(1)} = -\frac{\lambda_{yy}(1)}{\lambda_{yy}(0)}$$

- Recursão para ordens superiores segue a equação

$$\boldsymbol{\Lambda}_{k+1}\boldsymbol{\theta}^{(k+1)} = -\boldsymbol{\lambda}_{1:k+1}.$$

Estrutura da Matriz de Toeplitz Expansível

$$\mathbf{\Lambda}_{k+1} = \begin{bmatrix} \lambda_{yy}(0) & \lambda_{yy}(1) & \cdots & \lambda_{yy}(k-1) & \lambda_{yy}(k) \\ \lambda_{yy}(1) & \lambda_{yy}(0) & \cdots & \lambda_{yy}(k-2) & \lambda_{yy}(k-1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \lambda_{yy}(k-1) & \lambda_{yy}(k-2) & \cdots & \lambda_{yy}(0) & \lambda_{yy}(1) \\ \lambda_{yy}(k) & \lambda_{yy}(k-1) & \cdots & \lambda_{yy}(1) & \lambda_{yy}(0) \end{bmatrix}$$

Estrutura da Matriz de Toeplitz Expansível

$$\mathbf{\Lambda}_{k+1} = \begin{bmatrix} \lambda_{yy}(0) & \lambda_{yy}(1) & \cdots & \lambda_{yy}(k-1) & \lambda_{yy}(k) \\ \lambda_{yy}(1) & \lambda_{yy}(0) & \cdots & \lambda_{yy}(k-2) & \lambda_{yy}(k-1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \lambda_{yy}(k-1) & \lambda_{yy}(k-2) & \cdots & \lambda_{yy}(0) & \lambda_{yy}(1) \\ \lambda_{yy}(k) & \lambda_{yy}(k-1) & \cdots & \lambda_{yy}(1) & \lambda_{yy}(0) \end{bmatrix}$$

Estrutura da Matriz de Toeplitz Expansível

$$\mathbf{\Lambda}_{k+1} = \left[\begin{array}{cccc|c} \lambda_{yy}(0) & \lambda_{yy}(1) & \cdots & \lambda_{yy}(k-1) & \lambda_{yy}(k) \\ \lambda_{yy}(1) & \lambda_{yy}(0) & \cdots & \lambda_{yy}(k-2) & \lambda_{yy}(k-1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \lambda_{yy}(k-1) & \lambda_{yy}(k-2) & \cdots & \lambda_{yy}(0) & \lambda_{yy}(1) \\ \hline \lambda_{yy}(k) & \lambda_{yy}(k-1) & \cdots & \lambda_{yy}(1) & \lambda_{yy}(0) \end{array} \right]$$

Estrutura da Matriz de Toeplitz Expansível

$$\mathbf{\Lambda}_{k+1} = \left[\begin{array}{cccc|c} & & & & \lambda_{yy}(k) \\ & & & & \lambda_{yy}(k-1) \\ & & \mathbf{\Lambda}_k & & \vdots \\ & & & & \lambda_{yy}(1) \\ \hline \lambda_{yy}(k) & \lambda_{yy}(k-1) & \cdots & \lambda_{yy}(1) & \lambda_{yy}(0) \end{array} \right]$$

$$\mathbf{\Lambda}_{k+1} = \left[\begin{array}{c|c} \mathbf{\Lambda}_k & \\ \hline & \end{array} \right].$$

Estrutura da Matriz de Toeplitz Expansível

$$\mathbf{\Lambda}_{k+1} = \left[\begin{array}{cccc|c} & & & & \lambda_{yy}(k) \\ & & & & \lambda_{yy}(k-1) \\ & & & & \vdots \\ & & & & \lambda_{yy}(1) \\ \hline \lambda_{yy}(k) & \lambda_{yy}(k-1) & \cdots & \lambda_{yy}(1) & \lambda_{yy}(0) \end{array} \right]$$

$$\mathbf{\Lambda}_{k+1} = \left[\begin{array}{c|c} \mathbf{\Lambda}_k & \boldsymbol{\lambda}_{k:1} \end{array} \right].$$

Estrutura da Matriz de Toeplitz Expansível

$$\mathbf{\Lambda}_{k+1} = \left[\begin{array}{cccc|c} & & & & \lambda_{yy}(k) \\ & & & & \lambda_{yy}(k-1) \\ & & & & \vdots \\ & & \mathbf{\Lambda}_k & & \lambda_{yy}(1) \\ \hline \lambda_{yy}(k) & \lambda_{yy}(k-1) & \cdots & \lambda_{yy}(1) & \lambda_{yy}(0) \end{array} \right]$$

$$\mathbf{\Lambda}_{k+1} = \left[\begin{array}{c|c} \mathbf{\Lambda}_k & \boldsymbol{\lambda}_{k:1} \\ \hline \boldsymbol{\lambda}_{k:1}^T & \end{array} \right].$$

Estrutura da Matriz de Toeplitz Expansível

$$\mathbf{\Lambda}_{k+1} = \left[\begin{array}{cccc|c} & & & & \lambda_{yy}(k) \\ & & & & \lambda_{yy}(k-1) \\ & & & & \vdots \\ & & \mathbf{\Lambda}_k & & \lambda_{yy}(1) \\ \hline \lambda_{yy}(k) & \lambda_{yy}(k-1) & \cdots & \lambda_{yy}(1) & \lambda_{yy}(0) \end{array} \right]$$

$$\mathbf{\Lambda}_{k+1} = \left[\begin{array}{c|c} \mathbf{\Lambda}_k & \boldsymbol{\lambda}_{k:1} \\ \hline \boldsymbol{\lambda}_{k:1}^T & \lambda_{yy}(0) \end{array} \right].$$

Cálculo dos Coeficientes Recursivos

- Atualização dos parâmetros recursivos:

$$\boldsymbol{\theta}^{(k+1)} = \begin{bmatrix} \boldsymbol{\theta}^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_k \\ a_{k+1}^{(k+1)} \end{bmatrix}.$$

Cálculo dos Coeficientes Recursivos

- Atualização dos parâmetros recursivos:

$$\boldsymbol{\theta}^{(k+1)} = \begin{bmatrix} \boldsymbol{\theta}^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_k \\ a_{k+1}^{(k+1)} \end{bmatrix}.$$

- equação de Yule-Walker para a ordem $k + 1$:

$$\left[\begin{array}{c|c} \boldsymbol{\Lambda}_k & \boldsymbol{\lambda}_{k:1} \\ \hline \boldsymbol{\lambda}_{k:1}^T & \lambda_{yy}(0) \end{array} \right] \underbrace{\left(\begin{bmatrix} \boldsymbol{\theta}^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_k \\ a_{k+1}^{(k+1)} \end{bmatrix} \right)}_{=\boldsymbol{\theta}^{(k+1)}} = - \begin{bmatrix} \boldsymbol{\lambda}_{1:k} \\ \lambda_{yy}(k+1) \end{bmatrix}.$$

Cálculo dos Coeficientes Recursivos

- Atualização dos parâmetros recursivos:

$$\boldsymbol{\theta}^{(k+1)} = \begin{bmatrix} \boldsymbol{\theta}^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_k \\ a_{k+1}^{(k+1)} \end{bmatrix}.$$

- equação de Yule-Walker para a ordem $k + 1$:

$$\begin{bmatrix} \Lambda_k & \lambda_{k:1} \\ \lambda_{k:1}^T & \lambda_{yy}(0) \end{bmatrix} \left(\begin{bmatrix} \boldsymbol{\theta}^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_k \\ a_{k+1}^{(k+1)} \end{bmatrix} \right) = - \begin{bmatrix} \lambda_{1:k} \\ \lambda_{yy}(k+1) \end{bmatrix}.$$

$$\Lambda_k \boldsymbol{\theta}^{(k)} + \Lambda_k \boldsymbol{\epsilon}_k + \lambda_{k:1} a_{k+1}^{(k+1)} = -\lambda_{1:k}$$

Cálculo dos Coeficientes Recursivos

- Atualização dos parâmetros recursivos:

$$\boldsymbol{\theta}^{(k+1)} = \begin{bmatrix} \boldsymbol{\theta}^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_k \\ a_{k+1}^{(k+1)} \end{bmatrix}.$$

- equação de Yule-Walker para a ordem $k + 1$:

$$\left[\begin{array}{c|c} \boldsymbol{\Lambda}_k & \boldsymbol{\lambda}_{k:1} \\ \hline \boldsymbol{\lambda}_{k:1}^T & \lambda_{yy}(0) \end{array} \right] \left(\begin{bmatrix} \boldsymbol{\theta}^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_k \\ a_{k+1}^{(k+1)} \end{bmatrix} \right) = - \begin{bmatrix} \boldsymbol{\lambda}_{1:k} \\ \lambda_{yy}(k+1) \end{bmatrix}.$$

$$\boldsymbol{\Lambda}_k \boldsymbol{\theta}^{(k)} + \boldsymbol{\Lambda}_k \boldsymbol{\epsilon}_k + \boldsymbol{\lambda}_{k:1} a_{k+1}^{(k+1)} = -\boldsymbol{\lambda}_{1:k}$$

Equação de Yule-Walker ordem k : $-\boldsymbol{\lambda}_{1:k} = \boldsymbol{\Lambda}_k \boldsymbol{\theta}^{(k)}$

Cálculo dos Coeficientes Recursivos

- Atualização dos parâmetros recursivos:

$$\boldsymbol{\theta}^{(k+1)} = \begin{bmatrix} \boldsymbol{\theta}^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_k \\ a_{k+1}^{(k+1)} \end{bmatrix}.$$

- equação de Yule-Walker para a ordem $k + 1$:

$$\left[\begin{array}{c|c} \boldsymbol{\Lambda}_k & \boldsymbol{\lambda}_{k:1} \\ \hline \boldsymbol{\lambda}_{k:1}^T & \lambda_{yy}(0) \end{array} \right] \left(\begin{bmatrix} \boldsymbol{\theta}^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_k \\ a_{k+1}^{(k+1)} \end{bmatrix} \right) = - \begin{bmatrix} \boldsymbol{\lambda}_{1:k} \\ \lambda_{yy}(k+1) \end{bmatrix}.$$

$$\boldsymbol{\Lambda}_k \boldsymbol{\theta}^{(k)} + \boldsymbol{\Lambda}_k \boldsymbol{\epsilon}_k + \boldsymbol{\lambda}_{k:1} a_{k+1}^{(k+1)} = -\boldsymbol{\lambda}_{1:k}$$

Equação de Yule-Walker ordem k : $-\boldsymbol{\lambda}_{1:k} = \boldsymbol{\Lambda}_k \boldsymbol{\theta}^{(k)}$

$$\boldsymbol{\Lambda}_k \boldsymbol{\theta}^{(k)} + \boldsymbol{\Lambda}_k \boldsymbol{\epsilon}_k + \boldsymbol{\lambda}_{k:1} a_{k+1}^{(k+1)} = \boldsymbol{\Lambda}_k \boldsymbol{\theta}^{(k)}$$

Cálculo dos Coeficientes Recursivos

- Atualização dos parâmetros recursivos:

$$\boldsymbol{\theta}^{(k+1)} = \begin{bmatrix} \boldsymbol{\theta}^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_k \\ a_{k+1}^{(k+1)} \end{bmatrix}.$$

- equação de Yule-Walker para a ordem $k + 1$:

$$\left[\begin{array}{c|c} \boldsymbol{\Lambda}_k & \boldsymbol{\lambda}_{k:1} \\ \hline \boldsymbol{\lambda}_{k:1}^T & \lambda_{yy}(0) \end{array} \right] \left(\begin{bmatrix} \boldsymbol{\theta}^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_k \\ a_{k+1}^{(k+1)} \end{bmatrix} \right) = - \begin{bmatrix} \boldsymbol{\lambda}_{1:k} \\ \lambda_{yy}(k+1) \end{bmatrix}.$$

$$\boldsymbol{\Lambda}_k \boldsymbol{\theta}^{(k)} + \boldsymbol{\Lambda}_k \boldsymbol{\epsilon}_k + \boldsymbol{\lambda}_{k:1} a_{k+1}^{(k+1)} = -\boldsymbol{\lambda}_{1:k}$$

Equação de Yule-Walker ordem k : $-\boldsymbol{\lambda}_{1:k} = \boldsymbol{\Lambda}_k \boldsymbol{\theta}^{(k)}$

$$\boldsymbol{\Lambda}_k \boldsymbol{\theta}^{(k)} + \boldsymbol{\Lambda}_k \boldsymbol{\epsilon}_k + \boldsymbol{\lambda}_{k:1} a_{k+1}^{(k+1)} = \boldsymbol{\Lambda}_k \boldsymbol{\theta}^{(k)}$$

Cálculo dos Coeficientes Recursivos

- Atualização dos parâmetros recursivos:

$$\boldsymbol{\theta}^{(k+1)} = \begin{bmatrix} \boldsymbol{\theta}^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_k \\ a_{k+1}^{(k+1)} \end{bmatrix}.$$

- equação de Yule-Walker para a ordem $k + 1$:

$$\left[\begin{array}{c|c} \boldsymbol{\Lambda}_k & \boldsymbol{\lambda}_{k:1} \\ \hline \boldsymbol{\lambda}_{k:1}^T & \lambda_{yy}(0) \end{array} \right] \left(\begin{bmatrix} \boldsymbol{\theta}^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_k \\ a_{k+1}^{(k+1)} \end{bmatrix} \right) = - \begin{bmatrix} \boldsymbol{\lambda}_{1:k} \\ \lambda_{yy}(k+1) \end{bmatrix}.$$

$$\boldsymbol{\Lambda}_k \boldsymbol{\theta}^{(k)} + \boldsymbol{\Lambda}_k \boldsymbol{\epsilon}_k + \boldsymbol{\lambda}_{k:1} a_{k+1}^{(k+1)} = -\boldsymbol{\lambda}_{1:k}$$

Equação de Yule-Walker ordem k : $-\boldsymbol{\lambda}_{1:k} = \boldsymbol{\Lambda}_k \boldsymbol{\theta}^{(k)}$

$$\cancel{\boldsymbol{\Lambda}_k \boldsymbol{\theta}^{(k)}} + \boldsymbol{\Lambda}_k \boldsymbol{\epsilon}_k + \boldsymbol{\lambda}_{k:1} a_{k+1}^{(k+1)} = \cancel{\boldsymbol{\Lambda}_k \boldsymbol{\theta}^{(k)}}$$

Cálculo dos Coeficientes Recursivos

- Atualização dos parâmetros recursivos:

$$\boldsymbol{\theta}^{(k+1)} = \begin{bmatrix} \boldsymbol{\theta}^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_k \\ a_{k+1}^{(k+1)} \end{bmatrix}.$$

- equação de Yule-Walker para a ordem $k + 1$:

$$\left[\begin{array}{c|c} \boldsymbol{\Lambda}_k & \boldsymbol{\lambda}_{k:1} \\ \hline \boldsymbol{\lambda}_{k:1}^T & \lambda_{yy}(0) \end{array} \right] \left(\begin{bmatrix} \boldsymbol{\theta}^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_k \\ a_{k+1}^{(k+1)} \end{bmatrix} \right) = - \begin{bmatrix} \boldsymbol{\lambda}_{1:k} \\ \lambda_{yy}(k+1) \end{bmatrix}.$$

$$\boldsymbol{\Lambda}_k \boldsymbol{\theta}^{(k)} + \boldsymbol{\Lambda}_k \boldsymbol{\epsilon}_k + \boldsymbol{\lambda}_{k:1} a_{k+1}^{(k+1)} = -\boldsymbol{\lambda}_{1:k}$$

Equação de Yule-Walker ordem k : $-\boldsymbol{\lambda}_{1:k} = \boldsymbol{\Lambda}_k \boldsymbol{\theta}^{(k)}$

$$\boldsymbol{\Lambda}_k \boldsymbol{\epsilon}_k + \boldsymbol{\lambda}_{k:1} a_{k+1}^{(k+1)} = \mathbf{0}$$

Cálculo dos Coeficientes Recursivos

- Atualização dos parâmetros recursivos:

$$\boldsymbol{\theta}^{(k+1)} = \begin{bmatrix} \boldsymbol{\theta}^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_k \\ a_{k+1}^{(k+1)} \end{bmatrix}.$$

- equação de Yule-Walker para a ordem $k + 1$:

$$\left[\begin{array}{c|c} \boldsymbol{\Lambda}_k & \boldsymbol{\lambda}_{k:1} \\ \hline \boldsymbol{\lambda}_{k:1}^T & \lambda_{yy}(0) \end{array} \right] \left(\begin{bmatrix} \boldsymbol{\theta}^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_k \\ a_{k+1}^{(k+1)} \end{bmatrix} \right) = - \begin{bmatrix} \boldsymbol{\lambda}_{1:k} \\ \lambda_{yy}(k+1) \end{bmatrix}.$$

$$\boldsymbol{\Lambda}_k \boldsymbol{\theta}^{(k)} + \boldsymbol{\Lambda}_k \boldsymbol{\epsilon}_k + \boldsymbol{\lambda}_{k:1} a_{k+1}^{(k+1)} = -\boldsymbol{\lambda}_{1:k}$$

Equação de Yule-Walker ordem k : $-\boldsymbol{\lambda}_{1:k} = \boldsymbol{\Lambda}_k \boldsymbol{\theta}^{(k)}$

$$\boldsymbol{\Lambda}_k \boldsymbol{\epsilon}_k + \boldsymbol{\lambda}_{k:1} a_{k+1}^{(k+1)} = 0$$

$$\boldsymbol{\Lambda}_k \boldsymbol{\epsilon}_k = -\boldsymbol{\lambda}_{k:1} a_{k+1}^{(k+1)}$$

$\theta_R - \theta$ ao contrário

● Equações de Yule-Walker ao contrário

$$\left\{ \begin{array}{l} \lambda_{yy}(0)a_k^{(k)} + \lambda_{yy}(1)a_{k-1}^{(k)} + \cdots + \lambda_{yy}(k-1)a_1^{(k)} = -\lambda_{yy}(k) \\ \lambda_{yy}(1)a_k^{(k)} + \lambda_{yy}(0)a_{k-1}^{(k)} + \cdots + \lambda_{yy}(k-2)a_1^{(k)} = -\lambda_{yy}(k-1) \\ \vdots \\ \lambda_{yy}(k-2)a_k^{(k)} + \lambda_{yy}(k-3)a_{k-1}^{(k)} + \cdots + \lambda_{yy}(1)a_1^{(k)} = -\lambda_{yy}(2) \\ \lambda_{yy}(k-1)a_k^{(k)} + \lambda_{yy}(k-2)a_{k-1}^{(k)} + \cdots + \lambda_{yy}(0)a_1^{(k)} = -\lambda_{yy}(1) \end{array} \right.$$

$\theta_R - \theta$ ao contrário

• Equações de Yule-Walker ao contrário

$$\left\{ \begin{array}{l} \lambda_{yy}(0)a_k^{(k)} + \lambda_{yy}(1)a_{k-1}^{(k)} + \cdots + \lambda_{yy}(k-1)a_1^{(k)} = -\lambda_{yy}(k) \\ \lambda_{yy}(1)a_k^{(k)} + \lambda_{yy}(0)a_{k-1}^{(k)} + \cdots + \lambda_{yy}(k-2)a_1^{(k)} = -\lambda_{yy}(k-1) \\ \vdots \\ \lambda_{yy}(k-2)a_k^{(k)} + \lambda_{yy}(k-3)a_{k-1}^{(k)} + \cdots + \lambda_{yy}(1)a_1^{(k)} = -\lambda_{yy}(2) \\ \lambda_{yy}(k-1)a_k^{(k)} + \lambda_{yy}(k-2)a_{k-1}^{(k)} + \cdots + \lambda_{yy}(0)a_1^{(k)} = -\lambda_{yy}(1) \end{array} \right.$$

• Na forma matricial:

$$\underbrace{\begin{bmatrix} \lambda_{yy}(0) & \lambda_{yy}(1) & \cdots & \lambda_{yy}(k-1) \\ \lambda_{yy}(1) & \lambda_{yy}(0) & \cdots & \lambda_{yy}(k-2) \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{yy}(k-1) & \lambda_{yy}(k-2) & \cdots & \lambda_{yy}(0) \end{bmatrix}}_{=\Lambda_k} \underbrace{\begin{bmatrix} a_k^{(k)} \\ a_{k-1}^{(k)} \\ \vdots \\ a_1^{(k)} \end{bmatrix}}_{\theta_R^{(k)}} = - \underbrace{\begin{bmatrix} \lambda_{yy}(k) \\ \lambda_{yy}(k-1) \\ \vdots \\ \lambda_{yy}(1) \end{bmatrix}}_{\lambda_{k:1}}.$$

Cálculo do erro ϵ_k

- Equações de Yule-Walker ao contrário

$$\lambda_{k:1} = -\Lambda_k \theta_R^{(k)}$$

Cálculo do erro ϵ_k

- Equações de Yule-Walker ao contrário

$$\lambda_{k:1} = -\Lambda_k \theta_R^{(k)}$$

- Erro:

$$\Lambda_k \epsilon_k = -\lambda_{k:1} a_{k+1}^{(k+1)}$$

Cálculo do erro ϵ_k

- Equações de Yule-Walker ao contrário

$$\lambda_{k:1} = -\Lambda_k \theta_R^{(k)}$$

- Erro:

$$\Lambda_k \epsilon_k = -\lambda_{k:1} a_{k+1}^{(k+1)}$$

$$\Lambda_k \epsilon_k = \Lambda_k \theta_R^{(k)} a_{k+1}^{(k+1)}$$

Cálculo do erro ϵ_k

- Equações de Yule-Walker ao contrário

$$\lambda_{k:1} = -\Lambda_k \theta_R^{(k)}$$

- Erro:

$$\Lambda_k \epsilon_k = -\lambda_{k:1} a_{k+1}^{(k+1)}$$

$$\cancel{\Lambda_k} \epsilon_k = \cancel{\Lambda_k} \theta_R^{(k)} a_{k+1}^{(k+1)}$$

Cálculo do erro ϵ_k

- Equações de Yule-Walker ao contrário

$$\lambda_{k:1} = -\Lambda_k \theta_R^{(k)}$$

- Erro:

$$\Lambda_k \epsilon_k = -\lambda_{k:1} a_{k+1}^{(k+1)}$$

$$\cancel{\Lambda}_k \epsilon_k = \cancel{\Lambda}_k \theta_R^{(k)} a_{k+1}^{(k+1)}$$

$$\epsilon_k = \theta_R^{(k)} a_{k+1}^{(k+1)}$$

Cálculo do erro ϵ_k

- Equações de Yule-Walker ao contrário

$$\lambda_{k:1} = -\Lambda_k \theta_R^{(k)}$$

- Erro:

$$\Lambda_k \epsilon_k = -\lambda_{k:1} a_{k+1}^{(k+1)}$$

$$\cancel{\Lambda}_k \epsilon_k = \cancel{\Lambda}_k \theta_R^{(k)} a_{k+1}^{(k+1)}$$

$$\epsilon_k = \theta_R^{(k)} a_{k+1}^{(k+1)}$$

Falta calcular $a_{k+1}^{(k+1)}$.

Cálculo do erro $a_{k+1}^{(k+1)}$

Erro:

$$\epsilon_k = \theta_R^{(k)} a_{k+1}^{(k+1)}$$

Cálculo do erro $a_{k+1}^{(k+1)}$

Erro:

$$\epsilon_k = \theta_R^{(k)} a_{k+1}^{(k+1)}$$

Multiplicar à esquerda por $\lambda_{k:1}^T$:

$$\lambda_{k:1}^T \epsilon_k = \lambda_{k:1}^T \theta_R^{(k)} a_{k+1}^{(k+1)}$$

Cálculo do erro $a_{k+1}^{(k+1)}$

Erro:

$$\epsilon_k = \theta_R^{(k)} a_{k+1}^{(k+1)}$$

Multiplicar à esquerda por $\lambda_{k:1}^T$:

$$\lambda_{k:1}^T \epsilon_k = \lambda_{k:1}^T \theta_R^{(k)} a_{k+1}^{(k+1)}$$

$$\lambda_{k:1}^T \theta_R^{(k)} = \lambda_{1:k}^T \theta^{(k)}$$

Cálculo do erro $a_{k+1}^{(k+1)}$

Erro:

$$\epsilon_k = \theta_R^{(k)} a_{k+1}^{(k+1)}$$

Multiplicar à esquerda por $\lambda_{k:1}^T$:

$$\lambda_{k:1}^T \epsilon_k = \lambda_{k:1}^T \theta_R^{(k)} a_{k+1}^{(k+1)}$$

$$\lambda_{k:1}^T \theta_R^{(k)} = \lambda_{1:k}^T \theta^{(k)}$$

$$\lambda_{k:1}^T \epsilon_k = \lambda_{1:k}^T \theta^{(k)} a_{k+1}^{(k+1)}$$

Cálculo de $a_{k+1}^{(k+1)}$

- De novo, a equação de Yule-Walker para a ordem $k + 1$:

$$\left[\begin{array}{c|c} \Lambda_k & \lambda_{k:1} \\ \hline \lambda_{k:1}^T & \lambda_{yy}(0) \end{array} \right] \left(\begin{bmatrix} \theta^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \epsilon_k \\ a_{k+1}^{(k+1)} \end{bmatrix} \right) = - \begin{bmatrix} \lambda_{1:k} \\ \lambda_{yy}(k+1) \end{bmatrix}.$$

Cálculo de $a_{k+1}^{(k+1)}$

- De novo, a equação de Yule-Walker para a ordem $k + 1$:

$$\left[\begin{array}{c|c} \Lambda_k & \lambda_{k:1} \\ \hline \lambda_{k:1}^T & \lambda_{yy}(0) \end{array} \right] \left(\begin{bmatrix} \theta^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \epsilon_k \\ a_{k+1}^{(k+1)} \end{bmatrix} \right) = - \begin{bmatrix} \lambda_{1:k} \\ \lambda_{yy}(k+1) \end{bmatrix}.$$

$$\lambda_{k:1}^T \theta^{(k)} + \lambda_{k:1}^T \epsilon_k + \lambda_{yy}(0) a_{k+1}^{(k+1)} = -\lambda_{yy}(k+1)$$

Cálculo de $a_{k+1}^{(k+1)}$

- De novo, a equação de Yule-Walker para a ordem $k + 1$:

$$\left[\begin{array}{c|c} \Lambda_k & \lambda_{k:1} \\ \hline \lambda_{k:1}^T & \lambda_{yy}(0) \end{array} \right] \left(\begin{bmatrix} \boldsymbol{\theta}^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_k \\ a_{k+1}^{(k+1)} \end{bmatrix} \right) = - \begin{bmatrix} \lambda_{1:k} \\ \lambda_{yy}(k+1) \end{bmatrix}.$$

$$\lambda_{k:1}^T \boldsymbol{\theta}^{(k)} + \lambda_{k:1}^T \boldsymbol{\epsilon}_k + \lambda_{yy}(0) a_{k+1}^{(k+1)} = -\lambda_{yy}(k+1)$$

$$\lambda_{k:1}^T \boldsymbol{\epsilon}_k = \lambda_{1:k}^T \boldsymbol{\theta}^{(k)} a_{k+1}^{(k+1)}$$

Cálculo de $a_{k+1}^{(k+1)}$

- De novo, a equação de Yule-Walker para a ordem $k + 1$:

$$\left[\begin{array}{c|c} \Lambda_k & \lambda_{k:1} \\ \hline \lambda_{k:1}^T & \lambda_{yy}(0) \end{array} \right] \left(\begin{bmatrix} \theta^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \epsilon_k \\ a_{k+1}^{(k+1)} \end{bmatrix} \right) = - \begin{bmatrix} \lambda_{1:k} \\ \lambda_{yy}(k+1) \end{bmatrix}.$$

$$\lambda_{k:1}^T \theta^{(k)} + \lambda_{k:1}^T \epsilon_k + \lambda_{yy}(0) a_{k+1}^{(k+1)} = -\lambda_{yy}(k+1)$$

$$\lambda_{k:1}^T \epsilon_k = \lambda_{1:k}^T \theta^{(k)} a_{k+1}^{(k+1)}$$

$$\lambda_{k:1}^T \theta^{(k)} + \lambda_{1:k}^T \theta^{(k)} a_{k+1}^{(k+1)} + \lambda_{yy}(0) a_{k+1}^{(k+1)} = -\lambda_{yy}(k+1)$$

Cálculo de $a_{k+1}^{(k+1)}$

- De novo, a equação de Yule-Walker para a ordem $k + 1$:

$$\left[\begin{array}{c|c} \Lambda_k & \lambda_{k:1} \\ \hline \lambda_{k:1}^T & \lambda_{yy}(0) \end{array} \right] \left(\begin{bmatrix} \theta^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \epsilon_k \\ a_{k+1}^{(k+1)} \end{bmatrix} \right) = - \begin{bmatrix} \lambda_{1:k} \\ \lambda_{yy}(k+1) \end{bmatrix}.$$

$$\lambda_{k:1}^T \theta^{(k)} + \lambda_{k:1}^T \epsilon_k + \lambda_{yy}(0) a_{k+1}^{(k+1)} = -\lambda_{yy}(k+1)$$

$$\lambda_{k:1}^T \epsilon_k = \lambda_{1:k}^T \theta^{(k)} a_{k+1}^{(k+1)}$$

$$\lambda_{k:1}^T \theta^{(k)} + \lambda_{1:k}^T \theta^{(k)} a_{k+1}^{(k+1)} + \lambda_{yy}(0) a_{k+1}^{(k+1)} = -\lambda_{yy}(k+1)$$

Cálculo de $a_{k+1}^{(k+1)}$

- De novo, a equação de Yule-Walker para a ordem $k + 1$:

$$\left[\begin{array}{c|c} \Lambda_k & \lambda_{k:1} \\ \hline \lambda_{k:1}^T & \lambda_{yy}(0) \end{array} \right] \left(\begin{bmatrix} \theta^{(k)} \\ 0 \end{bmatrix} + \begin{bmatrix} \epsilon_k \\ a_{k+1}^{(k+1)} \end{bmatrix} \right) = - \begin{bmatrix} \lambda_{1:k} \\ \lambda_{yy}(k+1) \end{bmatrix}.$$

$$\lambda_{k:1}^T \theta^{(k)} + \lambda_{k:1}^T \epsilon_k + \lambda_{yy}(0) a_{k+1}^{(k+1)} = -\lambda_{yy}(k+1)$$

$$\lambda_{k:1}^T \epsilon_k = \lambda_{1:k}^T \theta^{(k)} a_{k+1}^{(k+1)}$$

$$\lambda_{k:1}^T \theta^{(k)} + \lambda_{1:k}^T \theta^{(k)} a_{k+1}^{(k+1)} + \lambda_{yy}(0) a_{k+1}^{(k+1)} = -\lambda_{yy}(k+1)$$

$$a_{k+1}^{(k+1)} = - \frac{\lambda_{yy}(k+1) + \lambda_{k:1}^T \theta^{(k)}}{\lambda_{yy}(0) + \lambda_{1:k}^T \theta^{(k)}}.$$

Algoritmo de Levinson-Durbin

Passo 1: Inicialização

$$\theta_1^{(1)} = a_1^{(1)} = -\frac{\lambda_{yy}(0)}{\lambda_{yy}(1)}$$

Passo 2: Recursão

Para $k = 1$ a n :

$$\bullet a_{k+1}^{(k+1)} = -\frac{\lambda_{yy}(k+1) + \lambda_{k:1}^T \theta^{(k)}}{\lambda_{yy}(0) + \lambda_{1:k}^T \theta^{(k)}}$$

$$\bullet \epsilon_k = \theta_R^{(k)} a_{k+1}^{(k+1)}$$

$$\bullet \begin{bmatrix} a_1^{(k+1)} \\ \vdots \\ a_k^{(k+1)} \end{bmatrix} = \begin{bmatrix} a_1^{(k)} \\ \vdots \\ a_k^{(k)} \end{bmatrix} + a_{k+1}^{(k+1)} \begin{bmatrix} a_k^{(k)} \\ \vdots \\ a_1^{(k)} \end{bmatrix}$$

Estimador do Erro de Previsão

O que são algoritmos de erro de previsão?

- Ferramentas para estimar modelos paramétricos de processos estocásticos ou de sistemas determinístico-estocásticos.
- Ajustam os parâmetros para minimizar a diferença entre os valores observados e previstos.
- Essenciais para melhorar a correspondência entre o modelo e os dados observados.

Propriedades dos Algoritmos de Erro de Previsão

Consistência das Estimativas

- As estimativas convergem para os valores verdadeiros à medida que o número de observações aumenta.
- A precisão influencia diretamente a qualidade das previsões e a compreensão do processo.

Optimalidade em Processos Gaussianos

- Minimiza o erro médio quadrático dos parâmetros em processos gaussianos.
- Ótimo, na medida em que não existe nenhum outro estimador com menor erro médio quadrático.

Previsor a Um Passo para um Processo AR

Previsor a um passo:

$$\hat{y}(t, a_1, \dots, a_n) = [1 - A(q^{-1})] y(t) = -a_1 y(t-1) - \dots - a_n y(t-n)$$

Erro de Previsão:

$$\hat{e}(t, a_1, \dots, a_n) = y(t) - \hat{y}(t, a_1, \dots, a_n)$$

Critério Quadrático e Mínimos Quadrados

O critério quadrático é uma função quadrática dos parâmetros:

$$V(a_1, \dots, a_n) = \frac{1}{2} \sum_{t=1}^N \hat{e}^2(t, a_1, \dots, a_n)$$

Pode ser minimizado usando o algoritmo de mínimos quadrados:

$$\Phi^T \Phi \theta = \Phi^T Y$$

$$\Phi = \begin{bmatrix} -y(n) & -y(n-1) & \cdots & -y(1) \\ -y(n+1) & -y(n) & \cdots & -y(2) \\ \vdots & \vdots & \vdots & \vdots \\ -y(t-1) & -y(t-2) & \cdots & -y(t-n) \\ \vdots & \vdots & \vdots & \vdots \\ -y(N-1) & -y(N-2) & \cdots & -y(N-n) \end{bmatrix} \in \mathbb{R}^{(N-n) \times n}$$

$$Y = [y(n+1) \quad y(n+2) \quad \cdots \quad y(N)] \in \mathbb{R}^{N-n}.$$

Sistema de Equações e Solução

Sistema de equações normais:

$$\Phi^T \Phi \theta = \Phi^T Y$$

Solução (quando $\Phi^T \Phi$ é não singular):

$$\hat{\theta} = (\Phi^T \Phi)^{-1} \Phi^T Y$$

Relação com o Estimador de Yule-Walker

Podemos rescrever a solução de mínimos quadrados na forma:

$$\hat{\theta} = \left(\frac{1}{N-n} \Phi^T \Phi \right)^{-1} \frac{1}{N-n} \Phi^T Y.$$

$$\frac{1}{N-n} \Phi^T \Phi = \frac{1}{N-n} \begin{bmatrix} \sum_{t=n}^{N-1} y^2(t) & \cdots & \sum_{t=n}^{N-1} y(t)y(t-n+1) \\ \sum_{t=n-1}^{N-2} y(t)y(t+1) & \cdots & \sum_{t=n-1}^{N-2} y(t)y(t-n+2) \\ \vdots & \vdots & \vdots \\ \sum_{t=1}^{N-n} y(t)y(t+n-1) & \cdots & \sum_{t=1}^{N-n} y^2(t) \end{bmatrix} =$$

$$= \begin{bmatrix} \hat{\lambda}_{yy}(0) & \hat{\lambda}_{yy}(1) & \cdots & \hat{\lambda}_{yy}(n-1) \\ \hat{\lambda}_{yy}(1) & \hat{\lambda}_{yy}(0) & \cdots & \hat{\lambda}_{yy}(n-2) \\ \vdots & \vdots & \vdots & \vdots \\ \hat{\lambda}_{yy}(n-1) & \hat{\lambda}_{yy}(n-2) & \cdots & \hat{\lambda}_{yy}(0) \end{bmatrix} = \hat{\Lambda}_n$$

Relação com o Estimador de Yule-Walker

$$\frac{1}{N-n} \Phi^T Y = -\frac{1}{N-n} \begin{bmatrix} \sum_{t=n}^{N-1} y(t)y(t+1) \\ \sum_{t=n-1}^{N-2} y(t)y(t+2) \\ \vdots \\ \sum_{t=1}^{N-n} y(t)y(t+n) \end{bmatrix} = - \begin{bmatrix} \hat{\lambda}_{yy}(1) \\ \hat{\lambda}_{yy}(2) \\ \vdots \\ \hat{\lambda}_{yy}(n) \end{bmatrix} = \hat{\lambda}_{1:n}$$

Equivalência assintótica:

$$\hat{\theta} = \hat{\Lambda}_n^{-1} \hat{\lambda}_{1:n} \xrightarrow{\text{probabilidade 1}} \Lambda_n^{-1} \lambda_{1:n} = \theta$$

Determinação da Ordem do Modelo: Correlação Parcial

Modelação de um Processo AR:

- Um processo AR $y(t)$ é modelado como uma combinação linear de cópias atrasadas de $y(t)$.
- A ordem do modelo indica quais atrasos influenciam $y(t)$.
- A função de autocorrelação parcial avalia essa influência.

Autocorrelação Parcial: Definição

Autocorrelação Parcial:

- Mede a correlação entre $y(t)$ e $y(t - k)$, removendo a influência de atrasos intermédios.
- Exemplo para $k = 1$:

$$\phi_{yy}(1) = \rho_{yy}(1) = \frac{E \{y(t)y(t-1)\}}{\sqrt{E \{y^2(t)\} E \{y^2(t-1)\}}} = \frac{\Lambda_{yy}(1)}{\lambda_{yy}(0)}$$

Correlação Parcial: Exemplo com $k = 2$

- Decomposição de $y(t-2)$ em componentes correlacionadas e não correlacionadas com $y(t-1)$:

$$y(t) = y(t-1)\hat{\alpha}_1^{(1)} + e_0^{(1)}(t)$$

$$\mathbf{E} \left\{ e_0^{(1)}(t)y(t-1) \right\} = 0$$

- Decomposição de $y(t-2)$ em componentes correlacionadas e não correlacionadas com $y(t-1)$:

$$y(t-2) = y(t-1)\hat{\beta}_1^{(1)} + e_2^{(1)}(t)$$

$$\mathbf{E} \left\{ e_2^{(1)}(t)y(t-1) \right\} = 0$$

- $e_0^{(1)}(t)$ e $e_2^{(1)}(t)$ não correlacionadas com $y(t-1)$.

Correlação Parcial: Exemplo com $k = 2$

Correlação parcial entre $y(t)$ e $y(t - 2)$:

$$\phi_{yy}(2) = \frac{\mathbf{E} \left\{ e_0^{(1)}(t) e_2^{(1)}(t) \right\}}{\sqrt{\mathbf{E} \left\{ e_0^{(1)2}(t) \right\} \mathbf{E} \left\{ e_2^{(1)2}(t) \right\}}}$$

Modelo para Atrasos de Ordem Geral

Expressões gerais para $y(t)$ e $y(t - k)$:

$$y(t) = [y(t-1) \quad \cdots \quad y(t-k+1)] \hat{\alpha}^{(k-1)} + e_0^{(k-1)}(t)$$

$$y(t-k) = [y(t-1) \quad \cdots \quad y(t-k+1)] \hat{\beta}^{(k-1)} + e_k^{(k-1)}(t)$$

- $e_0^{(k-1)}(t)$ e $e_k^{(k-1)}(t)$ correlacionadas com

$$[y(t-1) \quad \cdots \quad y(t-k+1)]$$

.

Estimativa de Coeficientes de Correlação Parcial

Correlação Parcial para $y(t)$ e $y(t - k)$:

$$\phi_{yy}(k) = \frac{E \left\{ e_0^{(k-1)}(t) e_k^{(k-1)}(t) \right\}}{\sqrt{E \left\{ e_0^{(k-1)2}(t) \right\} E \left\{ e_k^{(k-1)2}(t) \right\}}}$$

- As estimativas de $e_0^{(k-1)}(t)$ e $e_k^{(k-1)}(t)$ são os resíduos do estimadores de mínimos quadrados.
- São necessárias $2n$ regressões lineares para estimar n coeficientes de correlação parcial.

Estimativa de Coeficientes de Correlação Parcial

- As regressões são da forma:

$$\hat{\alpha}^{(k-1)} = \left(\frac{1}{N-k-1} X_f^{(k-1)T} X_f^{(k-1)} \right)^{-1} \frac{1}{N-k-1} X_f^{(k-1)T} Y_f^{(k-1)}$$

e

$$\hat{\beta}^{(k-1)} = \left(\frac{1}{N-k-1} X_b^{(k-1)T} X_b^{(k-1)} \right)^{-1} \frac{1}{N-k-1} X_b^{(k-1)T} Y_b^{(k-1)}$$

Estimativa de Coeficientes de Correlação Parcial

$$X_f^{(k-1)} = \begin{bmatrix} y(k-1) & y(k-2) & \cdots & y(1) \\ y(k) & y(k-1) & \cdots & y(2) \\ \vdots & \vdots & \vdots & \vdots \\ y(N) & y(N-1) & \cdots & y(N-k+1) \end{bmatrix}$$

$$Y_f^{(k-1)} = [y(k) \quad y(k+1) \quad \cdots \quad y(N)]$$

$$X_b^{(k-1)} = \begin{bmatrix} y(N) & y(N-1) & \cdots & y(N-k+2) \\ y(N-1) & y(N-2) & \cdots & y(N-k+1) \\ \vdots & \vdots & \vdots & \vdots \\ y(k) & y(k-1) & \cdots & y(2) \end{bmatrix}$$

$$Y_b^{(k-1)} = [y(N-k+1) \quad y(N-k) \quad \cdots \quad y(1)]^T.$$

Estimativa de Coeficientes de Correlação Parcial

$$\lim_{N \rightarrow \infty} \hat{\alpha}^{(k-1)} = \Lambda_{k-1}^{-1} \lambda_{1:k-1} = \alpha^{(k-1)}$$

$$\lim_{N \rightarrow \infty} \hat{\alpha}^{(k-1)} = \Lambda_{k-1}^{-1} \lambda_{k-1:1} = \alpha_R^{(k-1)}$$

Relação com o Algoritmo de Levinson-Durbin

Lema (Coeficiente de Autocorrelação Parcial)

Seja $y(t)$ um processo estocástico estacionário de média nula, $\Phi_{yy}(k)$ o Coeficiente de Autocorrelação Parcial entre $y(t)$ e $y(t-k)$ e $\alpha^{(k)} = [\alpha_1^{(k)} \ \dots \ \alpha_k^{(k)}]^T$ o vector de parâmetros do AR(k). Então

$$\Phi_{yy}(k) = -\alpha_k^{(k)}.$$

Demonstração

● Modelo AR(k-1):

$$\begin{aligned} y(t) &= \alpha_1^{(k-1)} q^{-1} y(t) + \dots + \alpha_{k-1}^{(k-1)} y(t) + e_0^{(k-1)}(t) \Leftrightarrow \\ \Leftrightarrow \left(1 - \alpha_1^{(k-1)} q^{-1} - \dots - \alpha_{k-1}^{(k-1)} q^{-k+1}\right) y(t) &= e_0^{(k-1)}(t) \Leftrightarrow \\ \Leftrightarrow \left(1 + a_1^{(k-1)} q^{-1} + \dots + a_{k-1}^{(k-1)} q^{-k+1}\right) y(t) &= e_0^{(k-1)}(t) \end{aligned}$$

Relação com o Algoritmo de Levinson-Durbin

Demonstração (continuação)

$$\alpha_i^{k-1} = -a_i^{k-1}, \quad i = 1, \dots, k-1$$

Isto implica que:

$$\theta^{(k)} = -\alpha^{(k)}$$

Os coeficientes $\alpha_k^{(k)}$ podem ser calculados pela fórmula de Levinson-Durbin:

$$\alpha_k^{(k)} = -a_k^{(k)} = \frac{\lambda_{yy}(k) - \lambda_{k-1:1}^T \alpha^{(k-1)}}{\lambda_{yy}(0) - \lambda_{1:k-1}^T \alpha^{(k-1)}}$$

Relação com o Algoritmo de Levinson-Durbin

Demonstração (continuação)

- Componentes de $y(t)$ e $y(t - k)$ não correlacionadas com atrasos intermédios:

$$e_0^{(k-1)}(t) = y(t) - \varphi^{(k-1)}(t)\alpha^{(k-1)}$$

$$e_k^{(k-1)}(t) = y(t - k) - \varphi^{(k-1)}(t)\alpha_R^{(k-1)}$$

com $\varphi^{(k-1)}(t) = [y(t - 1) \ \dots \ y(t - k + 1)]$

- Coeficiente de correlação parcial entre $y(t)$ e $y(t - k)$ (coeficiente de correlação entre $e_0^{(k-1)}(t)$ e $e_k^{(k-1)}(t)$):

$$\phi_{yy}(k) = \frac{\mathbf{E} \left\{ e_0^{(k-1)}(t) e_k^{(k-1)}(t) \right\}}{\sqrt{\mathbf{E} \left\{ e_0^{(k-1)}(t)^2 \right\} \mathbf{E} \left\{ e_k^{(k-1)}(t)^2 \right\}}}$$

Relação com o Algoritmo de Levinson-Durbin

Demonstração (continuação)

- Cálculo de $E \left\{ e_0^{(k-1)}(t) e_k^{(k-1)}(t) \right\}$:

$$\begin{aligned} E \left\{ e_0^{(k-1)}(t) e_k^{(k-1)}(t) \right\} &= E \left\{ e_0^{(k-1)}(t) \left(y(t-k) - \varphi^{(k-1)}(t) \boldsymbol{\alpha}_R^{(k-1)} \right) \right\} = \\ &= E \left\{ e_0^{(k-1)}(t) y(t-k) \right\} - E \left\{ e_0^{(k-1)}(t) \varphi^{(k-1)}(t) \right\} \boldsymbol{\alpha}_R^{(k-1)} \end{aligned}$$

$e_0^{(k-1)}(t)$ não é correlacionado com $y(t-1), \dots, y(t-k+1)$:

$$E \left\{ e_0^{(k-1)}(t) \varphi^{(k-1)}(t) \right\} = 0$$

$$\begin{aligned} E \left\{ e_0^{(k-1)}(t) e_k^{(k-1)}(t) \right\} &= E \left\{ e_0^{(k-1)}(t) y(t-k) \right\} = \\ &= E \left\{ \left(y(t) - \varphi^{(k-1)}(t) \boldsymbol{\alpha}^{(k-1)} \right) y(t-k) \right\} \end{aligned}$$

Relação com o Algoritmo de Levinson-Durbin

Demonstração (continuação)

$$E \left\{ e_0^{(k-1)}(t) e_k^{(k-1)}(t) \right\} = E \{ y(t) y(t-k) \} - E \left\{ \varphi^{(k-1)}(t) \alpha^{(k-1)} y(t-k) \right\}$$

$$\bullet E \{ y(t) y(t-k) \} = \lambda_{yy}(k)$$

$$\begin{aligned} \bullet E \left\{ \varphi^{(k-1)}(t) \alpha^{(k-1)} y(t-k) \right\} &= \\ &= \left[E \{ y(t-1) y(t-k) \} \quad \cdots \quad E \{ y(t-k+1) y(t-k) \} \right] \alpha^{(k-1)} = \\ &= \left[\lambda_{yy}(k-1) \quad \cdots \quad \lambda_{yy}(1) \right] \alpha^{(k-1)} = \lambda_{k-1:1} \alpha^{(k-1)} \end{aligned}$$

$$E \left\{ e_0^{(k-1)}(t) e_k^{(k-1)}(t) \right\} = \lambda_{yy}(k) - \lambda_{k-1:1} \alpha^{(k-1)}$$

Relação com o Algoritmo de Levinson-Durbin

Demonstração (continuação)

- Cálculo de $\mathbf{E} \left\{ e_0^{(k-1)^2}(t) \right\}$:

$$\begin{aligned}
 \mathbf{E} \left\{ e_0^{(k-1)^2}(t) \right\} &= \mathbf{E} \left\{ e_0^{(k-1)}(t) \left(y(t) - \varphi^{(k-1)}(t) \boldsymbol{\alpha}^{(k-1)} \right) \right\} = \\
 &= \mathbf{E} \left\{ e_0^{(k-1)}(t) y(t) \right\} = \\
 &= \mathbf{E} \left\{ \left(y(t) - \varphi^{(k-1)}(t) \boldsymbol{\alpha}^{(k-1)} \right) y(t) \right\} = \\
 &= \mathbf{E} \left\{ y^2(t) \right\} - \mathbf{E} \left\{ \varphi^{(k-1)}(t) \boldsymbol{\alpha}^{(k-1)} y(t) \right\}
 \end{aligned}$$

- $\mathbf{E} \left\{ y^2(t) \right\} = \lambda_{yy}(0)$

- $\mathbf{E} \left\{ \varphi^{(k-1)}(t) \boldsymbol{\alpha}^{(k-1)} y(t) \right\} = \boldsymbol{\lambda}_{1:k} \boldsymbol{\alpha}^{(k-1)}$

$$\mathbf{E} \left\{ e_0^{(k-1)^2}(t) \right\} = \lambda_{yy}(0) - \boldsymbol{\lambda}_{1:k} \boldsymbol{\alpha}^{(k-1)}$$

Relação com o Algoritmo de Levinson-Durbin

Demonstração (continuação)

- Cálculo de $E \left\{ e_k^{(k-1)^2}(t) \right\}$:

$$\begin{aligned}
 E \left\{ e_k^{(k-1)^2}(t) \right\} &= E \left\{ e_k^{(k-1)}(t) \left(y(t-k) - \varphi^{(k-1)}(t) \alpha_R^{(k-1)} \right) \right\} = \\
 &= E \left\{ e_0^{(k-1)}(t) y(t-k) \right\} = \\
 &= E \left\{ \left(y(t-k) - \varphi^{(k-1)}(t) \alpha_R^{(k-1)} \right) y(t-k) \right\} = \\
 &= E \left\{ y^2(t-k) \right\} - E \left\{ \varphi^{(k-1)}(t) \alpha_R^{(k-1)} y(t-k) \right\}
 \end{aligned}$$

- $E \left\{ y^2(t-k) \right\} = \lambda_{yy}(0)$

- $E \left\{ \varphi^{(k-1)}(t) \alpha_R^{(k-1)} y(t-k) \right\} = \lambda_{k:1} \alpha_R^{(k-1)} = \lambda_{1:k} \alpha^{(k-1)}$

$$E \left\{ e_0^{(k-1)^2}(t) \right\} = \lambda_{yy}(0) - \lambda_{1:k} \alpha^{(k-1)}$$

Relação com o Algoritmo de Levinson-Durbin

Demonstração (continuação)

$$\begin{aligned}\phi_{yy}(k) &= \frac{\mathbf{E} \left\{ e_0^{(k-1)}(t) e_k^{(k-1)}(t) \right\}}{\sqrt{\mathbf{E} \left\{ e_0^{(k-1)2}(t) \right\} \mathbf{E} \left\{ e_k^{(k-1)2}(t) \right\}}} \\ &= \frac{\lambda_{yy}(k) - \boldsymbol{\lambda}_{k-1:1} \boldsymbol{\alpha}^{(k-1)}}{\lambda_{yy}(0) - \boldsymbol{\lambda}_{1:k} \boldsymbol{\alpha}^{(k-1)}} = \alpha_k^{(k)} = -a_k^{(k)}\end{aligned}$$

□

Conclusão: Correlação Parcial e Ordem do Modelo

- A autocorrelação parcial é uma ferramenta fundamental para determinar a ordem de um modelo AR.
- O coeficiente de correlação parcial $\phi_{yy}(k)$ é igual a $-a_k^{(k)}$, simétrico do último parâmetro do modelo de regressão para a ordem k .
- Pode-se utilizar o algoritmo de Levinson-Durbin para calcular os coeficientes de correlação parcial.

Estimação Paramétrica - Processos MA

- Processos MA são descritos pela equação:

$$y(t) = C(q^{-1}) e(t),$$

- $C(q^{-1}) = 1 + c_1 q^{-1} + \dots + c_n q^{-n}$
- $e(t)$: Ruído branco com média nula e variância σ^2 .

Densidade Espectral de Potência dum Processo MA

- Os processos MA são estacionários
- A densidade espectral de potência é dada por:

$$\Phi_{yy}(\omega) = |C(e^{-j\omega})|^2 \sigma^2 = |1 + c_1 e^{-j\omega} + \dots + c_n e^{-j\omega n}|^2 \sigma^2$$

- Relacionada diretamente com o modelo
- Estimar $C(q^{-1})$ permite obter a densidade espectral do processo

Método da Função Geradora da Autocovariância

Exemplo: Processo MA(1)

- Modelo:

$$y(t) = e(t) + c_1 e(t-1)$$

- Parâmetros: c_1 e σ^2
- Equações de Yule-Walker para a covariância:

$$\lambda_{yy}(0) = (1 + c_1^2) \sigma^2$$

$$\lambda_{yy}(1) = c_1 \sigma^2 = \lambda_{yy}(-1)$$

$$\lambda_{yy}(\tau) = 0, \quad |\tau| > 1$$

Função Geradora de Autocovariância

Exemplo: Processo MA(1) (continuação)

- Função de transferência:

$$H(z) = 1 + c_1 z^{-1}$$

- Densidade espectral de $y(t)$:

$$\Phi_{yy}(e^{j\omega}) = H(e^{j\omega}) H(e^{-j\omega}) \sigma^2$$

- Função Geradora de Autocovariância (transformada z da autocovariância):

$$\Phi_{yy}(z) = H(z)H(z^{-1}) = \lambda_{yy}(-1)z + \lambda_{yy}(0) + \lambda_{yy}(1)z^{-1}$$

Zeros da Função Geradora de Autocovariância

Exemplo: Processo MA(1) (continuação)

- zeros da função geradora de autocovariância: z_1 e $\frac{1}{z_1}$ com:

$$z_1 = \begin{cases} -\frac{\lambda_{yy}(0)}{2\lambda_{yy}(1)} + \frac{\sqrt{\left(\frac{\lambda_{yy}(0)}{\lambda_{yy}(1)}\right)^2 - 4}}{2}, & \lambda_{yy}(1) > 0 \\ -\frac{\lambda_{yy}(0)}{2\lambda_{yy}(1)} - \frac{\sqrt{\left(\frac{\lambda_{yy}(0)}{2\lambda_{yy}(1)}\right)^2 - 4}}{2}, & \lambda_{yy}(1) < 0 \end{cases}$$

Zeros da Função Geradora de Autocovariância

Exemplo: Processo MA(1) (continuação)

- Coeficiente de autocorrelação $\rho_{yy}(1)$:

$$\rho_{yy}(1) = \frac{\lambda_{yy}(1)}{\lambda_{yy}(0)}$$

- z_1 em função de $r_{yy}(1)$:

$$z_1 = -\frac{1 - \sqrt{1 - 4\rho_{yy}^2(1)}}{2\rho_{yy}(1)}$$

Parâmetros do modelo

Exemplo: Processo MA(1) (continuação)

- Parâmetros do modelo:

$$c_1 = -z_1 = \frac{1 - \sqrt{1 - 4\rho_{yy}^2(1)}}{2\rho_{yy}(1)}$$
$$\sigma^2 = \frac{\lambda_{yy}(0)}{1 + c_1^2}$$

- Restrição:

$$1 - 4\rho_{yy}^2(1) \Rightarrow |\rho_{yy}(1)| < \frac{1}{2}$$

- Se $\rho_{yy}(1) = \pm \frac{1}{2}$, $z_1 = \pm 1$ (não é inversamente estável)
- Se $\rho_{yy}(1) > \frac{1}{2}$ c_1 é complexo...

Estimação dos parâmetros dum modelo MA(n)

Modelo:

$$y(t) = e(t) + c_1 e(t-1) + \cdots + c_n e(t-n)$$

- 1 Estimar a sequência de autocovariância $\tau = -n, \dots, n$:

$$\hat{\lambda}_{yy}(\tau) = \begin{cases} \frac{1}{N-\tau} \sum_{t=\tau+1}^N y(t)y(t-\tau), & 0 \leq \tau \leq n \\ \hat{\lambda}_{yy}(\tau) = \hat{\lambda}_{yy}(-\tau), & -n \leq \tau \leq -1 \\ 0, & |\tau| > n. \end{cases}$$

- 2 Calcular os zeros da função geradora de autocovariância:

$$\hat{\lambda}_{yy}(n)z^{2n} + \cdots + \hat{\lambda}_{yy}(1)z^{n+1} + \hat{\lambda}_{yy}(0)z^n + \hat{\lambda}_{yy}(1)z^{n-1} + \cdots + \hat{\lambda}_{yy}(n) = 0$$

- 3 Seleccionar os zeros $z_1, \dots, z_n$ que se encontram no interior do círculo unitário.

- 4 Fazer $\hat{H}(z) = (1 - z_1 z^{-1}) \cdots (1 - z_n^{-1} z_n)$

Estimação dos parâmetros dum modelo MA(n) (continuação)

5 Calcular

$$c_k = (-1)^k \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} z_{i_1} z_{i_2} \dots z_{i_k}, \quad k = 1, \dots, n$$

Ou seja, cada coeficiente c_k é a soma de todos os produtos dos zeros z_i tomados k a k , com sinal alternado determinado por $(-1)^k$.

6 Calcular:

$$\sigma^2 = \frac{\hat{\lambda}(0)}{1 + c_1^2 + \dots + c_n^2}$$

Inconveniente:

- Como é estimada a partir duma realização, a função geradora de autocovariância pode apresentar **um ou mais pares de polos na circunferência unitária** dando origem a modelos **com parâmetros complexos** e, por isso, **não admissíveis**.

Método da resposta impulsional truncada

- Função de transferência dum processo MA(n):

$$H(z) = 1 + c_1 z^{-1} + \dots + c_n z^{-n}$$

- 1 Aproximação de um processo MA(n) por um modelo AR(p):

- $$H(z) = \frac{1}{\left(\frac{1}{H(z)}\right)}$$

- Expansão de $\frac{1}{H(z)}$:

$$\frac{1}{H(z)} = 1 + a_1 z^{-1} + \dots + a_p z^{-p} + \dots$$

Método da resposta impulsional truncada

- Como fazer a expansão de $\frac{1}{H(z)}$?

1 Divisão longa: $1 \quad \underline{H(z)}$

2 Transformada inversa z :

$$a_k = \mathcal{Z}^{-1} \left\{ \frac{1}{H(z)} \right\}, \quad k = 1, \dots, p, \dots$$

- Truncar a expansão de $\frac{1}{H(z)}$ na potência $(-p)$:

$$\frac{1}{H(z)} \approx 1 + a_1 z^{-1} + \dots + a_p z^{-p} = A(z^{-1}) \Rightarrow H(z) \approx \frac{1}{A(z^{-1})}$$

- Aproximação AR(p):

$$y(t) + a_1 y(t-1) + \dots + a_p y(t-p) = e(t).$$

Método da resposta impulsional truncada

- Porque é que podemos truncar a expansão de $\frac{1}{H(z)}$?
- Resposta:
 - $H(z)$ é inversamente estável
 - Condição necessária e suficiente para a estabilidade de $\frac{1}{H(z)}$: $\sum_{k=1}^{\infty} |a_k| = C < \infty$
 - $\sum_{k=1}^{\infty} |a_k| = C < \infty \Rightarrow \lim_{p \rightarrow \infty} \sum_{k=p+1}^{\infty} |a_k| = 0$.

Método da resposta impulsional truncada

Exemplo: Aproximação dum modelo MA(1) por um AR(p)

- $H(z) = 1 + c_1 z^{-1}$

Método da resposta impulsional truncada

Exemplo: Aproximação dum modelo MA(1) por um AR(p)

- $H(z) = 1 + c_1 z^{-1}$

- Expansão de $\frac{1}{H(z)}$ (divisão de 1 por $H(z)$):

1

$$\begin{array}{r} 1 + c_1 z^{-1} \\ \hline \end{array}$$

Método da resposta impulsional truncada

Exemplo: Aproximação dum modelo MA(1) por um AR(p)

- $H(z) = 1 + c_1 z^{-1}$

- Expansão de $\frac{1}{H(z)}$ (divisão de 1 por $H(z)$):

1

$$\begin{array}{r} 1 + c_1 z^{-1} \\ \hline 1 \end{array}$$

Método da resposta impulsional truncada

Exemplo: Aproximação dum modelo MA(1) por um AR(p)

- $H(z) = 1 + c_1 z^{-1}$

- Expansão de $\frac{1}{H(z)}$ (divisão de 1 por $H(z)$):

$$\begin{array}{r|l} 1 & 1 + c_1 z^{-1} \\ 1 \quad c_1 z^{-1} & 1 \end{array}$$

Método da resposta impulsional truncada

Exemplo: Aproximação dum modelo MA(1) por um AR(p)

- $H(z) = 1 + c_1 z^{-1}$

- Expansão de $\frac{1}{H(z)}$ (divisão de 1 por $H(z)$):

$$\begin{array}{r|l} 1 & 1 + c_1 z^{-1} \\ 1 & c_1 z^{-1} \\ \hline 0 & -c_1 z^{-1} \end{array}$$

Método da resposta impulsional truncada

Exemplo: Aproximação dum modelo MA(1) por um AR(p)

- $H(z) = 1 + c_1 z^{-1}$

- Expansão de $\frac{1}{H(z)}$ (divisão de 1 por $H(z)$):

$$\begin{array}{r|l} 1 & 1 + c_1 z^{-1} \\ 1 \quad c_1 z^{-1} & 1 - c_1 z^{-1} \\ \hline 0 \quad -c_1 z^{-1} & \end{array}$$

Método da resposta impulsional truncada

Exemplo: Aproximação dum modelo MA(1) por um AR(p)

- $H(z) = 1 + c_1 z^{-1}$

- Expansão de $\frac{1}{H(z)}$ (divisão de 1 por $H(z)$):

$$\begin{array}{r|l}
 1 & 1 + c_1 z^{-1} \\
 1 \quad c_1 z^{-1} & \hline
 0 \quad -c_1 z^{-1} & 1 - c_1 z^{-1} \\
 \hline
 & -c_1 z^{-1} \quad -c_1^2 z^{-2}
 \end{array}$$

Método da resposta impulsional truncada

Exemplo: Aproximação dum modelo MA(1) por um AR(p)

- $H(z) = 1 + c_1 z^{-1}$

- Expansão de $\frac{1}{H(z)}$ (divisão de 1 por $H(z)$):

$$\begin{array}{r|l}
 1 & 1 + c_1 z^{-1} \\
 1 \quad c_1 z^{-1} & \hline
 0 \quad -c_1 z^{-1} & \\
 \quad -c_1 z^{-1} \quad -c_1^2 z^{-2} & \\
 \quad \quad 0 \quad c_1^2 z^{-2} &
 \end{array}$$

Método da resposta impulsional truncada

Exemplo: Aproximação dum modelo MA(1) por um AR(p)

- $H(z) = 1 + c_1 z^{-1}$

- Expansão de $\frac{1}{H(z)}$ (divisão de 1 por $H(z)$):

$$\begin{array}{r}
 1 \\
 1 \quad c_1 z^{-1} \\
 \hline
 0 \quad -c_1 z^{-1} \\
 \quad -c_1 z^{-1} \quad -c_1^2 z^{-2} \\
 \hline
 \quad \quad 0 \quad c_1^2 z^{-2} \\
 \quad \quad c_1^2 z^{-2} \quad c_1^3 z^{-3} \\
 \hline
 \quad \quad \quad 0 \quad -c_1^3 z^{-3} \\
 \quad \quad \quad \quad \vdots
 \end{array}
 \quad \bigg| \quad
 \begin{array}{l}
 1 + c_1 z^{-1} \\
 \hline
 1 - c_1 z^{-1} + c_1^2 z^{-2} - c_1^3 z^{-3} + \dots
 \end{array}$$

Método da resposta impulsional truncada

Exemplo: Aproximação dum modelo MA(1) por um AR(p)

- $H(z) = 1 + c_1 z^{-1}$

- Expansão de $\frac{1}{H(z)}$ (divisão de 1 por $H(z)$):

$$\begin{array}{r|l}
 1 & 1 + c_1 z^{-1} \\
 1 & 1 - c_1 z^{-1} + c_1^2 z^{-2} - c_1^3 z^{-3} + \dots \\
 \hline
 0 & -c_1 z^{-1} \\
 & -c_1 z^{-1} \quad -c_1^2 z^{-2} \\
 \hline
 & 0 \quad c_1^2 z^{-2} \\
 & c_1^2 z^{-2} \quad c_1^3 z^{-3} \\
 \hline
 & 0 \quad -c_1^3 z^{-3} \\
 & \vdots
 \end{array}$$

- $\frac{1}{H(z)} = 1 - c_1 z^{-1} + c_1^2 z^{-2} - c_1^3 z^{-3} + \dots + (-1)^p c_1^p z^{-p} + \dots$

Método da resposta impulsional truncada

Exemplo: Aproximação dum modelo MA(1) por um AR(p)

- Truncar a soma para $k \geq p$:

$$A(z^{-1}) = 1 - c_1 z^{-1} + c_1^2 z^{-2} - c_1^3 z^{-3} + \dots + (-c_1)^p z^{-p}$$

- Aproximação AR(p)

$$A(q^{-1})y(t) = e(t),$$

ou

$$y(t) = e(t) + c_1 y(t-1) - c_1^2 y(t-2) + c_1^3 y(t-3) - \dots - (-c_1)^p y(t-p)$$

Método da resposta impulsional truncada

- Obtida a aproximação AR(p) determinar a resposta impulsional:

$$\hat{c}_k = \mathcal{Z}^{-1} \left\{ \frac{1}{A(z^{-1})} \right\}$$

ou expandir $\frac{1}{A(z^{-1})}$:

$$\frac{1}{A(z^{-1})} = 1 + \hat{c}_1 z^{-1} + \dots + \hat{c}_n z^{-n} + \dots$$

- Truncar a resposta impulsional da aproximação AR(p) para $k = n$:

$$\hat{H}(z) = 1 + \hat{c}_1 z^{-1} + \dots + \hat{c}_n z^{-n}$$

Método da resposta impulsional truncada

- Modelo estimado:

$$y(t) = e(t) + \hat{c}_1 e(t-1) + \cdots + \hat{c}_n e(t-n)$$

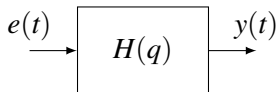
- Variância de $e(t)$:

$$\hat{\sigma}^2 = \frac{\hat{\lambda}_{yy}(0)}{1 + \hat{c}_1^2 + \cdots + \hat{c}_n^2}$$

- Se tivermos a (ou uma estimativa da) sequência de autocovariância do processo MA(n), podemos determinar a ordem da aproximação AR através dos coeficientes de correlação parcial e estimar os parâmetros do modelo AR através da resolução das equações de Yule-Walker

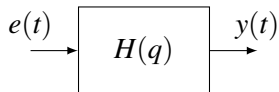
Método dos Mínimos Quadrados para Modelos MA

- Teorema da densidade espectral permite interpretar o processo estocástico estacionário $y(t)$ como a saída dum sistema linear, estável e invariante no tempo:



Método dos Mínimos Quadrados para Modelos MA

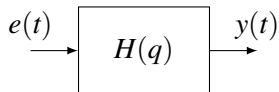
- Teorema da densidade espectral permite interpretar o processo estocástico estacionário $y(t)$ como a saída dum sistema linear, estável e invariante no tempo:



- Se $e(t)$ for conhecido, a estimação de $H(q)$ é um problema de identificação de sistemas.

Método dos Mínimos Quadrados para Modelos MA

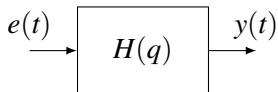
- Teorema da densidade espectral permite interpretar o processo estocástico estacionário $y(t)$ como a saída dum sistema linear, estável e invariante no tempo:



- Se $e(t)$ for conhecido, a estimação de $H(q)$ é um problema de identificação de sistemas.
- Pode ser resolvido utilizando o estimador de mínimos quadrados

Método dos Mínimos Quadrados para Modelos MA

- Teorema da densidade espectral permite interpretar o processo estocástico estacionário $y(t)$ como a saída dum sistema linear, estável e invariante no tempo:



- Se $e(t)$ for conhecido, a estimação de $H(q)$ é um problema de identificação de sistemas.
- Pode ser resolvido utilizando o estimador de mínimos quadrados

Como estimar $e(t)$?

Estimação de $e(t)$

- Aproximar o processo por um modelo $AR(p)$ de ordem p elevada:

$$y(t) = \hat{e}(t) - \hat{a}_1 y(t-1) - \dots - \hat{a}_p y(t-p)$$

Estimação de $e(t)$

- Aproximar o processo por um modelo $AR(p)$ de ordem p elevada:

$$y(t) = \hat{e}(t) - \hat{a}_1 y(t-1) - \dots - \hat{a}_p y(t-p)$$

- Determinação da ordem p :
 - 1 Análise dos coeficientes de correlação parcial
 - 2 Confirmar a ordem através dum teste de brancura de resíduos

Estimação de $e(t)$

- Aproximar o processo por um modelo AR(p) de ordem p elevada:

$$y(t) = \hat{e}(t) - \hat{a}_1 y(t-1) - \dots - \hat{a}_p y(t-p)$$

- Determinação da ordem p :
 - 1 Análise dos coeficientes de correlação parcial
 - 2 Confirmar a ordem através dum teste de brancura de resíduos
- Gerar $\hat{e}(t)$:

$$\hat{e}(t) = y(t) + \hat{a}_1 y(t-1) + \dots + \hat{a}_p y(t-p)$$

Identificação do modelo MA(n)

- Modelo MA(n)

$$y(t) = \hat{e}(t) + c_1\hat{e}(t-1) + \cdots + c_n\hat{e}(t-n)$$

Identificação do modelo MA(n)

- Modelo MA(n)

$$y(t) = \hat{e}(t) + c_1 \hat{e}(t-1) + \cdots + c_n \hat{e}(t-n)$$

- Vector de parâmetros:

$$\theta = [c_1 \quad \cdots \quad c_n]^T$$

Identificação do modelo MA(n)

- Modelo MA(n)

$$y(t) = \hat{e}(t) + c_1\hat{e}(t-1) + \cdots - c_n\hat{e}(t-n)$$

- Vector de parâmetros:

$$\theta = [c_1 \quad \cdots \quad c_n]^T$$

- Regressor:

$$\varphi(t) = [\hat{e}(t-1) \quad \cdots \quad \hat{e}(t-n)]^T$$

Identificação do modelo MA(n)

- Modelo MA(n)

$$y(t) = \hat{e}(t) + c_1\hat{e}(t-1) + \dots - c_n\hat{e}(t-n)$$

- Vector de parâmetros:

$$\theta = [c_1 \quad \dots \quad c_n]^T$$

- Regressor:

$$\varphi(t) = [\hat{e}(t-1) \quad \dots \quad \hat{e}(t-n)]^T$$

- Modelo de mínimos quadrados:

$$y(t) = \varphi^T(t)\theta + \hat{e}(t)$$

onde $\hat{e}(t)$ é ruído branco independente de $\varphi(t)$.

Identificação do modelo MA(n)

- Estimador de mínimos quadrados:

$$\hat{\theta} = (\Phi^T \Phi)^{-1} \Phi^T Y$$

com

$$\Phi = \begin{bmatrix} \varphi^T(n+1) \\ \vdots \\ \varphi(N) \end{bmatrix} = \begin{bmatrix} \hat{e}(n) & \cdots & \hat{e}(1) \\ \vdots & \vdots & \vdots \\ \hat{e}(N-1) & \cdots & \hat{e}(N-n) \end{bmatrix}$$

e

$$Y = [y(n+1) \quad \cdots \quad y(N)]^T$$

Método da correlação

- Estimador de mínimos quadrados:

$$\hat{\theta} = \left(\frac{1}{N-n-1} \Phi^T \Phi \right)^{-1} \frac{1}{N-n-1} \Phi^T Y$$

- $\frac{1}{N-n-1} \Phi^T \Phi =$

$$= \begin{bmatrix} \frac{1}{N-n-1} \sum_{t=n}^{N-1} \hat{e}^2(t) & \cdots & \frac{1}{N-n-1} \sum_{t=n}^{N-1} \hat{e}(t) \hat{e}(t-n+1) \\ \vdots & \vdots & \vdots \\ \frac{1}{N-n-1} \sum_{t=1}^{N-n} \hat{e}(t) \hat{e}(t+n-1) & \cdots & \frac{1}{N-n-1} \sum_{t=1}^{N-n} \hat{e}^2(t) \end{bmatrix}$$

Método da correlação

$$\bullet \lim_{n \rightarrow \infty} \frac{1}{N - n - 1} \Phi^T \Phi =$$

$$\begin{bmatrix} \lambda_{\hat{e}\hat{e}}(0) & \lambda_{\hat{e}\hat{e}}(1) & \cdots & \lambda_{\hat{e}\hat{e}}(n-1) \\ \lambda_{\hat{e}\hat{e}}(1) & \lambda_{\hat{e}\hat{e}}(0) & \cdots & \lambda_{\hat{e}\hat{e}}(n-2) \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{\hat{e}\hat{e}}(n-1) & \lambda_{\hat{e}\hat{e}}(n-2) & \cdots & \lambda_{\hat{e}\hat{e}}(0) \end{bmatrix}$$

- Como $\hat{e}(t)$ é ruído branco com variância $\hat{\sigma}^2(t)$:

$$\lambda_{\hat{e}\hat{e}}(\tau) = \begin{cases} \hat{\sigma}^2, & \tau = 0 \\ 0, & \tau \neq 0, \end{cases}$$

então

$$\lim_{n \rightarrow \infty} \frac{1}{N - n - 1} \Phi^T \Phi = \hat{\sigma}^2 \mathbf{I}_n.$$

Método da correlação

- $\Phi^T Y =$

$$\left[\frac{1}{N-n-1} \sum_{t=n}^{N-1} \hat{e}(t)y(t+1) \quad \cdots \quad \frac{1}{N-n-1} \sum_{t=1}^{N-n} \hat{e}(t)y(t+n) \right]^T$$

- $\lim_{N \rightarrow \infty} \frac{1}{N-n-1} \Phi^T \Phi = [\lambda_{y\hat{e}}(1) \quad \lambda_{y\hat{e}}(2) \quad \cdots \quad \lambda_{y\hat{e}}(n)]^T$

- $\lim_{N \rightarrow \infty} \hat{\theta} = \left[\frac{\lambda_{y\hat{e}}(1)}{\hat{\sigma}^2} \quad \frac{\lambda_{y\hat{e}}(2)}{\hat{\sigma}^2} \quad \cdots \quad \frac{\lambda_{y\hat{e}}(n)}{\hat{\sigma}^2} \right]^T$

- Temos, finalmente:

$$\hat{c}_k = \frac{\hat{\lambda}_{y\hat{e}}(k)}{\hat{\sigma}^2}, \quad k = 1, \dots, n.$$

Minimização do erro de previsão

- O método de minimização do erro de previsão minimiza o critério:

$$J(\theta) = \frac{1}{2} \sum_{t=1}^N \hat{e}^2(t, \theta)$$

- $\hat{e}(t, \theta)$ - erro de previsão
- $\theta = [c_1 \ \cdots \ c_n]^T$ - vector de parâmetros de $C(q^{-1})$.
- $\hat{\theta} = \min_{\theta} J(\theta)$ é o estimador do erro de previsão.
- $\hat{\theta} \xrightarrow{\text{probabilidade 1}} \theta_0$ - valor verdadeiro dos parâmetros
- $\hat{\theta}$ é um estimador fortemente consistente

Minimização do erro de previsão

- $J(\theta)$ é mínimo quando o seu gradiente é nulo:

$$\left(\frac{dJ(\theta)}{d\theta} \right)^T = \left[\frac{\partial J(\theta)}{\partial c_1} \quad \dots \quad \frac{\partial J(\theta)}{\partial c_n} \right]^T = \mathbf{0}.$$

- $\left(\frac{dJ(\theta)}{d\theta} \right)^T = \mathbf{0}$ não tem solução analítica.
- Método iterativo (descida de gradiente):

$$\theta^{(i+1)} = \theta^{(i)} - \eta \left(\frac{dJ(\theta)}{d\theta} \right)^T_{\theta=\theta^{(i)}}$$

- $\eta > 0$ - taxa de aprendizagem.

Cálculo do gradiente

- A derivada da função custo relativamente a c_k :

$$\frac{\partial J(\boldsymbol{\theta})}{\partial c_k} = \sum_{t=1}^N \frac{\partial \hat{e}(t, \boldsymbol{\theta})}{\partial c_k} \hat{e}(t, \boldsymbol{\theta}), \quad k = 1, \dots, n.$$

- $\hat{e}(t, \boldsymbol{\theta}) = y(t) - \hat{y}(t, \boldsymbol{\theta})$
- $\hat{y}(t, \boldsymbol{\theta})$ é a previsão de $y(t)$ a um passo:

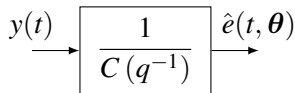
$$\hat{y}(t, \boldsymbol{\theta}) = \frac{C(q^{-1}) - 1}{C(q^{-1})} y(t).$$

- Equação às diferenças do erro de previsão:

$$\hat{e}(t, \boldsymbol{\theta}) = y(t) - \hat{y}(t, \boldsymbol{\theta}) = y(t) - c_1 \hat{e}(t-1, \boldsymbol{\theta}) - \dots - c_n \hat{e}(t-n, \boldsymbol{\theta})$$

Cálculo do gradiente

- $$\hat{e}(t, \theta) = \frac{1}{C(q^{-1})} y(t)$$

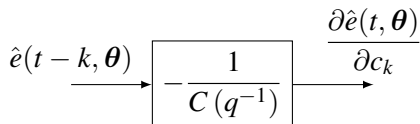


- $\hat{e}(t, \theta)$ é a saída dum sistema com função de transferência $\frac{1}{C(q^{-1})}$ excitado por $y(t)$
- Derivadas do erro em relação aos parâmetros c_k , $k = 1, \dots, n$:

$$\frac{\partial \hat{e}(t, \theta)}{\partial c_k} = -\hat{e}(t-k, \theta) - c_1 \frac{\partial \hat{e}(t-1, \theta)}{\partial c_k} - \dots - c_n \frac{\partial \hat{e}(t-n, \theta)}{\partial c_k},$$

Cálculo do gradiente

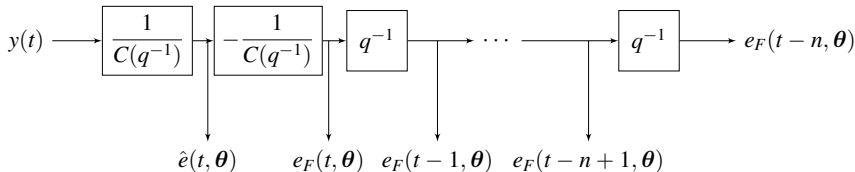
- $$\frac{\partial \hat{e}(t, \theta)}{\partial c_k} = -\frac{1}{C(q^{-1})} \hat{e}(t-k, \theta)$$



- $\frac{\partial \hat{e}(t, \theta)}{\partial c_k}$ é a saída dum sistema com função de transferência $-\frac{1}{C(q^{-1})}$ excitado por $\hat{e}(t, \theta)$

Cálculo do gradiente

- Definir: $e_F(t, \boldsymbol{\theta}) = -\frac{1}{C(q^{-1})}\hat{e}(t, \boldsymbol{\theta})$
- $\frac{\partial \hat{e}(t, \boldsymbol{\theta})}{\partial c_k} = e_F(t - k, \boldsymbol{\theta})$



- A derivada da função custo relativamente a c_k ($k^{\text{ésimo}}$ componente do gradiente):

$$\frac{\partial J(\boldsymbol{\theta})}{\partial c_k} = \sum_{t=1}^N e_F(t - k, \boldsymbol{\theta}) \hat{e}(t, \boldsymbol{\theta}), \quad k = 1, \dots, n.$$

Algoritmo

Minimização do erro de previsão:

Inicialização: Obter $\theta^{(0)}$ e definir $i = 0$.

Repetir até convergir:

- $\hat{e}(t, \theta^{(i)}) = \frac{1}{C^{(i)}(q^{-1})} y(t), \quad t = 1, \dots, N$
- $\hat{e}_F(t, \theta^{(i)}) = -\frac{1}{C^{(i)}(q^{-1})} \hat{e}(t, \theta^{(i)}), \quad t = 1, \dots, N$
- $\left(\frac{dJ(\theta)}{d\theta} \right)_{\theta=\theta^{(i)}}^T = \sum_{t=1}^N \begin{bmatrix} \hat{e}_F(t-1, \theta^{(i)}) \\ \vdots \\ \hat{e}_F(t-n, \theta^{(i)}) \end{bmatrix} \hat{e}(t, \theta^{(i)}).$
- Actualizar θ : $\theta^{(i+1)} = \theta^{(i)} - \eta \left(\frac{dJ(\theta)}{d\theta} \right)_{\theta=\theta^{(i)}}^T$
- incrementar i

Sequência de autocorrelação dum processo MA(n)

- Sequência de autocorrelação de um processo estocástico estacionário $y(t)$:

$$\rho_{yy}(\tau) = \frac{\lambda_{yy}(\tau)}{\lambda_{yy}(0)}$$

$\lambda_{yy}(\tau)$ - Sequência de autocovariância de $y(t)$.

- Num processo MA(n), $\lambda_{yy}(\tau) = 0$ para $|\tau| > n \Rightarrow \rho_{yy}(\tau) = 0$ para $|\tau| > n$
- A ordem do modelo dum processo MA pode ser determinada a partir da inspecção da sua sequência de autocorrelação:
 - Será igual ao último instante em que a $\rho_{yy}(\tau) \neq 0$.

Determinação da ordem dum processo MA(n) a partir da sequência autocorrelação

- **Problema:** Como $\rho_{yy}(\tau)$ tem que ser estimado a partir dos dados:

$\hat{\rho}_{yy}(\tau) = 0$ é um evento com probabilidade 0

$\hat{\rho}_{yy}(\tau)$ - Estimativa de $\rho_{yy}(\tau)$.

Determinação da ordem dum processo MA(n) a partir da sequência autocorrelação

- **Problema:** Como $\rho_{yy}(\tau)$ tem que ser estimado a partir dos dados:

$\hat{\rho}_{yy}(\tau) = 0$ é um evento com probabilidade 0

$\hat{\rho}_{yy}(\tau)$ - Estimativa de $\rho_{yy}(\tau)$.

- **Solução:** Se $\rho_{yy}(\tau) = 0$ então aproximadamente

$$\hat{\rho}_{yy}(\tau) \sim \mathcal{N}\left(0, \frac{1}{N}\right)$$

Determinação da ordem dum processo MA(n) a partir da sequência autocorrelação

- **Problema:** Como $\rho_{yy}(\tau)$ tem que ser estimado a partir dos dados:

$\hat{\rho}_{yy}(\tau) = 0$ é um evento com probabilidade 0

$\hat{\rho}_{yy}(\tau)$ - Estimativa de $\rho_{yy}(\tau)$.

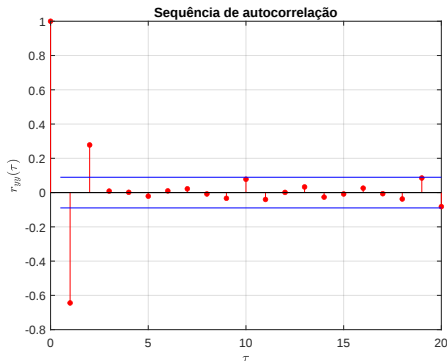
- **Solução:** Se $\rho_{yy}(\tau) = 0$ então aproximadamente

$$\hat{\rho}_{yy}(\tau) \sim \mathcal{N}\left(0, \frac{1}{N}\right)$$

- Consideramos que $\rho_{yy}(\tau) = 0$ com uma *confiança* $(1 - \alpha) \times 100\%$ quando $\hat{\rho}_{yy}(\tau)$ estiver dentro do intervalo de confiança $[-K_\alpha, K_\alpha]$ que, debaixo da distribuição $\mathcal{N}\left(0, \frac{1}{N}\right)$ tem uma probabilidade de ocorrência $1 - \alpha$

Determinação da ordem dum processo MA(n) a partir da sequência autocorrelação

- Se $K_\alpha = 1.96$ o intervalo tem 95% de confiança.
- A ordem do modelo será o último valor de τ para o qual $\hat{\rho}_{yy}(\tau)$ está fora do intervalo de confiança.



Estimação Patamétrica - Processos ARMA

- Um processo ARMA(n,m) é descrito pelo modelo

$$A(q^{-1})y(t) = C(q^{-1})e(t)$$

- $A(q^{-1}) = 1 + a_1q^{-1} + \dots + a_nq^{-n}$
- $C(q^{-1}) = 1 + c_1q^{-1} + \dots + c_mq^{-m}$
- $e(t)$: Ruído branco com média nula e variância σ^2 .

Densidade Espectral dum Processo ARMA

- Os processos ARMA são estacionários
- A densidade espectral é dada por

$$\Phi_{yy}(\omega) = \frac{|C(e^{-j\omega})|^2}{|A(e^{-j\omega})|^2} \sigma^2 = \frac{|1 + c_1 e^{-j\omega} + \dots + c_m e^{-j\omega m}|^2}{|1 + a_1 e^{-j\omega} + \dots + a_n e^{-j\omega n}|^2} \sigma^2$$

- Relacionada diretamente com o modelo
- Estimar $A(q^{-1})$ e $C(q^{-1})$ permite obter a densidade espectral do processo
- A estimação de modelos ARMA é complexa devido à presença simultânea das componentes AR e MA.

Método dos Esfasamentos Superiores

- Permite estimar os parâmetros da componente $AR(n)$ sem a influência da componente $MA(m)$:

Método dos Esfasamentos Superiores

- Permite estimar os parâmetros da componente AR(n) sem a influência da componente MA(m):
- Os coeficientes da componente AR são estimados através as equações de Yule-Walker para $\tau > m$:

$$\begin{aligned}\lambda_{yy}(m)a_1 + \lambda_{yy}(m-1)a_2 + \cdots \lambda_{yy}(m+1-n)a_n &= -\lambda_{yy}(m+1) \\ \lambda_{yy}(m+1)a_1 + \lambda_{yy}(m)a_2 + \cdots \lambda_{yy}(m+2-n)a_n &= -\lambda_{yy}(m+2) \\ &\vdots \\ \lambda_{yy}(m+n)a_1 + \lambda_{yy}(m+n-1)a_2 + \cdots \lambda_{yy}(m)a_n &= -\lambda_{yy}(m+n+1)\end{aligned}$$

Método dos Esfasamentos Superiores

- Permite estimar os parâmetros da componente AR(n) sem a influência da componente MA(m):
- Os coeficientes da componente AR são estimados através as equações de Yule-Walker para $\tau > m$:

$$\begin{aligned}\lambda_{yy}(m)a_1 + \lambda_{yy}(m-1)a_2 + \cdots \lambda_{yy}(m+1-n)a_n &= -\lambda_{yy}(m+1) \\ \lambda_{yy}(m+1)a_1 + \lambda_{yy}(m)a_2 + \cdots \lambda_{yy}(m+2-n)a_n &= -\lambda_{yy}(m+2) \\ &\vdots \\ \lambda_{yy}(m+n)a_1 + \lambda_{yy}(m+n-1)a_2 + \cdots \lambda_{yy}(m)a_n &= -\lambda_{yy}(m+n+1)\end{aligned}$$

- Os resíduos do modelo AR aproximam-se de um processo MA

Método dos Esfasamentos Superiores

- Permite estimar os parâmetros da componente AR(n) sem a influência da componente MA(m):
- Os coeficientes da componente AR são estimados através as equações de Yule-Walker para $\tau > m$:

$$\begin{aligned}\lambda_{yy}(m)a_1 + \lambda_{yy}(m-1)a_2 + \cdots \lambda_{yy}(m+1-n)a_n &= -\lambda_{yy}(m+1) \\ \lambda_{yy}(m+1)a_1 + \lambda_{yy}(m)a_2 + \cdots \lambda_{yy}(m+2-n)a_n &= -\lambda_{yy}(m+2) \\ &\vdots \\ \lambda_{yy}(m+n)a_1 + \lambda_{yy}(m+n-1)a_2 + \cdots \lambda_{yy}(m)a_n &= -\lambda_{yy}(m+n+1)\end{aligned}$$

- Os resíduos do modelo AR aproximam-se de um processo MA
- O modelo MA pode ser estimado a partir destes resíduos utilizando os métodos que estudamos anteriormente.

Ajuste da Resposta Impulsional

- Método idêntico ao da resposta impulsional truncada para estimar modelos MA:
 - A série é aproximada por um modelo $AR(p)$
 - Igualam-se os primeiros $n + m + 1$ valores da resposta impulsional do modelo $ARMA(m, n)$ com os de uma aproximação $AR(p)$.
 - Alternativamente, utiliza-se um estimador de mínimos quadrados caso desejemos aproximar um número de pontos maior do que $n + m + 1$.
- Seja $h(0), h(1), \dots, h(n + m), \dots$ a resposta impulsional da aproximação $AR(p)$

Ajuste da Resposta Impulsional

- Resposta impulsional do modelo ARMA(n, m):

$$y(t) = e(t) + c_1 e(t-1) + \dots + c_m e(t-m) - a_1 y(t-1) - \dots - a_n y(t-n)$$

com

$$e(t) = \begin{cases} 1, & t = 0 \\ 0, & t \neq 0 \end{cases}$$

$$h(0) = e(0) = 1$$

$$h(1) = c_1 - a_1 h(0)$$

$$h(2) = c_2 - a_1 h(1) - a_2 h(0)$$

$$\vdots$$

$$h(m) = -c_m - a_1 h(m-1) - a_2 h(m-2) - \dots - a_m h(0)$$

$$\vdots$$

$$h(n+m) = -a_1 h(n+m-1) - \dots - a_n h(m).$$

Ajuste da Resposta Impulsional

- Resposta impulsional do modelo ARMA(n, m):

$$y(t) = e(t) + c_1 e(t-1) + \dots + c_m e(t-m) - a_1 y(t-1) - \dots - a_n y(t-n)$$

com

$$e(t) = \begin{cases} 1, & t = 0 \\ 0, & t \neq 0 \end{cases}$$

$$h(0) = e(0) = 1$$

$$h(1) = c_1 - a_1 h(0)$$

$$h(2) = c_2 - a_1 h(1) - a_2 h(0)$$

$$\vdots$$

$$h(m) = -c_m - a_1 h(m-1) - a_2 h(m-2) - \dots - a_m h(0)$$

$$\vdots$$

$$h(n+m) = -a_1 h(n+m-1) - \dots - a_n h(m).$$

Cálculo dos parâmetros

- Equações na forma matricial

$$\left[\begin{array}{c|c} \mathbf{I}_m & -\mathcal{H}_{n+m} \\ \hline \mathbf{0}_{n \times m} & \end{array} \right] \boldsymbol{\theta} = \mathbf{h}_{1:n+m} \Leftrightarrow \boldsymbol{\theta} = \left[\begin{array}{c|c} \mathbf{I}_m & -\mathcal{H}_{n+m} \\ \hline \mathbf{0}_{n \times m} & \end{array} \right]^{-1} \mathbf{h}_{1:n+m}$$

$$\bullet \mathcal{H}_{n+m} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ h(1) & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ h(n) & h(n-1) & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ h(n+m) & h(n+m-1) & \cdots & h(m) \end{bmatrix} \in \mathbb{R}^{n+m \times n},$$

- $\boldsymbol{\theta}$ é o vector de parâmetros:

$$\boldsymbol{\theta} = [c_1 \quad \cdots \quad c_m \quad a_1 \quad \cdots \quad a_n]^T \in \mathbb{R}^{n+m}$$

$$\bullet \mathbf{h}_{1:n+m} = [h(1) \quad \cdots \quad h(n+m)]^T \in \mathbb{R}^{n+m}$$

Método dos mínimos quadrados

- expande a abordagem do Método de Mínimos Quadrados para Modelos MA aos modelos ARMA:
 - Aproxima a série temporal por um modelo $AR(p)$
 - Estima o modelo ARMA através do estimador de mínimos quadrados considerando os resíduos como o sinal de entrada
- Equações:

$$y(t) = \hat{e}(t) + c_1 \hat{e}(t-1) + \dots + c_m \hat{e}(t-m) - a_1 y(t-1) - \dots - a_n y(t-n)$$

- $\hat{e}(t)$ - Resíduos do modelo $AR(p)$

Método dos mínimos quadrados

- Forma vetorial:

$$y(t) = \varphi^T(t)\theta + \hat{e}(t)$$

- Regressor:

$$\varphi(t) = [\hat{e}(t-1) \quad \cdots \quad e(t-m) \quad -y(t-1) \quad \cdots \quad -y(t-n)]^T$$

- Parâmetros:

$$\theta = [c_1 \quad \cdots \quad c_m \quad a_1 \quad \cdots \quad a_n]^T$$

Método dos mínimos quadrados

- Estimador:

$$\hat{\theta} = (\Phi^T \Phi)^{-1} \Phi^T Y$$

- Matriz dos regressores:

$$\Phi = \begin{bmatrix} \hat{e}(r) & \cdots & \hat{e}(r-m) & -y(r) & \cdots & -y(r-n) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \hat{e}(N-1) & \cdots & \hat{e}(N-1-m) & -y(N-1) & \cdots & -y(N-1-n) \end{bmatrix}$$

- Vector Y das amostras:

$$Y = [y(r+1) \quad \cdots \quad y(N)]^T$$

- $r = \max(m, n)$.

Minimização do Erro de Previsão

- Minimiza o critério:

$$J(\theta) = \frac{1}{2} \sum_{t=1}^N \hat{e}^2(t, \theta)$$

- $\hat{e}(t, \theta)$ - erro de previsão
- $\theta = [c_1 \ \cdots \ c_m \ a_1 \ \cdots \ a_n]^T$ - vector de parâmetros
- $\hat{\theta} \xrightarrow{\text{probabilidade 1}} \theta_0$ - valor verdadeiro dos parâmetros
- $\hat{\theta}$ é um estimador fortemente consistente
- Solução é anular o gradiente:

$$\left(\frac{dJ(\theta)}{d\theta} \right)^T = \left[\frac{\partial J(\theta)}{\partial c_1} \ \cdots \ \frac{\partial J(\theta)}{\partial c_m} \ \frac{\partial J(\theta)}{\partial a_1} \ \cdots \ \frac{\partial J(\theta)}{\partial a_n} \right]^T = \mathbf{0}.$$

Minimização do Erro de Previsão

- Método de descida de gradiente):

$$\boldsymbol{\theta}^{(i+1)} = \boldsymbol{\theta}^{(i)} - \eta \left(\frac{dJ(\boldsymbol{\theta})}{d\boldsymbol{\theta}} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(i+1)}}^T$$

- $\eta > 0$ - taxa de aprendizagem.

- $$\frac{\partial J(\boldsymbol{\theta})}{\partial c_k} = \sum_{t=1}^N \frac{\partial \hat{e}(t, \boldsymbol{\theta})}{\partial c_k} \hat{e}(t, \boldsymbol{\theta}), \quad k = 1, \dots, m$$

- $$\frac{\partial J(\boldsymbol{\theta})}{\partial a_k} = \sum_{t=1}^N \frac{\partial \hat{e}(t, \boldsymbol{\theta})}{\partial a_k} \hat{e}(t, \boldsymbol{\theta}), \quad k = 1, \dots, n$$

- Erro de previsão: $\hat{e}(t, \boldsymbol{\theta}) = y(t) - \hat{y}(t, \boldsymbol{\theta})$

- Previsão:
$$\hat{y}(t, \boldsymbol{\theta}) = \frac{C(q^{-1}) - A(q^{-1})}{C(q^{-1})} y(t)$$

Minimização do Erro de Previsão

- Equações às diferenças:

$$\hat{y}(t, \theta) = c_1 \hat{e}(t-1, \theta) + \dots + c_m \hat{e}(t-m, \theta) - a_1 y(t-1) - \dots - a_n y(t-n)$$

$$\begin{aligned} \hat{e}(t, \theta) = & -c_1 \hat{e}(t-1, \theta) - \dots - c_m \hat{e}(t-m, \theta) + \\ & + y(t) + a_1 y(t-1) + \dots + a_n y(t-n) \end{aligned}$$

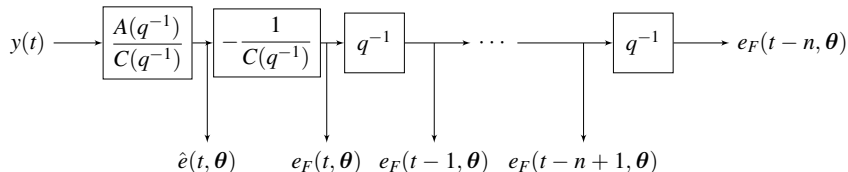
- Derivadas do erro de previsão:

$$\begin{aligned} \frac{\partial \hat{e}(t, \theta)}{\partial c_k} &= -\hat{e}(t-k, \theta) - c_1 \frac{\partial \hat{e}(t-1, \theta)}{\partial c_k} - \dots - c_m \frac{\partial \hat{e}(t-m, \theta)}{\partial c_k} = \\ &= -\frac{1}{C(q^{-1})} \hat{e}(t-k, \theta), \quad k = 1, \dots, m. \end{aligned}$$

$$\begin{aligned} \frac{\partial \hat{e}(t, \theta)}{\partial a_k} &= y(t-k, \theta) - c_1 \frac{\partial \hat{e}(t-1, \theta)}{\partial c_k} - \dots - c_m \frac{\partial \hat{e}(t-m, \theta)}{\partial c_k} = \\ &= \frac{1}{C(q^{-1})} y(t-k, \theta), \quad k = 1, \dots, m. \end{aligned}$$

Minimização do Erro de Previsão

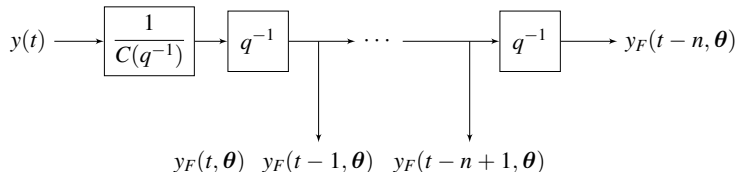
- $\frac{\partial \hat{e}(t, \theta)}{\partial c_k} = \hat{e}_F(t - k, \theta)$
- $\hat{e}_F(t, \theta) = -\frac{1}{C(q^{-1})} \hat{e}(t, \theta)$



Minimização do Erro de Previsão

- $\frac{\partial \hat{e}(t, \theta)}{\partial a_k} = y_F(t - k, \theta)$

- $y_F(t, \theta) = \frac{1}{C(q^{-1})}y(t)$



Algoritmo

Minimização do erro de previsão:

Inicialização: Obter $\theta^{(0)}$ e definir $i = 0$.

Repetir até convergir:

- $\hat{e}(t, \theta^{(i)}) = \frac{A^{(i)}(q^{-1})}{C^{(i)}(q^{-1})} y(t),$
 $t = 1, \dots, N$
- $\hat{e}_F(t, \theta^{(i)}) = -\frac{1}{C^{(i)}(q^{-1})} \hat{e}(t, \theta^{(i)}),$
 $t = 1, \dots, N$
- $y_F(t, \theta^{(i)}) = \frac{1}{C^{(i)}(q^{-1})} \hat{e}(t, \theta^{(i)}),$
 $t = 1, \dots, N$

Algoritmo (continuação)

Minimização do erro de previsão (continuação):

- Gradiente do custo:

$$\left(\frac{dJ(\theta)}{d\theta} \right)_{\theta=\theta^{(i)}}^T = \sum_{t=1}^N \begin{bmatrix} \hat{e}_F(t-1, \theta^{(i)}) \\ \vdots \\ \hat{e}_F(t-m, \theta^{(i)}) \\ y_F(t-1, \theta^{(i)}) \\ \vdots \\ y_F(t-n, \theta^{(i)}) \end{bmatrix} \hat{e}(t, \theta^{(i)}).$$

- Actualizar parâmetros:

$$\theta^{(i+1)} = \theta^{(i)} - \eta \left(\frac{dJ(\theta)}{d\theta} \right)_{\theta=\theta^{(i)}}^T$$

- Incrementar i .