

se exemplar com a redação final da tese
defendida por NÁDIA DOLORES GIMENEZ
pela Comissão
Julgadora em 19 07 1994
Basílio E. A. Milani
Orientador

Universidade Estadual de Campinas
FACULDADE DE ENGENHARIA ELÉTRICA
DEPARTAMENTO DE TELEMÁTICA



**Estimação Adaptativa de Parâmetros
Através de Mínimos Quadrados
Sujeito a Restrições Lineares de Igualdade**

Tese apresentada à Faculdade de Engenharia Elétrica da Universidade Estadual de
Campinas, como parte dos requisitos exigidos para a obtenção do título de
Mestre em Engenharia Elétrica.

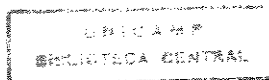
por

Nádia Dolores Gimenez
Matemática - UFSCAR - São Carlos

em 19 de julho de 1994 perante a banca examinadora

Basílio E. A. Milani
Basílio E. A. Milani - Orientador

Samuel Siqueira Bueno - IA/CTI
Akebo Yamakami - FEE/UNICAMP



Para Dona Tica, minha Avó e
para Neide, minha Mãe.

*“Sem esforço de nossa parte, jamais atingiremos o alto da montanha,
pois só atingirá o cimo da montanha se estiver decidido
a enfrentar o esforço da caminhada”.*

Livro - “Minutos de Sabedoria”

Agradecimentos

Meus sinceros agradecimentos a todos aqueles que, direta ou indiretamente, contribuíram de alguma forma para a realização deste trabalho, e em especial:

- Ao Professor Dr. Basílio E. A. Milani pela oportunidade a mim concedida e pela orientação, incentivo, paciência e amizade;
- Ao meu pai Germinal e ao resto da “*cambada*” que eu **tanto amo**: Ninho, Rui, Aydil, Vitão, Guiga e Henricão. Obrigada por tudo;
- Ao meu grande amigo e “*irmão*” Osvaldo Júnior, pelo seu apoio e incentivo, pelos seus conselhos, pela sua amizade e pela força sempre presente e sincera;
- Aos meus amigos Cearenses, Paulistas, Baianos, Mineiros e tantos outros pelo companheirismo e ajuda nestes anos;
- Aos meus amigos Tieteenses e Campineiros pelo incentivo, motivação e bons momentos;
- Aos professores, amigos e funcionários do departamento de Telemática e da CPG pela compreensão e atendimento sempre prestados;
- À Banca Examinadora, pela aceitação e pela participação ;
- Ao governo brasileiro, através do CNPq, pelo suporte financeiro;
- À DEUS, por tudo.

RESUMO

Este trabalho trata da estimação adaptativa de parâmetros através de mínimos quadrados lineares sujeitos a restrições lineares de igualdade. São propostos algoritmos recursivos para solução de problemas de mínimos quadrados com ponderação exponencial de dados e janela móvel de dados, combinando equações normais e fatoração ortogonal com espaço nulo da matriz de restrições e multiplicadores de Lagrange. Um exemplo numérico ilustra o desempenho dos algoritmos propostos no projeto de um filtro de resposta finita (FIR) para detecção de frequências presentes em um sinal com ruído.

ABSTRACT

This work is concerned with the adaptive parameter estimation via a linear least-squares subject to linear equality constraints. Recursive algorithms are proposed for solution of least-squares problems with exponential data weighting and sliding data window, by combining normal equations and orthogonal factorization with the null space of the constraint matrix and Lagrange multipliers. A numerical example illustrates the proposed algorithm performance in the design of a finite impulsive response filter (FIR) for frequency detection in a signal with noise.

Conteúdo

1	Introdução Geral	2
2	Estimação de Parâmetros por Mínimos Quadrados Lineares Sujeito a Restrições de Igualdade Lineares	4
2.1	Introdução	4
2.2	Definição do Problema	4
2.3	Métodos utilizando o Espaço Nulo de C ($\mathcal{N}_{(C)}$)	6
2.3.1	Equações Normais	7
2.3.2	Fatoração Ortogonal	9
2.4	Métodos utilizando Multiplicadores de Lagrange (λ)	11
2.4.1	Equações Normais	12
2.4.2	Fatoração Ortogonal	13
2.5	Análise Comparativa	16
3	Estimação Adaptativa de Parâmetros com Restrições de Igualdade utilizando Ponderação Exponencial dos Dados	18
3.1	Introdução	18
3.2	Definição do Problema	18
3.3	Métodos utilizando o Espaço Nulo de C ($\mathcal{N}_{(C)}$)	20
3.3.1	Equações Normais	22
3.3.2	Fatoração Ortogonal	25
3.4	Métodos utilizando Multiplicadores de Lagrange (λ)	28
3.4.1	Equações Normais	29
3.4.2	Fatoração Ortogonal	32
3.5	Análise Comparativa	35
4	Estimação Adaptativa de Parâmetros com Restrições de Igualdade utilizando Janela Móvel de Dados	37
4.1	Introdução	37
4.2	Definição do Problema	37
4.3	Métodos utilizando o Espaço Nulo de C ($\mathcal{N}_{(C)}$)	39

4.3.1	Equações Normais	40
4.3.2	Fatoração Ortogonal	44
4.4	Métodos utilizando Multiplicadores de Lagrange (λ)	48
4.4.1	Equações Normais	50
4.4.2	Fatoração Ortogonal	54
4.5	Análise Comparativa	60
5	Exemplo Numérico	62
5.1	Introdução	62
5.2	Formulação do Problema	62
5.3	Resultados Obtidos	64
A		80
A.1	Lema de Inversão de Matrizes	80
A.2	Fatoração QR utilizando matrizes de rotação de Givens	81
	Bibliografia	84

Capítulo 1

Introdução Geral

Estimação de parâmetros por mínimos quadrados lineares, é de grande importância em muitas aplicações, tais como: Análise de séries temporais, comunicações, análise espectral, controle, etc. [2], [3], [6], [8].

A estimação adaptativa por mínimos quadrados de sistemas lineares com parâmetros variantes no tempo, é geralmente obtida através de algoritmos recursivos utilizando as “Equações Normais de Mínimos Quadrados” com Ponderação Exponencial dos dados, para levar em conta a maior importância das informações atuais sobre as do passado [3] [6].

A utilização conjunta de equações normais e ponderação exponencial leva a algoritmos recursivos simples, com baixo custo computacional, mas com qualidade numérica precária, quando utilizados em problemas de mínimos quadrados mal condicionados, porque tanto a montagem das equações normais, quanto a utilização da Ponderação Exponencial, pioram sobremaneira o condicionamento original, inerente ao problema a ser resolvido [5], [7].

Formas alternativas de contornar este problema são: a utilização de Fatores Ortogonais QR ao invés de “Equações Normais” e usar Janela Móvel no lugar de Ponderação Exponencial dos Dados [16].

Devido a problemas de ordem prática sobre intervalo de validade dos parâmetros ou técnicas de síntese de sistemas de controle e/ou filtros é muito comum aparecer problemas de mínimos quadrados sujeito a restrições nos parâmetros. Em particular é bastante comum restrições lineares de igualdade e/ou desigualdade.

Duas formas básicas, bastante difundidas [15], de tratamento de problemas de mínimos quadrados com restrições lineares de igualdade, são a utilização do espaço nulo da matriz de restrições ou os multiplicadores de Lagrange. A escolha de uma dessas métodos envolve considerações sobre o número de restrições, condicionamento do problema de mínimos quadrados, forma de solução do problema de mínimos quadrados, etc [15].

Neste trabalho trataremos da solução recursiva de problemas de mínimos quadrados lineares sujeitos a restrições lineares de igualdade nos parâmetros. Serão considerados algoritmos baseados nas equações normais, fatoração ortogonal combinados com ponderação exponencial e janela móvel de dados. Para tratamento das restrições serão consideradas métodos baseadas no espaço nulo da matriz de restrições e multiplicadores de Lagrange. Os algoritmos desenvolvidos serão analisados, comparados entre si e ilustrados utilizando um exemplo significativo de síntese de filtros FIR (Resposta Impulsiva Finita).

A matéria se distribue ao longo dos capítulos da seguinte forma:

- Capítulo 2: trata da solução do problema de mínimos quadrados lineares, sujeito a restrições de igualdade lineares. São apresentados algoritmos conhecidos, não recursivos, para solução do problema, utilizando equações normais e fatoração ortogonal combinadas com espaço nulo da matriz de restrições e multiplicadores de Lagrange.
- Capítulo 3: trata da solução do problema de mínimos quadrados lineares com ponderação exponencial dos dados, sujeito a restrições de igualdade lineares. São apresentados algoritmos recursivos utilizando equações normais e fatoração ortogonal combinadas com espaço nulo da matriz de restrições e multiplicadores de Lagrange.
- Capítulo 4: trata da solução do problema de mínimos quadrados lineares com janela móvel dos dados, sujeito a restrições de igualdade lineares. São apresentados algoritmos recursivos utilizando equações normais e fatoração ortogonal combinadas com espaço nulo da matriz de restrições e multiplicadores de Lagrange.
- Capítulo 5: para ilustrar e comparar o desempenho dos algoritmos apresentados nos capítulos 3 e 4, neste capítulo é apresentado um exemplo numérico significativo, correspondente a um problema de síntese de um filtro de resposta impulsiva finita (FIR) para detecção de frequências presentes em um sinal com ruídos.

Capítulo 2

Estimação de Parâmetros por Mínimos Quadrados Lineares Sujeito a Restrições de Igualdade Lineares

2.1 Introdução

Neste capítulo, o problema de estimação de parâmetros utilizando mínimos quadrados lineares sujeito a restrições de igualdade lineares é definido. Para a solução de problema de mínimo quadrado são apresentadas abordagens, combinando equações normais e fatorações ortogonais com o uso do espaço nulo da matriz de restrições e do multiplicador de lagrange, para tratamento das restrições.

2.2 Definição do Problema

Consideremos o problema de estimar um vetor de parâmetros h de dimensões $(n, 1)$, solução da equação de regressão linear

$$y(i) = a(i)h(i) + \omega(i) \tag{2.1}$$

$$a(i) = [a_{i1} \ a_{i2} \ a_{i3} \ \dots \ a_{in}]$$

sujeito a seguinte restrição de igualdade:

$$Ch(i) = d \quad (2.2)$$

onde : $a(i)$ é um vetor de entrada de dimensões $(1, n)$, $\omega(i)$ é um ruído branco com média zero, $y(i)$ é um sinal de referência no instante i , C é uma matriz de dimensões (r, n) de posto r , $r < n$ e d é um vetor de dimensões $(r, 1)$.

Assumindo h estacionário no intervalo $0 < k < i$, a estimação por mínimos quadrados de h no instante i , corresponde à solução do seguinte problema de otimização :

$$\min_{h(i)} J(i) = [h(i) - h(0)]^T A_0 [h(i) - h(0)] + \sum_{k=1}^i [y(k) - a(k)h(i)]^2 \quad (2.3)$$

s.a

$$Ch(i) = d \quad (2.4)$$

onde : A_0 é uma matriz definida positiva representando o peso da estimativa a priori $h(0)$ relativamente aos dados no intervalo $0 < k < i$.

É fácil verificar que, podemos formular o problema de mínimos quadrados com restrições (PMQR) da seguinte forma matricial [16]:

$$\min_{h(i)} J(i) = \left\| \begin{bmatrix} R(0) \\ A(i) \end{bmatrix} [h(i) - h(0)] - \begin{bmatrix} 0 \\ Y(i) - A(i)h(0) \end{bmatrix} \right\|_2 \quad (2.5)$$

s.a

$$C[h(i) - h(0)] = 0$$

$$Ch(0) = d \quad (2.6)$$

onde :

$$A(i) = \begin{bmatrix} a(1) \\ a(2) \\ \vdots \\ a(i) \end{bmatrix} \text{ matriz de dados de entrada de dimensões } (i, n),$$

$$Y(i) = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(i) \end{bmatrix} \text{ vetor dos sinais de referência de dimensões } (i, 1),$$

$A_0 = R(0)^T R(0)$, $R(0)$ matriz triangular superior de dimensões (n, n) obtida da fatoração de Cholesky de A_0 .

Podemos obter a solução do PMQR (2.5) de quatro formas principais: métodos utilizando o espaço nulo combinado com equações normais ou fatoração ortogonal e métodos utilizando multiplicadores de Lagrange combinado com equações normais ou fatoração ortogonal. Veremos nas próximas seções como cada método resolve o PMQR (2.5) juntamente com uma análise de suas vantagens e desvantagens computacionais.

2.3 Métodos utilizando o Espaço Nulo de C ($\mathcal{N}_{(C)}$)

Seja

$$Z = \mathcal{N}_{(C)} \quad (2.7)$$

onde: Z é uma matriz ortogonal de dimensões $(n, n-r)$ correspondente a uma base do espaço nulo de C .

Considerando o problema de minimização na forma matricial (2.5) e assumindo a seguinte mudança de variáveis:

$$h(i) - h(0) = Zp \quad (2.8)$$

onde: p é um vetor de dimensões $(n-r, 1)$, a restrição

$$C[h(i) - h(0)] = 0 \quad (2.9)$$

é automaticamente satisfeita para qualquer p , temos o seguinte problema de minimização irrestrito:

$$\min_{p(i)} J(i) = \left\| \begin{bmatrix} R(0) \\ A(i) \end{bmatrix} Zp - \begin{bmatrix} 0 \\ Y(i) - A(i)h(0) \end{bmatrix} \right\|_2 \quad (2.10)$$

para solução deste problema de mínimos quadrados é bem conhecido que podemos utilizar o método das equações normais ou fatoração ortogonal [15], que descrevemos sumariamente a seguir.

2.3.1 Equações Normais

Para obter a solução do problema (2.10) podemos impor a condição de ponto estacionário:

$$\nabla J(i) = 0 \quad (2.11)$$

onde $\nabla J(i)$ é o vetor gradiente de $J(i)$.

Isto corresponde a [5], [6], [8],

$$\nabla J(i) = Z^T \begin{bmatrix} R(0) \\ A(i) \end{bmatrix}^T \left\{ \begin{bmatrix} 0 \\ Y(i) - A(i)h(0) \end{bmatrix} - \begin{bmatrix} R(0) \\ A(i) \end{bmatrix} Zp \right\} = 0 \quad (2.12)$$

Rearranjando os termos chegamos à “equação normal”

$$Z^T \begin{bmatrix} R(0) \\ A(i) \end{bmatrix}^T \begin{bmatrix} R(0) \\ A(i) \end{bmatrix} Zp = Z^T \begin{bmatrix} R(0) \\ A(i) \end{bmatrix}^T \begin{bmatrix} 0 \\ Y(i) - A(i)h(0) \end{bmatrix} \quad (2.13)$$

que podemos escrever na forma condensada :

$$(XZ)^T (XZ)p = (XZ)^T b \quad (2.14)$$

onde :

$$X = \begin{bmatrix} R(0) \\ A(i) \end{bmatrix} \quad (2.15)$$

$$b = \begin{bmatrix} 0 \\ Y(i) - A(i)h(0) \end{bmatrix} \quad (2.16)$$

Assim a solução de (2.10) é dada por:

$$h(i) = h(0) + Zp \quad (2.17)$$

onde p é obtido da solução do sistemas de equações lineares (2.14) de ordem $(n - r)$, ou seja, da dimensão do espaço nulo de C .

Podemos ver que o esforço computacional para obtenção de $h(i)$ é devido principalmente a :

- Produto de matrizes XZ : $n(n + i)(n - r)$ FLOPS
- Produto de matrizes $(XZ)^T(XZ)$: $(n + i)(n - r)^2$ FLOPS
- Produto de matrizes $A(i)h(0)$: ni FLOPS
- Produto de matrizes $(XZ)^Tb$: $(n - r)i$ FLOPS
- Produto de matrizes Zp : $(n - r)n$ FLOPS
- Solução do sistema de equações (2.13) pelo método de Cholesky : $\frac{(n - r)^3}{6} + \frac{(n - r)^2}{2}$ FLOPS

onde FLOPS é o número de operações de multiplicações e divisões de ponto flutuante.

Totalizando, temos o seguinte esforço computacional:

$$FLOPS = (n - r)[n(n + i) + n + i] + (n - r)^2 \left[(n + i) + \frac{(n - r)}{6} + \frac{1}{2} \right] + ni \quad (2.18)$$

A precisão da solução do sistema de equações (2.14) depende do número condição da matriz do sistema $(XZ)^T(XZ)$.

É fácil verificar que [5] :

$$\text{cond}[(XZ)^T(XZ)] = \text{cond}[(XZ)]^2 \quad (2.19)$$

onde $\text{cond}[\cdot]$ é o número condição de $[\cdot]$.

2.3.2 Fatoração Ortogonal

Para obter solução do problema (2.10) vamos nos basear nas seguintes propriedades das transformações ortogonais [14]:

$$\|Qy\|_2 = \|y\|_2 \quad (2.20)$$

$$\text{cond}(QG) = \text{cond}(G) \quad (2.21)$$

onde :

y é um vetor de dimensões $(n, 1)$,

Q é uma matriz ortogonal de dimensões (n, n) , $(Q^T Q = I_n)$,

G é uma matriz com n linhas,

$\|\cdot\|_2$ é o quadrado da norma euclidiana e

$\text{cond}(\cdot)$ é o número condição de $(\cdot)$.

A propriedade (2.20) permite transformar o índice de desempenho quadrático do problema (2.10) da seguinte forma :

$$\begin{aligned} \min_{p(i)} J(i) &= \left\| Q \begin{bmatrix} R(0) \\ A(i) \end{bmatrix} Zp - Q \begin{bmatrix} 0 \\ Y(i) - A(i)h(0) \end{bmatrix} \right\|_2 \\ \min_{p(i)} J(i) &= \left\| \begin{bmatrix} R \\ 0 \end{bmatrix} p - \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \right\|_2 \end{aligned} \quad (2.22)$$

sendo,

$$Q \begin{bmatrix} R(0) \\ A(i) \end{bmatrix} Z = QXZ = \begin{bmatrix} R \\ 0 \end{bmatrix} ; Q \begin{bmatrix} 0 \\ Y(i) - A(i)h(0) \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad (2.23)$$

onde :

Q matriz ortogonal de dimensões $(n+i, n+i)$,

R matriz triangular superior de dimensões $(n-r, n-r)$,

b_1 vetor de dimensões $(n-r, 1)$ e

b_2 vetor de dimensões $(i+r, 1)$.

A matriz Q pode ser sempre obtida de forma eficiente e numericamente estável através do produto de uma sequência de matrizes de Householder ou Givens [14].

Para a obtenção da matriz $\begin{bmatrix} R \\ 0 \end{bmatrix}$ e do vetor $\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$, não é necessário construir Q explicitamente [14], [5], [Apêndice A.2].

Devido à estrutura de $\begin{bmatrix} R \\ 0 \end{bmatrix}$, verificamos que a solução do (2.10) é dada por:

$$h(i) = h(0) + Zp \quad (2.24)$$

onde: p é obtido da solução do seguinte sistemas de equações de dimensões $(n-r)$

$$Rp = b_1 \quad (2.25)$$

O esforço computacional para obtenção de $h(i)$ é devido à obtenção de Z (2.7), XZ , R e b_1 (2.23), solução do sistema triangular de equações (2.25) e Zp (2.24). Utilizando matrizes de Householder para obtenção de Z , R e b_1 e substituição regressiva para solução do sistema (2.25), o esforço computacional para obtenção de $h(i)$ é estimado em :

- Obtenção de Z : $\frac{r^2}{3}(3n-r) + \frac{(1+n)}{2}(n-r)$ FLOPS
- Produto de matrizes XZ : $n(n+i)(n-r)$ FLOPS

- Obtenção de R : $\frac{(n-r)^2}{3}(3n+3i-n-r)$ FLOPS
- Obtenção de b_1 : $2(n-r)(n+i)$ FLOPS
- Produto de matrizes Zp : $(n-r)n$ FLOPS
- Solução do sistema de equações (2.23): $\frac{(n-r)}{2}(n-r+1)$ FLOPS

onde FLOPS é o número de operações de multiplicações e divisões de ponto flutuante.

Totalizando, temos o seguinte esforço computacional:

$$FLOPS = \frac{r^2}{3}(3n-r) + \frac{(n-r)^2}{3}(3n+3i-n-r) + \frac{(n-r)}{2}(2n^2+2ni+8n+4i-r+2) \quad (2.26)$$

A precisão da solução do sistemas de equações (2.25) depende do número condição de R .

De (2.21) e (2.23) é fácil verificar que:

$$cond(QXZ) = cond(XZ) = cond(R) \quad (2.27)$$

2.4 Métodos utilizando Multiplicadores de Lagrange (λ)

Para solução do problema (2.5), as equações básicas do método dos multiplicadores de Lagrange são:

$$\nabla J(i) + C^T \lambda = 0 \quad (2.28)$$

$$C[h(i) - h(0)] = 0 \quad (2.29)$$

A seguir descreveremos os métodos de solução utilizando equações normais e fatoração ortogonal.

2.4.1 Equações Normais

É bem conhecido que [15]:

$$\nabla J(i) = X^T X[h(i) - h(0)] - X^T b \quad (2.30)$$

onde X é definido por (2.15) e b é definido por (2.16); portanto para obter a solução do problema (2.5), usando (2.28) e (2.29), temos:

$$X^T X[h(i) - h(0)] - X^T b + C^T \lambda = 0 \quad (2.31)$$

$$C[h(i) - h(0)] = 0 \quad (2.32)$$

Multiplicando (2.31) por $(X^T X)^{-1}$, temos:

$$[h(i) - h(0)] - (X^T X)^{-1} X^T b + (X^T X)^{-1} C^T \lambda = 0 \quad (2.33)$$

Multiplicando (2.33) por C e substituindo o resultado (2.32), temos:

$$-C(X^T X)^{-1} X^T b + C(X^T X)^{-1} C^T \lambda = 0 \quad (2.34)$$

portanto, λ é obtido da solução do seguinte sistema de equações de ordem r

$$[C(X^T X)^{-1} C^T] \lambda = C(X^T X)^{-1} X^T b \quad (2.35)$$

e a solução de (2.5) é dada por:

$$h(i) = h(0) + \hat{h}(i) - Y \lambda \quad (2.36)$$

onde: $\hat{h}(i)$ e Y são soluções do seguinte sistema matricial de equação de ordem n

$$(X^T X)[\hat{h}(i), Y] = [X^T b, C^T] \quad (2.37)$$

De (2.35), (2.36) e (2.37), podemos ver que o esforço computacional para obtenção de $h(i)$ é devido principalmente a:

- Produto de matrizes $X^T X : n^2(n + i)$ FLOPS
- Produto de matrizes $X^T b : n(n + i)$ FLOPS

- Produto de matrizes $Y\lambda$: nr FLOPS
- Produto de matrizes $CY, C\hat{h}(i)$: $nr^2 + nr$ FLOPS
- Solução do sistema de equações (2.35) : $\frac{r^3}{6} + \frac{r^2}{2}$ FLOPS
- Solução do sistema de equações (2.37) : $\frac{n^3}{3} + \frac{(2r+3)n^2}{3} + n$ FLOPS

onde FLOPS é o número de operações de multiplicações e divisões de ponto flutuante.

Totalizando, temos o seguinte esforço computacional:

$$FLOPS = n^2 \left[\frac{(2r+3)}{3} + (n+i) \right] + n \left[\frac{n^2}{3} + r^2 + 2r + n + i + 1 \right] + \frac{r^2}{2} \left[\frac{r}{3} + 1 \right] \quad (2.38)$$

A precisão da solução do problema de mínimos quadrados depende principalmente da solução do sistema de equações (2.37) e portanto do número condição da matriz $(X^T X)$. É fácil verificar [5] que:

$$cond(X^T X) = cond(X)^2 \quad (2.39)$$

2.4.2 Fatoração Ortogonal

Para obter solução do problema (2.5), usando a propriedade (2.20) temos:

$$\min_{h(i)} J(i) = \left\| Q \begin{bmatrix} R(0) \\ A(i) \end{bmatrix} [h(i) - h(0)] - Q \begin{bmatrix} 0 \\ Y(i) - A(i)h(0) \end{bmatrix} \right\|_2$$

$$\min_{h(i)} J(i) = \left\| \begin{bmatrix} R \\ 0 \end{bmatrix} [h(i) - h(0)] - \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \right\|_2 \quad (2.40)$$

s.a

$$C[h(i) - h(0)] = 0 \quad (2.41)$$

sendo, Q uma matriz ortogonal tal que:

$$Q \begin{bmatrix} R(0) \\ A(i) \end{bmatrix} = QX = \begin{bmatrix} R \\ 0 \end{bmatrix}; \quad Q \begin{bmatrix} 0 \\ Y(i) - A(i)h(0) \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad (2.42)$$

Usando (2.28) e (2.29), temos:

$$R^T R[h(i) - h(0)] - R^T b_1 + C^T \lambda = 0 \quad (2.43)$$

$$C[h(i) - h(0)] = 0 \quad (2.44)$$

Multiplicando (2.43) por $(R^T)^{-1}$, temos:

$$R[h(i) - h(0)] - b_1 + (R^T)^{-1} C^T \lambda = 0 \quad (2.45)$$

Multiplicando (2.45) por R^{-1} , temos:

$$h(i) - h(0) - R^{-1} b_1 + R^{-1} (R^T)^{-1} C^T \lambda = 0 \quad (2.46)$$

Multiplicando (2.46) por C e substituindo o resultado (2.44), temos:

$$-CR^{-1} b_1 + CR^{-1} (R^T)^{-1} C^T \lambda = 0 \quad (2.47)$$

portanto, é fácil verificar que λ é a solução do seguinte problema de mínimos quadrados de ordem r :

$$\min_{\lambda} \|Y\lambda - b_1\|_2 \quad (2.48)$$

onde: Y é solução do seguinte sistema de equações matriciais:

$$R^T Y = C^T \quad (2.49)$$

Portanto, a solução de (2.5) é dada por:

$$h(i) = h(0) + \hat{h}(i) - Y\lambda \quad (2.50)$$

onde: $\hat{h}(i)$ é dado pela solução do seguinte sistema de equações de ordem n :

$$R\hat{h}(i) = b_1 \quad (2.51)$$

De (2.48), (2.49), (2.50) e (2.51), é fácil verificar que o esforço computacional requerido é devido principalmente a:

- Obtenção de $R(0)$ (fatoração de Cholesky) : $\frac{n^3}{6}$ FLOPS
- Obtenção de R e b_1 : $n^2(n + i - \frac{n}{3} + 2)$ FLOPS
- Produto de matrizes $Y\lambda$: nr FLOPS
- Solução do sistema matricial triangular (2.49) : $r\frac{n^2}{2}$ FLOPS
- Solução do sistema triangular (2.51) : $\frac{n^2}{2}$ FLOPS
- Solução do problema de mínimos quadrados (2.48) : $r^2(n - \frac{r}{6}) + \frac{5}{2}r^2$ FLOPS

onde FLOPS é o número de operações de multiplicações e divisões de ponto flutuante.

Totalizando, temos o seguinte esforço computacional:

$$FLOPS = \frac{n^3}{6} + n^2 \left[n + i - \frac{n}{3} + 2 + \frac{r}{2} + \frac{1}{2} \right] + r^2 \left[n - \frac{r}{6} + \frac{5}{2} \right] + nr \quad (2.52)$$

A precisão da solução do problema de mínimos quadrados depende da solução do sistema de equações (2.49) e (2.51), e portanto do número condição da matriz R e R^T .

De (2.42) é fácil verificar que:

$$\text{cond } Q \begin{bmatrix} R(0) \\ A(i) \end{bmatrix} = \text{cond}(QX) = \text{cond}(R) = \text{cond}(R^T) \quad (2.53)$$

2.5 Análise Comparativa

Para comparar o esforço computacional requerido pelos vários métodos, vamos considerar um problema com as seguintes características:

$$n = 10; \quad r = 2; \quad i = 20$$

De (2.18), (2.26), (2.38), (2.52), temos a seguinte estimativa de FLOPS:

- Z com Equações Normais: 4437
- Z com Fatoração QR : 4741
- λ com Equações Normais: 3960
- λ com Fatoração QR : 3252

Para $r=5$, temos os seguintes FLOPS:

- Z com Equações Normais: 2283
- Z com Fatoração QR : 2725
- λ com Equações Normais: 4460
- λ com Fatoração QR : 3675

Desta forma, podemos concluir que:

- Para baixos valores de r/n , os algoritmos utilizando multiplicadores de Lagrange λ requerem menor esforço computacional que seus similares utilizando o espaço nulo Z . Para $r/n > 1/2$ essa relação é invertida.
- Com o espaço nulo Z , o uso das equações normais requer menor esforço computacional que o uso da fatoração QR. Com os multiplicadores de Lagrange λ essa relação é invertida.

De (2.19), (2.27), (2.39), (2.53), temos que a precisão dos resultados obtidos dependem principalmente de:

- Z com Equações Normais: $\text{cond}(XZ)^2$

- Z com Fatoração QR : $\text{cond}(XZ)$
- λ com Equações Normais: $\text{cond}(X)^2$
- λ com Fatoração QR : $\text{cond}(X)$

[15]: Como Z é uma matriz ortogonal com $\text{posto}(Z) < \text{posto}(X)$, é bem conhecido

$$\text{cond}(XZ) \leq \text{cond}(X) \quad (2.54)$$

Desta forma, podemos concluir que:

- O condicionamento numérico dos algoritmos utilizando fatoração QR é muito melhor que seus similares utilizando equações normais. O uso das equações normais não é recomendado para problemas inerentemente mal condicionados ($\text{cond}(X)$ elevado).
- O condicionamento numérico dos algoritmos utilizando o espaço nulo de Z é superior ao de seus similares utilizando multiplicadores de Lagrange.

Capítulo 3

Estimação Adaptativa de Parâmetros com Restrições de Igualdade utilizando Ponderação Exponencial dos Dados

3.1 Introdução

Neste capítulo, vamos apresentar algoritmos recursivos para a solução do problema de mínimos quadrados lineares com ponderação exponencial dos dados sujeito a restrições de igualdade, utilizando métodos baseados no espaço nulo de C e multiplicadores de Lagrange, combinadas com equações normais e fatoração ortogonal.

3.2 Definição do Problema

Com a finalidade de seguir as não estacionaridades de um processo, um fator de esquecimento exponencial W é introduzido ao PMQR (2.3). Da mesma forma, considere o problema de estimar um vetor de parâmetros h de dimensões $(n, 1)$ com variação temporal lenta, solução da equação de regressão linear dado em (2.1). Assumindo h constante em $0 < k \leq i$, a estimação por mínimos quadrados de h no

instante i , corresponde à solução do seguinte problema :

$$\min_{h(i)} J(i) = [h(i) - h(0)]^T A_0 [h(i) - h(0)] + \sum_{k=1}^i W^{2(i-k)} [y(k) - a(k)h(i)]^2 \quad (3.1)$$

s.a

$$\begin{aligned} Ch(0) &= d \\ Ch(i) &= d \end{aligned} \quad (3.2)$$

onde : $A_0 = R(0)^T W^{2i} R(0)$ é uma matriz definida positiva representando o peso da estimativa a priori $h(0)$ relativamente aos dados no intervalo $0 < k \leq i$ e W é o fator de esquecimento $0 < W < 1$.

Sendo assim podemos formular o nosso problema de mínimos quadrados não recursivo sujeito a restrições , com ponderação exponencial dos dados PMQRPE, na seguinte forma matricial:

$$\min_{h(i)} J(i) = \left\| \begin{bmatrix} W^i R(0) \\ \bar{W}(i) A(i) \end{bmatrix} [h(i) - h(0)] - \begin{bmatrix} 0 \\ \bar{W}(i) [Y(i) - A(i)h(0)] \end{bmatrix} \right\|_2 \quad (3.3)$$

s.a

$$Ch(i) = d \quad (3.4)$$

$$\text{onde: } \bar{W}(i) = \begin{bmatrix} W^{i-1} & & & \\ & \ddots & & \\ & & W^2 & \\ & & & W \\ & & & & 1 \end{bmatrix} \text{ de dimensões } (i, i)$$

$$A(i) = \begin{bmatrix} a(1) \\ a(2) \\ \vdots \\ a(i) \end{bmatrix} \text{ matriz de dados de entrada de dimensões } (i, n)$$

$$Y(i) = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(i) \end{bmatrix} \text{ vetor dos sinais de referência de dimensões } (i, 1)$$

Fazendo

$$\bar{X}(i) = \begin{bmatrix} W^i R(0) \\ \bar{W}(i) A(i) \end{bmatrix} \quad (3.5)$$

$$b(i) = \begin{bmatrix} 0 \\ \bar{W}(i)(Y(i) - A(i)h(0)) \end{bmatrix} \quad (3.6)$$

e assumindo $h(i-1)$ solução do PMQRPE, no instante $(i-1)$, pode ser verificado por inspeção que podemos reescrever (3.3) da seguinte forma recursiva:

$$\min_{h(i)} J(i) = \left\| \begin{bmatrix} W\bar{X}(i-1) \\ a(i) \end{bmatrix} [h(i) - h(i-1)] - \begin{bmatrix} 0 \\ y(i) - a(i)h(i-1) \end{bmatrix} \right\|_2 \quad (3.7)$$

s.a

$$C[h(i) - h(i-1)] = 0$$

$$Ch(0) = d \quad (3.8)$$

Para a solução do problema (3.7), (3.8) podemos usar os seguintes métodos:

- métodos utilizando espaço nulo de C combinado com equações normais ou fatoração ortogonal
- métodos utilizando multiplicador de Lagrange combinado com equações normais ou fatoração ortogonal.

Em seguida descreveremos com detalhes os algoritmos recursivos para a solução das alternativas mencionadas.

3.3 Métodos utilizando o Espaço Nulo de C ($\mathcal{N}_{(C)}$)

Seja

$$Z = \mathcal{N}_{(C)} \quad (3.9)$$

onde: Z é uma matriz ortogonal de dimensões $(n, n-r)$ correspondente a uma base do espaço nulo de C .

Considerando o problema de minimização na forma matricial (3.7), (3.8) e assumindo a seguinte mudança de variáveis:

$$h(i) - h(i-1) = Zp(i) \quad (3.10)$$

onde: p é um vetor de dimensões $(n-r, 1)$, a restrição

$$C[h(i) - h(i-1)] = 0 \quad (3.11)$$

é automaticamente satisfeita para qualquer p . Assim sendo, temos o seguinte problema de minimização irrestrito:

$$\min_{p(i)} J(i) = \left\| \begin{bmatrix} W\bar{X}(i-1)Z \\ a(i)Z \end{bmatrix} p(i) - \begin{bmatrix} 0 \\ y(i) - a(i)h(i-1) \end{bmatrix} \right\|_2 \quad (3.12)$$

Obtida a solução de (3.12) em (3.10) temos que a solução do PMQRPE é dada por:

$$h(i) = h(i-1) + Zp(i) \quad (3.13)$$

De (3.5), (3.7), é fácil verificar que a precisão da solução do PMQRPE sem considerar as restrições (3.8) depende de $\text{cond}[\bar{X}(i)]$, o número condição da matriz $\bar{X}(i)$ [15].

De (3.5), (3.12), é fácil verificar que a precisão da solução do PMQRPE com espaço nulo Z (3.12) depende de $\text{cond}[\bar{X}(i)Z]$. Como Z é uma matriz ortogonal com $\text{posto}(Z) \leq \text{posto}(\bar{X})$, temos [15]

$$\text{cond}[\bar{X}(i)Z] \leq \text{cond}[\bar{X}(i)] \quad (3.14)$$

Portanto, o uso do espaço nulo Z melhora o condicionamento original do PMQRPE sem restrições.

Para a obtenção da solução recursiva do PMQRPE, basta a solução recursiva de (3.12), que é um problema de mínimos quadrados sem restrições, incorporando (3.13) no processo recursivo. A seguir apresentaremos sumariamente esses métodos.

3.3.1 Equações Normais

Utilizando as equações normais [15], a solução do problema de mínimos quadrados (3.12) com (3.13) é dada pela solução da seguinte equação :

$$[W^2 Z^T \bar{X}^T(i-1) \bar{X}(i-1) Z + Z^T a^T(i) a(i) Z] p(i) = Z^T a^T(i) [y(i) - a(i) h(i-1)] \quad (3.15)$$

$$h(i) = h(i-1) + Z p(i) \quad (3.16)$$

definindo

$$S(i-1) = [W^2 Z^T \bar{X}^T(i-1) \bar{X}(i-1) Z]^{-1} \quad (3.17)$$

$$S(i) = [W^2 Z^T \bar{X}^T(i-1) \bar{X}(i-1) Z + Z^T a^T(i) a(i) Z]^{-1} \quad (3.18)$$

Pelo Lema de Inversão de Matrizes no Apêndice A.1 temos

$$S(i) = \frac{1}{W^2} \left[S(i-1) - \frac{S(i-1) Z^T a^T(i) a(i) Z S(i-1)}{W^2 + a(i) Z S(i-1) Z^T a^T(i)} \right] \quad (3.19)$$

Definindo

$$v(i) = S(i-1) Z^T a^T(i) \quad (3.20)$$

$$g(i) = \frac{v(i)}{W^2 + a(i) Z v(i)} \quad (3.21)$$

temos então :

$$S(i) = \frac{1}{W^2} [S(i-1) - g(i) v^T(i)] \quad (3.22)$$

De (3.15), temos:

$$p(i) = S(i) Z^T a^T(i) e(i) \quad (3.23)$$

onde

$$e(i) = y(i) - a(i) h(i-1) \quad (3.24)$$

Substituindo (3.22) em (3.23), temos:

$$p(i) = \frac{1}{W^2} [S(i-1) - g(i)v^T(i)] Z^T a(i)^T e(i) \quad (3.25)$$

Substituindo (3.20) em (3.25), temos:

$$p(i) = \frac{1}{W^2} [v(i) - g(i)v^T(i)Z^T a(i)^T] e(i) \quad (3.26)$$

Substituindo (3.21) em (3.26), é fácil verificar que:

$$p(i) = g(i)e(i) \quad (3.27)$$

Desta forma, considerando as equações (3.13), (3.20), (3.21), (3.22) e (3.25), o cálculo recursivo do PMQRPE pode ser racionalizado através do algoritmo descrito na Tabela 1a, aqui denominado de APNZ.

Conforme pode ser visto na Tabela 1a, o esforço computacional requerido por iteração do algoritmo APNZ é devido principalmente às operações correspondentes às equações (3.31) a (3.35). Para essas equações, o número de FLOPS é dado na Tabela 1b.

De (3.15), (3.18) é fácil verificar que a precisão da solução fornecida pelo algoritmo APNZ depende de $\text{cond}[S(i)^{-1}]$. De (3.5), (3.7), (3.18), temos

$$S(i)^{-1} = Z^T \bar{X}^T(i) \bar{X}(i) Z \quad (3.28)$$

e conseqüentemente [15]

$$\text{cond}[S(i)^{-1}] = \text{cond}[\bar{X}(i)Z]^2 \quad (3.29)$$

TABELA 1a: APNZ - Algoritmo de Mínimos Quadrados utilizando Equações Normais, Ponderação Exponencial dos Dados com Espaço Nulo de C .

Inicialização

$$S(0) = (Z^T A_0 Z)^{-1} \quad (3.30)$$

Atualizações

- ganho de adaptação

$$v(i) = S(i-1)Z^T a^T(i) \quad (3.31)$$

$$g(i) = \frac{v(i)}{W^2 + a(i)Zv(i)} \quad (3.32)$$

- erro de estimação

$$e(i) = y(i) - a(i)h(i-1) \quad (3.33)$$

- parâmetros do sistema

$$h(i) = h(i-1) + Zg(i)e(i) \quad (3.34)$$

- matriz inversa de autocorrelação

$$S(i) = \frac{1}{W^2} [S(i-1) - g(i)v^T(i)] \quad (3.35)$$

TABELA 1b: Esforço computacional do APNZ

Equação	FLOPS
(3.31) :	$(n-r)^2 + n(n-r)$
(3.32) :	$2(n-r) + 1$
(3.33) :	n
(3.34) :	$n(n-r) + (n-r)$
(3.35) :	$2(n-r)^2$
Total :	$2n(n-r) + 3(n-r)^2 + 4n - 3r + 1$

3.3.2 Fatoração Ortogonal

Para o PMQRPE (3.12), utilizando a transformação ortogonal $Q(i-1)$, sem alterar o índice de desempenho $J(i)$, podemos colocá-lo na seguinte forma:

$$J(i) = \left\| \begin{bmatrix} Q(i-1) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} W\bar{X}(i-1) \\ a(i) \end{bmatrix} Zp(i) - \begin{bmatrix} Q(i-1) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ y(i) - a(i)h(i-1) \end{bmatrix} \right\|_2 \quad (3.36)$$

Assumindo $Q(i-1)$, tal que:

$$Q(i-1)\bar{X}(i-1)Z = \begin{bmatrix} R(i-1) \\ 0 \end{bmatrix} \quad (3.37)$$

onde $R(i-1)$ tem dimensões $(n-r, n-r)$ com posto completo, podemos transformar (3.36) em:

$$J(i) = \left\| \begin{bmatrix} WR(i-1) \\ a(i)Z \end{bmatrix} p(i) - \begin{bmatrix} 0 \\ y(i) - a(i)h(i-1) \end{bmatrix} \right\|_2 \quad (3.38)$$

Aplicando em (3.38) uma matriz ortogonal $Q(i)$ tal que:

$$Q(i) \begin{bmatrix} WR(i-1) \\ a(i)Z \end{bmatrix} = \begin{bmatrix} R(i) \\ 0 \end{bmatrix} \quad (3.39)$$

$$Q(i) \begin{bmatrix} 0 \\ y(i) - a(i)h(i-1) \end{bmatrix} = \begin{bmatrix} b_1(i) \\ b_2(i) \end{bmatrix} \quad (3.40)$$

e assumindo a equação (3.13), podemos colocar o PMQRPE na seguinte forma:

$$J(i) = \left\| \begin{bmatrix} R(i) \\ 0 \end{bmatrix} p(i) - \begin{bmatrix} b_1(i) \\ b_2(i) \end{bmatrix} \right\|_2 \quad (3.41)$$

$$h(i) = h(i-1) + Zp(i) \quad (3.42)$$

Devido a estrutura da matriz $\begin{bmatrix} R(i) \\ 0 \end{bmatrix}$ a solução de (3.41), (3.42) e consequentemente do PMQRPE, é obtida de:

$$R(i)p(i) = b_1(i) \quad (3.43)$$

$$h(i) = h(i-1) + Zp(i) \quad (3.44)$$

É importante notar que para a obtenção de $b_1(i)$ e $R(i)$ não é necessário construir $Q(i)$ explicitamente, bastando aplicar a rotação de Givens diretamente em relação a $\begin{bmatrix} WR(i-1) \\ a(i)Z \end{bmatrix}$ e $\begin{bmatrix} 0 \\ y(i) - a(i)h(i-1) \end{bmatrix}$.

Desta forma, considerando as equações (3.39), (3.40), (3.43), (3.44), o cálculo recursivo do PMQRPE pode ser racionalizado através do algoritmo descrito na Tabela 2a, aqui denominado de APQZ.

Conforme pode ser visto na Tabela 2a, o esforço computacional requerido por iteração do algoritmo APQZ é devido principalmente às operações correspondentes às equações (3.49) a (3.52). Para essas equações, o número de FLOPS é dado na Tabela 2b.

De (3.43) temos que a precisão da solução fornecida pelo algoritmo APQZ depende de $\text{cond}[R(i)]$. De (3.5), (3.12), (3.37), (3.39) temos:

$$Q\bar{X}(i)Z = \begin{bmatrix} R(i) \\ 0 \end{bmatrix} \quad (3.45)$$

onde Q é uma matriz ortogonal.

Consequentemente [15],

$$\text{cond}[R(i)] = \text{cond}[\bar{X}(i)Z] \quad (3.46)$$

TABELA 2a: APQZ - Algoritmo de Mínimos Quadrados utilizando Fatoração Ortogonal, Ponderação Exponencial dos Dados com Espaço Nulo de C .

Inicialização

- Obter $R(0)$ tal que

$$R^T(0)R(0) = Z^T A_0 Z \quad (3.47)$$

$$h(0) \text{ tal que } Ch(0) = d \quad (3.48)$$

Atualizações

- Determinar $R(i), b_1(i)$ e $b_2(i)$:

$$Q(i) \begin{bmatrix} W R(i-1) \\ a(i)Z \end{bmatrix} = \begin{bmatrix} R(i) \\ 0 \end{bmatrix} \quad (3.49)$$

$$Q(i) \begin{bmatrix} 0 \\ y(i) - a(i)h(i-1) \end{bmatrix} = \begin{bmatrix} b_1(i) \\ b_2(i) \end{bmatrix} \quad (3.50)$$

- Obter $p(i)$ da solução do sistema:

$$R(i)p(i) = b_1(i) \quad (3.51)$$

- Determinar parâmetros $h(i)$:

$$h(i) = h(i-1) + Zp(i) \quad (3.52)$$

Obs.: $Q(i)$ corresponde ao produto de uma sequência de matrizes de rotação de Givens (Apêndice A2). Sua construção explícita e armazenamento não é necessário.

TABELA 2b: Esforço computacional do APQZ

Equação	FLOPS
(3.49) :	$\frac{9}{2}(n-r)^2 + \frac{1}{2}(n-r)(2n-9)$
(3.50) :	$4(n-r)$
(3.51) :	$\frac{1}{2}(n-r)(n-r+1)$
(3.52) :	$n(n-r)$
Total :	$5(n-r)^2 + \frac{3}{2}(n-r)n$

3.4 Métodos utilizando Multiplicadores de Lagrange (λ)

Para solução do problema (3.7), (3.8) as equações básicas da técnica dos multiplicadores de Lagrange são:

$$\nabla J(i) + C^T \lambda(i) = 0 \quad (3.53)$$

$$C[h(i) - h(i-1)] = 0 \quad (3.54)$$

onde: $\lambda(i)$ é um vetor de dimensões $(r, 1)$ correspondente aos multiplicadores de Lagrange.

É bem conhecido [15] que:

$$\nabla J(i) = \bar{X}(i)^T \bar{X}(i)[h(i) - h(i-1)] - \bar{X}(i)^T b(i) \quad (3.55)$$

onde:

$$\bar{X}(i) = \begin{bmatrix} WX(i-1) \\ a(i) \end{bmatrix} \quad (3.56)$$

$$b(i) = \begin{bmatrix} 0 \\ y(i) - a(i)h(i-i) \end{bmatrix} \quad (3.57)$$

Substituindo (3.55) em (3.53), temos o seguinte sistema de equações :

$$\bar{X}(i)^T \bar{X}(i)[h(i) - h(i-1)] = \bar{X}(i)^T b(i) - C^T \lambda(i) \quad (3.58)$$

$$C[h(i) - h(i-1)] = 0 \quad (3.59)$$

Pode ser verificado, sem dificuldade, que podemos obter $\lambda(i)$ das equações acima da seguinte forma:

$$[C(\bar{X}(i)^T \bar{X}(i))^{-1} C^T] \lambda(i) = C(\bar{X}(i)^T \bar{X}(i))^{-1} \bar{X}(i)^T b(i) \quad (3.60)$$

Tendo $\lambda(i)$, de (3.58), podemos obter $h(i)$ da seguinte forma:

$$h(i) = h(i-1) + (\bar{X}(i)^T \bar{X}(i))^{-1} \bar{X}(i)^T b(i) - (\bar{X}(i)^T \bar{X}(i))^{-1} C^T \lambda(i) \quad (3.61)$$

notando que [15]:

$$(\overline{X}(i)^T \overline{X}(i))^{-1} \overline{X}(i)^T b(i) = \hat{h}(i) - \hat{h}(i-1) = \hat{d}_h(i) \quad (3.62)$$

onde $\hat{h}(i)$ é a solução do PMQRPE (3.7) sem considerar a restrição (3.8). E assumindo

$$Y(i) = (\overline{X}(i)^T \overline{X}(i))^{-1} C^T \quad (3.63)$$

Podemos reescrever (3.60) e (3.61) da seguinte forma:

$$[CY(i)]\lambda(i) = C\hat{d}_h(i) \quad (3.64)$$

$$h(i) = h(i-1) + \hat{d}_h(i) - Y(i)\lambda(i) \quad (3.65)$$

Das equações (3.64), (3.65), vemos que para obter recursivamente $h(i)$, devemos calcular recursivamente $\hat{d}_h(i)$ e $Y(i)$.

A seguir vamos mostrar como obter este cálculo recursivo utilizando equações normais e fatoração ortogonal.

3.4.1 Equações Normais

Definindo

$$S(i-1) = [W^2 \overline{X}^T(i-1) \overline{X}(i-1)]^{-1} \quad (3.66)$$

$$S(i) = [\overline{X}(i)^T \overline{X}(i-1)]^{-1} = [W^2 \overline{X}^T(i-1) \overline{X}(i-1) + a^T(i)a(i)]^{-1} \quad (3.67)$$

$$v(i) = S(i-1)a^T(i) \quad (3.68)$$

$$g(i) = \frac{v(i)}{W^2 + a(i)v(i)} \quad (3.69)$$

$$e(i) = y(i) - a(i)h(i-1) \quad (3.70)$$

Utilizando Lema de Inversão de Matrizes (Apêndice A.1) e procedimento análogo ao utilizado na seção 3.3.1, pode ser verificado que:

$$S(i) = \frac{1}{W^2} [S(i-1) - g(i)v^T(i)] \quad (3.71)$$

$$\hat{d}_h(i) = g(i)e(i) \quad (3.72)$$

De (3.63), (3.67) e (3.71), temos:

$$Y(i) = \frac{1}{W^2}[S(i-1) - g(i)v^T(i)]C^T \quad (3.73)$$

consequentemente

$$Y(i) = \frac{1}{W^2}[Y(i-1) - g(i)u^T(i)] \quad (3.74)$$

onde

$$u(i) = Cv(i) \quad (3.75)$$

Dessa forma, considerando as equações (3.64), (3.65), (3.71), (3.72), (3.74), (3.75), podemos racionalizar o cálculo recursivo do PMQRPE com multiplicador de Lagrange, através do algoritmo descrito na Tabela 3a, aqui denominado de APNL.

Conforme pode ser visto na Tabela 3a, o esforço computacional requerido por iteração do algoritmo APNL é devido principalmente às operações correspondentes às equações (3.81) a (3.89). Para essas equações, o número de FLOPS é dado na Tabela 3b.

Para o algoritmo APNL, temos:

- de (3.62), (3.63), (3.67), temos que a precisão do cálculo de $\hat{d}_h(i)$ e $Y(i)$ dependem de

$$\text{cond}[S(i)^{-1}] = \text{cond}[\bar{X}(i)^T \bar{X}(i)] = \text{cond}[\bar{X}(i)]^2 \quad (3.76)$$

- de (3.63), (3.64), temos que a precisão do cálculo de $\lambda(i)$ depende de

$$\text{cond}[CS(i)C^T] \quad (3.77)$$

TABELA 3a: APNL - Algoritmo de Mínimos Quadrados utilizando Equações Normais, Ponderação Exponencial dos Dados com Multiplicadores de Lagrange.

Inicialização

$$S(0) = (A_0)^{-1} \quad (3.78)$$

$$Y(0) = S(0)C^T \quad (3.79)$$

$$h(0) \quad (3.80)$$

Atualizações

$$v(i) = S(i-1)a^T(i) \quad (3.81)$$

$$u(i) = Cv(i) \quad (3.82)$$

$$g(i) = \frac{v(i)}{W^2 + a(i)v(i)} \quad (3.83)$$

$$e(i) = y(i) - a(i)h(i-1) \quad (3.84)$$

$$\hat{d}_h(i) = g(i)e(i) \quad (3.85)$$

$$Y(i) = \frac{1}{W^2} [Y(i-1) - g(i)u^T(i)] \quad (3.86)$$

Resolve o sistema

$$CY(i)\lambda(i) = C\hat{d}_h(i) \quad (3.87)$$

$$h(i) = h(i-1) + \hat{d}_h(i) - Y(i)\lambda(i) \quad (3.88)$$

Matriz Inversa de Autocorrelação

$$S(i) = \frac{1}{W^2} [S(i-1) - g(i)v^T(i)] \quad (3.89)$$

TABELA 3b: Esforço computacional do APNL

Equação	FLOPS
(3.81) :	n^2
(3.82) :	nr
(3.83) :	$2n + 1$
(3.84) :	n
(3.85) :	n
(3.86) :	$2nr$
(3.87) :	$\frac{1}{6}r^3 + nr^2 + nr$
(3.88) :	nr
(3.89) :	$2n^2$
Total :	$\frac{1}{6}r^3 + nr^2 + 3n^2 + 5nr + 4n + 1$

3.4.2 Fatoração Ortogonal

Assumindo uma transformação ortogonal $Q(i)$, tal que:

$$Q(i)\overline{X}(i) = Q(i) \begin{bmatrix} W\overline{X}(i-1) \\ a(i) \end{bmatrix} = \begin{bmatrix} R(i) \\ 0 \end{bmatrix} \quad (3.90)$$

$$Q(i) \begin{bmatrix} 0 \\ y(i) - a(i)h(i-1) \end{bmatrix} = \begin{bmatrix} b_1(i) \\ b_2(i) \end{bmatrix} \quad (3.91)$$

podemos colocar o PMQRPE (3.7) sem considerar as restrições (3.8), na seguinte forma:

$$\min_{h(i)} J(i) = \left\| \begin{bmatrix} R(i) \\ 0 \end{bmatrix} [\hat{h}(i) - \hat{h}(i-1)] - \begin{bmatrix} b_1(i) \\ b_2(i) \end{bmatrix} \right\|_2 \quad (3.92)$$

É fácil verificar, pela estrutura das matrizes envolvidas, que:

$$\hat{d}_h(i) = [\hat{h}(i) - \hat{h}(i-1)] = R(i)^{-1}b_1(i) \quad (3.93)$$

Substituindo (3.63), (3.90), (3.93), em (3.64), temos:

$$[CR(i)^{-1}][(R(i)^T)^{-1}C^T]\lambda = C[R(i)^{-1}]b_1(i) \quad (3.94)$$

É fácil verificar que (3.94) é a equação normal do seguinte problema de mínimos quadrados:

$$\|\bar{Y}(i)\lambda(i) - b_1(i)\|_2 \quad (3.95)$$

onde $\bar{Y}(i)$ é solução do seguinte sistema de equações

$$R(i)^T \bar{Y}(i) = C^T \quad (3.96)$$

Assim sendo, para obter $\lambda(i)$, basta resolver o problema de mínimos quadrados (3.95), o qual pode ser feito utilizando transformações ortogonais.

Substituindo (3.63), (3.90), (3.93) em (3.65), temos:

$$h(i) = h(i-1) + R(i)^{-1}b_1(i) - R(i)^{-1}\bar{Y}(i)\lambda(i) \quad (3.97)$$

Portanto

$$h(i) = h(i-1) + dh \quad (3.98)$$

onde dh é a solução do seguinte sistema de equações

$$R(i)dh = b_1(i) - \bar{Y}(i)\lambda(i) \quad (3.99)$$

Dessa forma, considerando as equações (3.90), (3.91), (3.95), (3.96), (3.98) e (3.99), podemos racionalizar o cálculo recursivo do PMQRPE com multiplicador de Lagrange, através do algoritmo descrito na Tabela 4a, aqui denominado de APQL.

Conforme pode ser visto na Tabela 4a, o esforço computacional requerido por iteração do algoritmo APQL é devido principalmente às operações correspondentes às equações (3.103) a (3.107). Para essas equações, o número de FLOPS é dado na Tabela 4b.

De (3.56), (3.90), (3.92), (3.96), temos que a precisão do cálculo de $\hat{d}_h(i)$ e $Y(i)$ dependem de

$$\text{cond}[R(i)] = \text{cond}[\bar{X}(i)] \quad (3.100)$$

De (3.95), temos que o cálculo de $\lambda(i)$ depende de

$$\text{cond}[\bar{Y}(i)] = \text{cond}[(R(i)^{-1})^T C^T] \quad (3.101)$$

TABELA 4a: APQL - Algoritmo de Mínimos Quadrados utilizando Fatoração Ortogonal, Ponderação Exponencial dos Dados com Multiplicadores de Lagrange.

Inicialização

$$R(0), h(0) \quad (3.102)$$

Atualizações

- Determinar $R(i)$, $b_1(i)$ e $b_2(i)$:

$$Q(i) \begin{bmatrix} WR(i-1) \\ a(i) \end{bmatrix} = \begin{bmatrix} R(i) \\ 0 \end{bmatrix} \quad (3.103)$$

$$Q(i) \begin{bmatrix} 0 \\ y(i) - a(i)h(i-1) \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad (3.104)$$

- Determinar $\bar{Y}(i)$, resolvendo o seguinte sistema de equações triangular

$$R(i)^T \bar{Y}(i) = C^T \quad (3.105)$$

- Determinar $\lambda(i)$, resolvendo o seguinte problema de mínimos quadrados

$$\|\bar{Y}(i)\lambda(i) - b_1(i)\|_2 \quad (3.106)$$

- Determinar dh , resolvendo o seguinte sistema de equações triangular

$$R(i)dh = b_1(i) - \bar{Y}(i)\lambda(i) \quad (3.107)$$

- Determinar os parâmetros $h(i)$:

$$h(i) = h(i-1) + dh \quad (3.108)$$

TABELA 4b: Esforço computacional do **APQL**

Equação	FLOPS
(3.103) :	$\frac{5}{2}n^2 + \frac{15}{2}n$
(3.104) :	$2n$
(3.105) :	$nr(n+1)/2$
(3.106) :	$\frac{1}{6}r^3 + r^2n + rn$
(3.107) :	$nr + n(n+1)/2$
Total :	$\frac{1}{6}r^3 + r^2n + \frac{1}{2}n^2(r+6) + \frac{5}{2}rn + 10n$

3.5 Análise Comparativa

Para comparar o esforço computacional por passo requerido pelos vários métodos, vamos considerar um problema com as seguintes características:

$$n = 10; \quad r = 2$$

Das Tabelas **1b**, **2b**, **3b**, **4b**, temos a seguinte estimativa de FLOPS:

- APNZ: 387
- APQZ: 440
- APNL: 482
- APQL: 591

Desta forma, podemos concluir que:

- os algoritmos utilizando o espaço nulo Z requerem menor esforço computacional que seus similares utilizando multiplicadores de Lagrange λ .
- os algoritmos utilizando equações normais requerem menor esforço computacional que seus similares utilizando fatoração QR.

De (3.29), (3.46), (3.76), (3.100), temos que a precisão dos resultados obtidos dependem principalmente de:

- APNZ: $\text{cond}[\overline{X}(i)Z]^2$
- APQZ: $\text{cond}[\overline{X}(i)Z]$
- APNL: $\text{cond}[\overline{X}(i)]^2$
- APQL: $\text{cond}[\overline{X}(i)]$

Desta forma, podemos concluir que:

- O uso de fatoração QR não altera o condicionamento original do PMQRPE e os algoritmos que a utilizam possuem condicionamento numérico muito superior aos seus similares utilizando equações normais. O uso desta última não é recomendado para problemas inerentemente mal condicionados ($\text{cond}[\overline{X}(i)]$ elevado).
- O condicionamento numérico dos algoritmos utilizando o espaço nulo Z é superior ao de seus similares utilizando multiplicadores de Lagrange λ .

Capítulo 4

Estimação Adaptativa de Parâmetros com Restrições de Igualdade utilizando Janela Móvel de Dados

4.1 Introdução

Neste capítulo, vamos apresentar algoritmos recursivos para a solução do problema de mínimos quadrados lineares com janela móvel de dados sujeito a restrições de igualdade, utilizando métodos baseados no espaço nulo de C e multiplicadores de Lagrange, combinadas com equações normais e fatoração ortogonal.

4.2 Definição do Problema

Com a finalidade de seguir as não estacionaridades de um processo, podemos considerar o PMQR (2.3) restrito a uma janela móvel de dados de comprimento ld . Considere o problema de estimar um vetor de parâmetros h de dimensões $(n, 1)$ solução da equação de regressão linear dado em (2.1). Assumindo h constante em $i - ld + 1 \leq k \leq i$, a estimação por mínimos quadrados de h no instante i , corresponde

à solução do seguinte problema:

$$\min_{h(i)} J(i) = [h(i) - h(0)]^T A_0 [h(i) - h(0)] + \sum_{k=i-l_d+1}^i [y(k) - a(k)h(i)]^2 \quad (4.1)$$

s.a

$$Ch(0) = d$$

$$Ch(i) = d \quad (4.2)$$

onde : $A_0 = R(0)^T R(0)$ é uma matriz definida positiva representando o peso da estimativa a priori $h(0)$ relativamente aos dados da janela corrente, e $y(k) = h(k) = 0$, para $k < 0$.

Sendo assim, podemos definir o problema de mínimos quadrados não recursivo sujeito a restrições com janela móvel de dados PMQRJM na seguinte forma matricial:

$$\min_{h(i)} J(i) = \left\| \begin{bmatrix} R(0) \\ A(i) \end{bmatrix} [h(i) - h(0)] - \begin{bmatrix} 0 \\ Y(i) - A(i)h(0) \end{bmatrix} \right\|_2 \quad (4.3)$$

s.a

$$Ch(i) = d \quad (4.4)$$

onde

$$A(i) = \begin{bmatrix} a(i-l_d+1) \\ \vdots \\ a(i) \end{bmatrix} \text{ matriz de dados de entrada de dimensões } (l_d, n)$$

$$Y(i) = \begin{bmatrix} y(i-l_d+1) \\ \vdots \\ y(i) \end{bmatrix} \text{ vetor dos sinais de referência de dimensões } (l_d, 1)$$

De (4.1), podemos verificar que:

$$J(i) = J(i-1) + [y(i) - a(i)h(i)]^2 - [y(i-l_d) - a(i-l_d)h(i)]^2 \quad (4.5)$$

Fazendo

$$X(i-1) = \begin{bmatrix} R(0) \\ A(i-1) \end{bmatrix} \quad (4.6)$$

e assumindo $h(i-1)$ a solução de PMQRJM no instante $(i-1)$, pode ser verificado por inspeção que podemos reescrever (4.3) recursivamente da seguinte forma:

$$\min_{h(i)} J(i) = \left\| \begin{bmatrix} X(i-1) \\ a(i) \\ ja(i-l d) \end{bmatrix} [h(i) - h(i-1)] - \begin{bmatrix} 0 \\ y(i) - a(i)h(i-1) \\ j[y(i-l d) - a(i-l d)h(i-1)] \end{bmatrix} \right\|_2 \quad (4.7)$$

s.a

$$Ch(0) = d$$

$$C[h(i) - h(i-1)] = 0 \quad (4.8)$$

onde $j = \sqrt{-1}$.

Para a solução do problema (4.7), (4.8) podemos usar os seguintes métodos:

- métodos utilizando espaço nulo de C combinado com equações normais ou fatoração ortogonal
- métodos utilizando multiplicador de Lagrange combinado com equações normais ou fatoração ortogonal.

Em seguida descreveremos com detalhes os algoritmos recursivos para a solução das alternativas mencionadas.

4.3 Métodos utilizando o Espaço Nulo de C ($\mathcal{N}_{(C)}$)

Seja

$$Z = \mathcal{N}_{(C)} \quad (4.9)$$

onde: Z é uma matriz ortogonal de dimensões $(n, n-r)$ correspondente a uma base do espaço nulo de C .

Considerando o problema de minimização na forma matricial (4.7), (4.8) e assumindo a seguinte mudança de variáveis:

$$h(i) - h(i-1) = Zp(i) \quad (4.10)$$

onde: p é um vetor de dimensões $(n - r, 1)$, a restrição

$$C[h(i) - h(i - 1)] = 0 \quad (4.11)$$

é automaticamente satisfeita para qualquer p . Assim sendo, temos o seguinte problema de minimização irrestrito:

$$\min_{p(i)} J(i) = \left\| \begin{bmatrix} X(i-1) \\ a(i) \\ ja(i-l d) \end{bmatrix} Zp(i) - \begin{bmatrix} 0 \\ y(i) - a(i)h(i-1) \\ j[y(i-l d) - a(i-l d)h(i-1)] \end{bmatrix} \right\|_2 \quad (4.12)$$

onde $j = \sqrt{-1}$.

Obtida a solução de (4.12), em (4.10) temos que a solução do PMQRJM é dada por:

$$h(i) = h(i-1) + Zp(i) \quad (4.13)$$

Desta forma, para a obtenção da solução recursiva do PMQRJM, basta obter a solução recursiva de (4.12), que é um problema de mínimos quadrados sem restrições, incorporando (4.13) no processo recursivo. A seguir apresentaremos sumariamente esses métodos.

4.3.1 Equações Normais

Utilizando as equações normais [15], a solução do problema de mínimos quadrados (4.12) com (4.13) é dada pela solução da seguinte equação :

$$\begin{aligned} & [Z^T X^T(i-1)X(i-1)Z + Z^T a^T(i)a(i)Z - Z^T a^T(i-l d)a(i-l d)Z]p(i) = \\ & Z^T a^T(i)[y(i) - a(i)h(i-1)] - Z^T a^T(i-l d)[y(i-l d) - a(i-l d)h(i-1)] \end{aligned} \quad (4.14)$$

$$h(i) = h(i-1) + Zp(i) \quad (4.15)$$

Para resolver recursivamente a equação (4.14), (4.15), podemos dividir o problema em duas etapas :

$$[ZX^T(i-1)X(i-1)Z + Za^T(i)a(i)Z]\tilde{p}(i) = Z^T a^T(i)[y(i) - a(i)h(i-1)] \quad (4.16)$$

$$\tilde{h}(i) = h(i-1) + Z\tilde{p}(i) \quad (4.17)$$

$$\begin{aligned} [Z^T \tilde{X}^T(i) \tilde{X}(i) Z - Z^T a^T(i-l_d) a(i-l_d) Z] p(i) = \\ - Z^T a^T(i-l_d) [y(i-l_d) - a(i-l_d) \tilde{h}(i)] \end{aligned} \quad (4.18)$$

$$h(i) = \tilde{h}(i) + Zp(i) \quad (4.19)$$

onde $\tilde{X}(i) = \begin{bmatrix} X(i-1) \\ a(i) \end{bmatrix}$

Para a solução da 1a. etapa, podemos utilizar do seguinte procedimento:
Definindo

$$S(i-1) = [Z^T X^T(i-1) X(i-1) Z]^{-1} \quad (4.20)$$

$$\tilde{S}(i) = [Z^T X^T(i-1) X(i-1) Z + Z^T a^T(i) a(i) Z]^{-1} \quad (4.21)$$

Pelo Lema de Inversão de Matrizes no Apêndice A.1 temos

$$\tilde{S}(i) = \left[S(i-1) - \frac{S(i-1) Z^T a^T(i) a(i) Z S(i-1)}{1 + a(i) Z S(i-1) Z^T a^T(i)} \right] \quad (4.22)$$

Definindo

$$\tilde{v}(i) = S(i-1) Z^T a^T(i) \quad (4.23)$$

$$\tilde{g}(i) = \frac{\tilde{v}(i)}{1 + a(i) Z \tilde{v}(i)} \quad (4.24)$$

devemos ter, analogamente ao desenvolvimento visto na seção 3.3.1:

$$\tilde{S}(i) = S(i-1) - \tilde{g}(i) \tilde{v}^T(i) \quad (4.25)$$

$$\tilde{p}(i) = \tilde{g}(i) \tilde{e}(i) \quad (4.26)$$

e conseqüentemente de (4.17), temos:

$$\tilde{h}(i) = h(i-1) + Z\tilde{g}(i) \tilde{e}(i) \quad (4.27)$$

Para a solução da 2a. etapa, temos o seguinte:

Definindo

$$S(i) = [Z^T \widetilde{X}^T(i) \widetilde{X}(i) Z - Z^T a^T(i - ld) a(i - ld) Z] \quad (4.28)$$

$$v(i) = \widetilde{S}(i) Z^T a^T(i - ld) \quad (4.29)$$

$$g(i) = \frac{v(i)}{1 - a(i) Z v(i)} \quad (4.30)$$

devemos ter, utilizando o Lema de Inversão de Matrizes no Apêndice A.1 e analogamente ao desenvolvido na seção 3.3.1:

$$S(i) = \widetilde{S}(i) + g(i) v^T(i) \quad (4.31)$$

$$p(i) = -g(i) e(i) \quad (4.32)$$

$$h(i) = \widetilde{h}(i) + Z g(i) e(i) \quad (4.33)$$

Desta forma, considerando as equações (4.20), (4.23), (4.24), (4.25), (4.27), (4.31), (4.33), o cálculo recursivo do PMQRJM, pode ser racionalizado através do algoritmo descrito na Tabela 5a, aqui denominado de AJNZ.

Conforme pode ser visto na Tabela 5a, o esforço computacional requerido por iteração do algoritmo AJNZ é devido principalmente às operações correspondentes às equações (4.38) a (4.47). Para essas equações, o número de FLOPS é dado na Tabela 5b.

De (4.14), (4.16), (4.21), (4.28) é fácil verificar que a precisão da solução fornecida pelo algoritmo AJNZ depende de $\text{cond}[S(i)^{-1}]$. De (4.6), (4.7), (4.28), temos

$$S(i)^{-1} = Z^T X^T(i) X(i) Z \quad (4.34)$$

e consequentemente [15]

$$\text{cond}[S(i)^{-1}] = \text{cond}[X(i) Z]^2 \quad (4.35)$$

TABELA 5a: AJNZ - Algoritmo de Mínimos Quadrados utilizando Equações Normais, Janela Móvel de Dados com Espaço Nulo de C .

Inicialização

$$S(0) = (Z^T A_0 Z)^{-1} \quad (4.36)$$

$$h(0) \text{ tal que } Ch(0) = d \quad (4.37)$$

Atualizações

Para $i=1$ até final dos dados faça

- ganho de adaptação

$$v(i) = S(i-1)Z^T a^T(i) \quad (4.38)$$

$$g(i) = \frac{v(i)}{1 + a(i)Zv(i)} \quad (4.39)$$

- erro de estimação

$$e(i) = y(i) - a(i)h(i-1) \quad (4.40)$$

- parâmetros do sistema

$$h(i) = h(i-1) + Zg(i)e(i) \quad (4.41)$$

- matriz inversa de autocorrelação

$$S(i) = S(i-1) - g(i)v^T(i) \quad (4.42)$$

- Caso $i > ld$

—ganho de adaptação

$$v(i) = S(i)Z^T a^T(i - ld) \quad (4.43)$$

$$g(i) = \frac{v(i)}{1 - a(i - ld)Zv(i)} \quad (4.44)$$

— erro de estimação

$$e(i) = y(i - ld) - a(i - ld)h(i) \quad (4.45)$$

— parâmetros do sistema

$$h(i) = h(i) - Zg(i)e(i) \quad (4.46)$$

— matriz inversa de autocorrelação

$$S(i) = S(i) + g(i)v^T(i) \quad (4.47)$$

TABELA 5b: Esforço computacional do **AJNZ**

Equação	FLOPS
(4.38) e (4.43) :	$2(n - r)^2 + 2n(n - r)$
(4.39) e (4.44) :	$4(n - r)$
(4.40) e (4.45) :	$2n$
(4.41) e (4.46) :	$2n(n - r) + 2(n - r)$
(4.42) e (4.47) :	$2(n - r)^2$
Total :	$4n(n - r) + 4(n - r)^2 + 8n - 6r$

4.3.2 Fatoração Ortogonal

Para o PMQRJM (4.12), (4.13), utilizando a transformação ortogonal $\check{Q}$, sem alterar o índice de desempenho $J(i)$, podemos colocá-lo na seguinte forma:

$$J(i) = \left\| \begin{bmatrix} \check{Q} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X(i - 1) \\ a(i) \\ ja(i - ld) \end{bmatrix} Zp(i) - \begin{bmatrix} \check{Q} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ y(i) - a(i)h(i - 1) \\ j[y(i - ld) - a(i - ld)h(i - 1)] \end{bmatrix} \right\|_2 \quad (4.48)$$

Assumindo $\check{Q}$, tal que:

$$\check{Q}X(i-1)Z = \begin{bmatrix} R(i-1) \\ 0 \end{bmatrix} \quad (4.49)$$

onde $R(i-1)$ tem dimensões $(n-r, n-r)$ com posto completo, podemos transformar (4.48) em:

$$J(i) = \left\| \begin{bmatrix} R(i-1) \\ a(i)Z \\ ja(i-l d)Z \end{bmatrix} p(i) - \begin{bmatrix} 0 \\ y(i) - a(i)h(i-1) \\ j[y(i-l d) - a(i-l d)h(i-1)] \end{bmatrix} \right\|_2 \quad (4.50)$$

Aplicando em (4.50) uma matriz ortogonal

$$\begin{bmatrix} \tilde{Q} & 0 \\ 0 & 1 \end{bmatrix}$$

tal que:

$$\tilde{Q} \begin{bmatrix} R(i-1) \\ a(i)Z \end{bmatrix} = \begin{bmatrix} \tilde{R}(i) \\ 0 \end{bmatrix} \quad (4.51)$$

$$\tilde{Q} \begin{bmatrix} 0 \\ y(i) - a(i)h(i-1) \end{bmatrix} = \begin{bmatrix} \tilde{b}_1(i) \\ \tilde{b}_2(i) \end{bmatrix} \quad (4.52)$$

onde $\tilde{R}(i)$ tem dimensões $(n-r, n-r)$ com posto completo e $\tilde{b}_1(i)$ tem dimensões $(n-r, 1)$, podemos colocar o PMQRJM na seguinte forma:

$$J(i) = \left\| \begin{bmatrix} \tilde{R}(i) \\ 0 \\ ja(i-l d)Z \end{bmatrix} p(i) - \begin{bmatrix} \tilde{b}_1(i) \\ \tilde{b}_2(i) \\ j[y(i-l d) - a(i-l d)h(i-1)] \end{bmatrix} \right\|_2 \quad (4.53)$$

Aplicando em (4.53) uma matriz ortogonal Q tal que:

$$Q \begin{bmatrix} \tilde{R}(i) \\ ja(i-l d)Z \end{bmatrix} = \begin{bmatrix} R(i) \\ 0 \end{bmatrix} \quad (4.54)$$

$$Q \begin{bmatrix} \tilde{b}_1(i) \\ j[y(i-l d) - a(i-l d)h(i-1)] \end{bmatrix} = \begin{bmatrix} b_1(i) \\ jb_2(i) \end{bmatrix} \quad (4.55)$$

onde $R(i)$ tem dimensões $(n-r, n-r)$ com posto completo e $b_1(i)$ tem dimensões $(n-r, 1)$, podemos colocar o PMQRJM na seguinte forma:

$$J(i) = \left\| \begin{bmatrix} R(i) \\ 0 \end{bmatrix} p(i) - \begin{bmatrix} b_1(i) \\ j b_2(i) \end{bmatrix} \right\|_2 \quad (4.56)$$

$$h(i) = h(i-1) + Zp(i) \quad (4.57)$$

Devido a estrutura da matriz $\begin{bmatrix} R(i) \\ 0 \end{bmatrix}$ a solução de (4.56), (4.57) e consequentemente do PMQRJM, é obtida de:

$$R(i)p(i) = b_1(i) \quad (4.58)$$

$$h(i) = h(i-1) + Zp(i) \quad (4.59)$$

Dessa forma, considerando as equações (4.51), (4.52), (4.54), (4.55), (4.58) e (4.59), podemos racionalizar o cálculo recursivo do PMQRJM com espaço nulo, através do algoritmo descrito na Tabela 6a, aqui denominado de AJQZ.

Conforme pode ser visto na Tabela 6a, o esforço computacional requerido por iteração do algoritmo AJQZ é devido principalmente às operações correspondentes às equações (4.64) a (4.68). Para essas equações, o número de FLOPS é dado na Tabela 6b.

De (4.58) temos que a precisão da solução fornecida pelo algoritmo AJQZ depende de $\text{cond}[R(i)]$. De (4.6), (4.12), (4.50), (4.51), (4.54) temos

$$Q_{JM}X(i)Z = \begin{bmatrix} R(i) \\ 0 \end{bmatrix} \quad (4.60)$$

onde Q_{JM} é uma matriz ortogonal.

Consequentemente [15],

$$\text{cond}[R(i)] = \text{cond}[X(i)Z] \quad (4.61)$$

ou seja, o condicionamento original do problema (4.12) não é alterado.

TABELA 6a: AJQZ - Algoritmo de Mínimos Quadrados utilizando Fatoração Ortogonal, Janela Móvel dos Dados com Espaço Nulo de C .

Inicialização

- Obter $R(0)$ tal que

$$R^T(0)R(0) = Z^T A_0 Z \quad (4.62)$$

$$h(0) \text{ tal que } Ch(0) = d \quad (4.63)$$

Atualizações

Para $i=1$ até final dos dados faça

- Determinar $R(i), b_1(i)$ e $b_2(i)$:

$$\tilde{Q} \begin{bmatrix} R(i-1) \\ a(i)Z \end{bmatrix} = \begin{bmatrix} R(i) \\ 0 \end{bmatrix} \quad (4.64)$$

$$\tilde{Q} \begin{bmatrix} 0 \\ y(i) - a(i)h(i-1) \end{bmatrix} = \begin{bmatrix} b_1(i) \\ b_2(i) \end{bmatrix} \quad (4.65)$$

- Se $i > ld$

— Determinar $R(i), b_1(i)$ e $b_2(i)$:

$$Q \begin{bmatrix} R(i) \\ ja(i-ld)Z \end{bmatrix} = \begin{bmatrix} R(i) \\ 0 \end{bmatrix} \quad (4.66)$$

$$Q \begin{bmatrix} b_1(i) \\ j[y(i-ld) - a(i-ld)h(i-1)] \end{bmatrix} = \begin{bmatrix} b_1(i) \\ b_2(i) \end{bmatrix} \quad (4.67)$$

- Obter $p(i)$ da solução do sistema:

$$R(i)p(i) = b_1(i) \quad (4.68)$$

- Determinar parâmetros $h(i)$:

$$h(i) = h(i-1) + Zp(i) \quad (4.69)$$

É importante notar que as transformações ortogonais $\tilde{Q}$ e Q podem ser obtidas de forma eficiente e numericamente estável, através de uma sequência de transformações de Givens, conforme apresentado na Apêndice A.2. A construção explícita de $\tilde{Q}$ e Q não é necessária.

TABELA 6b: Esforço computacional do **AJQZ**

Equação	FLOPS
(4.64) e (4.66) :	$8(n-r)^2 + 2(n-r)(n-4)$
(4.65) e (4.67) :	$8(n-r)$
(4.68) :	$\frac{1}{2}(n-r)(n-r+1)$
(4.69) :	$n(n-r)$
Total :	$\frac{17}{2}(n-r)^2 + (n-r)(3n + \frac{1}{2})$

4.4 Métodos utilizando Multiplicadores de Lagrange (λ)

Para solução do problema (4.7), (4.8) as equações básicas da técnica dos multiplicadores de Lagrange são:

$$\nabla J(i) + C^T \lambda(i) = 0 \quad (4.70)$$

$$C[h(i) - h(i-1)] = 0 \quad (4.71)$$

onde: $\lambda(i)$ é um vetor de dimensões $(r, 1)$ correspondente aos multiplicadores de Lagrange.

É bem conhecido [15] que:

$$\nabla J(i) = X(i)^T X(i)[h(i) - h(i-1)] - X(i)^T b(i) \quad (4.72)$$

onde:

$$X(i) = \begin{bmatrix} X(i-1) \\ a(i) \\ ja(i-l d) \end{bmatrix} \quad (4.73)$$

$$b(i) = \begin{bmatrix} 0 \\ y(i) - a(i)h(i-i) \\ j[y(i-l d) - a(i-l d)h(i-1)] \end{bmatrix} \quad (4.74)$$

Substituindo (4.72) em (4.70), temos o seguinte sistema de equações :

$$X(i)^T X(i)[h(i) - h(i-1)] = X(i)^T b(i) - C^T \lambda(i) \quad (4.75)$$

$$C[h(i) - h(i-1)] = 0 \quad (4.76)$$

Pode ser verificado, sem dificuldade, analogamente a seção 2.4.1, que podemos obter $\lambda(i)$ das equações acima da seguinte forma:

$$[C(X(i)^T X(i))^{-1} C^T] \lambda(i) = C(X(i)^T X(i))^{-1} X(i)^T b(i) \quad (4.77)$$

Tendo $\lambda(i)$, de (4.75), podemos obter $h(i)$ da seguinte forma:

$$h(i) = h(i-1) + (X(i)^T X(i))^{-1} X(i)^T b(i) - (X(i)^T X(i))^{-1} C^T \lambda(i) \quad (4.78)$$

notando que:

$$(X(i)^T X(i))^{-1} X(i)^T b(i) = \hat{h}(i) - \hat{h}(i-1) = \hat{d}_h(i) \quad (4.79)$$

onde $\hat{h}(i)$ é a solução do PMQRJM (4.7) sem considerar a restrição (4.8). E assumindo:

$$Y(i) = (X(i)^T X(i))^{-1} C^T \quad (4.80)$$

Podemos reescrever (4.77), (4.78), (4.79), (4.80) da seguinte forma:

$$(X(i)^T X(i)) \hat{d}_h(i) = X(i)^T b(i) \quad (4.81)$$

$$(X(i)^T X(i)) Y(i) = C^T \quad (4.82)$$

$$[CY(i)]\lambda(i) = C\hat{d}_h(i) \quad (4.83)$$

$$h(i) = h(i-1) + \hat{d}_h(i) - Y(i)\lambda(i) \quad (4.84)$$

Substituindo (4.73), (4.74) em (4.81), (4.82), temos:

$$\begin{aligned} & [X^T(i-1)X(i-1) + a^T(i)a(i) - a^T(i-l_d)a(i-l_d)]\hat{d}_h(i) = \\ & a^T(i)[y(i) - a(i)h(i-1)] - a^T(i-l_d)[y(i-l_d) - a(i-l_d)h(i-1)] \end{aligned} \quad (4.85)$$

$$[X^T(i-1)X(i-1) + a^T(i)a(i) - a^T(i-l_d)a(i-l_d)]Y(i) = C^T \quad (4.86)$$

Das equações (4.83), (4.84), vemos que para obter recursivamente $h(i)$, devemos calcular recursivamente $\hat{d}_h(i)$ (4.85) e $Y(i)$ (4.86).

A seguir vamos mostrar como obter este cálculo recursivo utilizando equações normais e fatoração ortogonal.

4.4.1 Equações Normais

Para resolver recursivamente a equação (4.85), podemos dividir o problema em duas etapas:

$$[X^T(i-1)X(i-1) + a^T(i)a(i)]\hat{d}_{h1} = a^T(i)[y(i) - a(i)h(i-1)] \quad (4.87)$$

$$\tilde{h}(i) = h(i-1) + \hat{d}_{h1} \quad (4.88)$$

$$[\widetilde{X}^T(i)\widetilde{X}(i) - a^T(i-l_d)a(i-l_d)]\hat{d}_{h2} = -a^T(i-l_d)[y(i-l_d) - a(i-l_d)\tilde{h}(i)] \quad (4.89)$$

$$\hat{d}_h(i) = \hat{d}_{h1} + \hat{d}_{h2} \quad (4.90)$$

fazendo

$$\widetilde{X}(i) = \begin{bmatrix} X(i-1) \\ a(i) \end{bmatrix} \quad (4.91)$$

Definindo

$$S(i-1) = [X^T(i-1)X(i-1)]^{-1} \quad (4.92)$$

$$\tilde{S}(i) = [X(i-1)^T X(i-1) + a^T(i)a(i)]^{-1} \quad (4.93)$$

$$\tilde{v}(i) = S(i-1)a^T(i) \quad (4.94)$$

$$\tilde{u}(i) = C\tilde{v}(i) \quad (4.95)$$

$$\tilde{g}(i) = \frac{\tilde{v}(i)}{1 + a(i)\tilde{v}(i)} \quad (4.96)$$

$$\tilde{e}(i) = y(i) - a(i)h(i-1) \quad (4.97)$$

Utilizando Lema de Inversão de Matrizes (Apêndice A.1) e analogamente a seção 4.3.1, que:

$$\tilde{S}(i) = S(i-1) - \tilde{g}(i)\tilde{v}(i) \quad (4.98)$$

$$\hat{d}_{h1} = \tilde{g}(i)\tilde{e}(i) \quad (4.99)$$

$$\tilde{Y}(i) = Y(i-1) - \tilde{g}(i)\tilde{u}^T(i) \quad (4.100)$$

$$\tilde{h}(i) = h(i-1) + \hat{d}_{h1} \quad (4.101)$$

Definindo

$$S(i) = [\tilde{X}(i)^T \tilde{X}(i) - a^T(i-l_d)a(i-l_d)]^{-1} \quad (4.102)$$

$$v(i) = \tilde{S}(i)a^T(i-l_d) \quad (4.103)$$

$$u(i) = Cv(i) \quad (4.104)$$

$$g(i) = \frac{v(i)}{1 - a(i)v(i)} \quad (4.105)$$

$$e(i) = y(i-l_d) - a(i-l_d)\tilde{h}(i) \quad (4.106)$$

Utilizando o Lema de Inversão de Matrizes (Apêndice A.1) e analogamente à seção 4.3.1, temos:

$$S(i) = \tilde{S}(i) + g(i)v^T(i) \quad (4.107)$$

$$\hat{d}_{h2} = -g(i)e(i) \quad (4.108)$$

$$\hat{d}_h(i) = \hat{d}_{h1} + \hat{d}_{h2} \quad (4.109)$$

$$Y(i) = [\tilde{Y}(i) + g(i)u^T(i)] \quad (4.110)$$

Utilizando (4.109) e (4.110), de (4.83) podemos obter $\lambda(i)$ através da solução do sistema de equações

$$CY(i)\lambda(i) = C\hat{d}_h(i) \quad (4.111)$$

e conseqüentemente, de (4.84), obtemos

$$h(i) = h(i-1) + \hat{d}_h(i) - Y(i)\lambda(i) \quad (4.112)$$

Dessa forma, considerando as equações de (4.93) a (4.112), podemos racionalizar o cálculo recursivo do PMQRJM, através do algoritmo descrito na Tabela 7a, aqui denominado de AJNL.

Conforme pode ser visto na Tabela 7a, o esforço computacional requerido por iteração do algoritmo AJNL é devido principalmente às operações correspondentes às equações (4.118) a (4.136). Para essas equações, o número de FLOPS é dado na Tabela 7b.

Para o algoritmo AJNL, temos:

- de (4.79), (4.80), (4.102), temos que a precisão do cálculo de $\hat{d}_h(i)$ e $Y(i)$ dependem de

$$\text{cond}[S(i)^{-1}] = \text{cond}[X(i)^T X(i)] = \text{cond}[X(i)]^2 \quad (4.113)$$

- de (4.80), (4.83), temos que a precisão do cálculo de $\lambda(i)$ depende de

$$\text{cond}[CS(i)C^T] \quad (4.114)$$

TABELA 7a: AJNL - Algoritmo de Mínimos Quadrados utilizando Equações Normais, Janela Móvel de Dados com Multiplicadores de Lagrange.

Inicialização

$$S(0) = (A_0)^{-1} \quad (4.115)$$

$$Y(0) = S(0)C^T \quad (4.116)$$

$$h(0) \text{ tal que } Ch(0) = d \quad (4.117)$$

Atualizações

Para $i=1$ até final dos dados fazer

$$v(i) = S(i-1)a^T(i) \quad (4.118)$$

$$g(i) = \frac{v(i)}{1 + a(i)v(i)} \quad (4.119)$$

$$u(i) = Cv(i) \quad (4.120)$$

$$e(i) = y(i) - a(i)h(i-1) \quad (4.121)$$

$$\hat{d}_h(i) = g(i)e(i) \quad (4.122)$$

$$h(i) = h(i-1) + d_h(i) \quad (4.123)$$

$$S(i) = S(i-1) - g(i)v^T(i) \quad (4.124)$$

$$Y(i) = Y(i-1) - g(i)u^T(i) \quad (4.125)$$

• Se $i > ld$

$$v(i) = S(i)a^T(i-ld) \quad (4.126)$$

$$g(i) = \frac{v(i)}{1 - a(i-ld)v(i)} \quad (4.127)$$

$$u(i) = Cv(i) \quad (4.128)$$

$$e(i) = y(i-ld) - a(i-ld)h(i) \quad (4.129)$$

$$d_{h2} = -g(i)e(i) \quad (4.130)$$

$$h(i) = h(i) + d_{h2} \quad (4.131)$$

$$S(i) = S(i) + g(i)v^T(i) \quad (4.132)$$

$$\hat{d}_h(i) = \hat{d}_h(i) + d_{h2} \quad (4.133)$$

$$Y(i) = Y(i-1) + g(i)u^T(i) \quad (4.134)$$

- Resolver o sistema

$$CY(i)\lambda(i) = C\hat{d}_h(i) \quad (4.135)$$

Fazer

$$h(i) = h(i-1) + \hat{d}_h(i) - Y(i)\lambda(i) \quad (4.136)$$

TABELA 7b: Esforço computacional do AJNL

Equação	FLOPS
(4.118) e (4.126) :	$2n^2$
(4.119) e (4.127) :	$4n$
(4.120) e (4.128) :	$2nr$
(4.121) e (4.129) :	$2n$
(4.122) e (4.130) :	$2n$
(4.124) e (4.132) :	$2n^2$
(4.125) e (4.134) :	$2nr$
(4.135) :	$\frac{1}{6}r^3 + nr^2 + nr$
(4.136) :	nr
Total :	$\frac{1}{6}r^3 + nr^2 + 6nr + 4n^2 + 8n$

4.4.2 Fatoração Ortogonal

Em forma matricial, a equação (4.85), corresponde ao seguinte problema de mínimos quadrados sem restrições

$$\min_{\hat{d}_h(i)} J(i) = \left\| \begin{bmatrix} X(i-1) \\ a(i) \\ ja(i-l_d) \end{bmatrix} \hat{d}_h(i) - \begin{bmatrix} 0 \\ y(i) - a(i)h(i-1) \\ j[y(i-l_d) - a(i-l_d)h(i-1)] \end{bmatrix} \right\|_2 \quad (4.137)$$

Aplicando em (4.137), uma transformação ortogonal

$$\begin{bmatrix} \check{Q} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

tal que:

$$\check{Q}X(i-1) = \begin{bmatrix} R(i-1) \\ 0 \end{bmatrix} \quad (4.138)$$

onde $R(i-1)$ tem dimensões (n, n) com posto completo, temos:

$$\min_{\hat{d}_h(i)} J(i) = \left\| \begin{bmatrix} R(i-1) \\ 0 \\ a(i) \\ ja(i-l d) \end{bmatrix} \hat{d}_h(i) - \begin{bmatrix} 0 \\ y(i) - a(i)h(i-1) \\ j[y(i-l d) - a(i-l d)h(i-1)] \end{bmatrix} \right\|_2 \quad (4.139)$$

Aplicando em (4.139) uma matriz ortogonal

$$\begin{bmatrix} \tilde{Q} & 0 \\ 0 & 1 \end{bmatrix}$$

tal que:

$$\tilde{Q} \begin{bmatrix} R(i-1) \\ a(i) \end{bmatrix} = \begin{bmatrix} \tilde{R}(i) \\ 0 \end{bmatrix} \quad (4.140)$$

$$\tilde{Q} \begin{bmatrix} 0 \\ y(i) - a(i)h(i-1) \end{bmatrix} = \begin{bmatrix} \tilde{b}_1(i) \\ \tilde{b}_2(i) \end{bmatrix} \quad (4.141)$$

onde $\tilde{R}(i)$ tem dimensões (n, n) com posto completo e $\tilde{b}_1(i)$ tem dimensões $(n, 1)$, temos:

$$\min_{\hat{d}_h(i)} J(i) = \left\| \begin{bmatrix} \tilde{R}(i) \\ 0 \\ ja(i-l d) \end{bmatrix} \hat{d}_h(i) - \begin{bmatrix} \tilde{b}_1(i) \\ \tilde{b}_2(i) \\ j[y(i-l d) - a(i-l d)h(i-1)] \end{bmatrix} \right\|_2 \quad (4.142)$$

Aplicando em (4.142) uma matriz ortogonal Q tal que:

$$Q \begin{bmatrix} \tilde{R}(i) \\ ja(i - ld) \end{bmatrix} = \begin{bmatrix} R(i) \\ 0 \end{bmatrix} \quad (4.143)$$

$$Q \begin{bmatrix} \tilde{b}_1(i) \\ j[y(i - ld) - a(i - ld)h(i - 1)] \end{bmatrix} = \begin{bmatrix} b_1(i) \\ jb_2(i) \end{bmatrix} \quad (4.144)$$

onde $R(i)$ tem dimensões (n, n) com posto completo e $b_1(i)$ tem dimensões $(n, 1)$, temos finalmente:

$$\min_{\hat{d}_h(i)} J(i) = \left\| \begin{bmatrix} R(i) \\ 0 \end{bmatrix} \hat{d}_h(i) - \begin{bmatrix} b_1(i) \\ jb_2(i) \end{bmatrix} \right\|_2 \quad (4.145)$$

Pela estrutura das matrizes envolvidas, temos $\hat{d}_h(i)$ solução do seguinte sistema de equações

$$R(i)\hat{d}_h(i) = b_1(i) \quad (4.146)$$

Com as transformações ortogonais (4.138), (4.140), (4.143) foi obtido

$$Q \begin{bmatrix} \tilde{Q} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \check{Q} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X(i-1) \\ 0 \\ a(i) \\ ja(i - ld) \end{bmatrix} = \begin{bmatrix} R(i) \\ 0 \end{bmatrix} \quad (4.147)$$

Como transformações Q , $\tilde{Q}$, $\check{Q}$ são ortogonais, devemos ter:

$$X^T(i-1)X(i-1) + a^T(i)a(i) - a^T(i-ld)a(i-ld) = R^T(i)R(i) \quad (4.148)$$

Substituindo (4.148) em (4.86) temos:

$$R^T(i)R(i)Y(i) = C^T$$

$$Y(i) = R(i)^{-1}(R(i)^T)^{-1}C^T \quad (4.149)$$

Substituindo (4.146) e (4.149) em (4.83), temos:

$$[CR(i)^{-1}](R(i)^T)^{-1}C^T\lambda(i) = [CR(i)^{-1}]b_1(i) \quad (4.150)$$

Em forma matricial a equação (4.150) corresponde ao seguinte problema de mínimos quadrados

$$\|\tilde{Y}(i)\lambda(i) - b_1(i)\|_2 \quad (4.151)$$

com $\tilde{Y}(i)$ solução do sistema de equações

$$R^T(i)\tilde{Y}(i) = C^T \quad (4.152)$$

Seja

$$E(i) = \tilde{Y}(i)\lambda(i) - b_1(i) \quad (4.153)$$

que corresponde ao vetor resíduo do problema (4.151).

Multiplicando (4.84) por $R(i)$, temos:

$$R(i)[h(i) - h(i-1)] = R(i)\hat{d}_h(i) - R(i)Y(i)\lambda(i) \quad (4.154)$$

Substituindo (4.146) e (4.149) em (4.154), temos:

$$R(i)[h(i) - h(i-1)] = b_1(i) - (R(i)^T)^{-1}C^T\lambda(i) \quad (4.155)$$

Substituindo (4.152) e em seguida (4.153) em (4.155), temos:

$$R(i)[h(i) - h(i-1)] = -E(i) \quad (4.156)$$

Dessa forma, considerando as equações de (4.140), (4.141), (4.143), (4.144), (4.151), (4.152), (4.153) e (4.156), podemos racionalizar o cálculo recursivo do PMQRJM, através do algoritmo descrito na Tabela 8a, aqui denominado de AJQL.

Conforme pode ser visto na Tabela 8a, o esforço computacional requerido por iteração do algoritmo AJQL é devido principalmente às operações correspondentes às equações (4.159) a (4.165). Para essas equações, o número de FLOPS é dado na Tabela 8b.

- De (4.73), (4.74), (4.140), (4.143), (4.145), (4.152), temos que a precisão do cálculo de $\hat{d}_h(i)$ e $Y(i)$ dependem de

$$\text{cond}[R(i)] = \text{cond}[X(i)] \quad (4.157)$$

- De (4.151), (4.152) temos que o cálculo de $\lambda(i)$ depende de

$$\text{cond}[\tilde{Y}(i)] = \text{cond}[R(i)^{-1}C^T] \quad (4.158)$$

TABELA 8a: AJQL - Algoritmo de Mínimos Quadrados utilizando Fatoração Ortogonal, Janela Móvel dos Dados com Multiplicadores de Lagrange.

Inicialização

- Obter $R(0)$, $h(0)$

Atualizações

Para $i=1$ até final dos dados faça

- Determinar $R(i)$, $b_1(i)$ e $b_2(i)$:

$$\tilde{Q} \begin{bmatrix} R(i-1) \\ a(i) \end{bmatrix} = \begin{bmatrix} R(i) \\ 0 \end{bmatrix} \quad (4.159)$$

$$\tilde{Q} \begin{bmatrix} 0 \\ y(i) - a(i)h(i-1) \end{bmatrix} = \begin{bmatrix} b_1(i) \\ b_2(i) \end{bmatrix} \quad (4.160)$$

- Se $i > ld$
— Determinar $R(i)$, $b_1(i)$ e $b_2(i)$:

$$Q \begin{bmatrix} R(i) \\ ja(i-ld) \end{bmatrix} = \begin{bmatrix} R(i) \\ 0 \end{bmatrix} \quad (4.161)$$

$$Q \begin{bmatrix} b_1(i) \\ j[y(i-ld) - a(i-ld)h(i-1)] \end{bmatrix} = \begin{bmatrix} b_1(i) \\ jb_2(i) \end{bmatrix} \quad (4.162)$$

- Determinar $\tilde{Y}(i)$, resolvendo o seguinte sistema de equações triangular

$$R(i)^T \tilde{Y}(i) = C^T \quad (4.163)$$

- Determinar $E(i)$ correspondente ao vetor resíduo do seguinte problema de mínimos quadrados lineares

$$\|\tilde{Y}(i)\lambda(i) - b_1(i)\|_2 \quad (4.164)$$

- Determinar dh , resolvendo o seguinte sistema de equações triangular

$$R(i)dh = -E(i) \quad (4.165)$$

- Determinar os parâmetros $h(i)$

$$h(i) = h(i-1) + dh \quad (4.166)$$

É importante notar que as transformações ortogonais $\tilde{Q}$ e Q podem ser obtidas de forma eficiente e numericamente estável, através de uma sequência de transformações de Givens, conforme apresentado no Apêndice A.2. A construção explícita de $\tilde{Q}$ e Q não é necessária.

TABELA 8b: Esforço computacional do AJQL

Equação	FLOPS
(4.159) e (4.161) :	$4n^2 + 14n$
(4.160) e (4.162) :	$4n$
(4.163) :	$nr(n+1)/2$
(4.164) :	$\frac{1}{6}r^3 + r^2n + rn$
(4.165) :	$nr + n(n+1)/2$
Total :	$\frac{1}{6}r^3 + r^2n + \frac{1}{2}n^2(8+r) + \frac{5}{2}rn + \frac{37}{2}n$

4.5 Análise Comparativa

Para comparar o esforço computacional por passo requerido pelos vários métodos, vamos considerar um problema com as seguintes características:

$$n = 10; \quad r = 2$$

Das Tabelas 5b, 6b, 7b, 8b, temos a seguinte estimativa de FOLPS:

- AJNZ: 644
- AJQZ: 788
- AJNL: 641
- AJQL: 776

Para $r=5$, temos os seguintes FLOPS:

- AJNZ: 350
- AJQZ: 365
- AJNL: 1050
- AJQL: 1230

Desta forma, vemos que para valores baixos de r/n , o esforço computacional requerido pelos algoritmos que utilizam o espaço nulo Z é maior que o requerido por seus similares utilizando multiplicadores de Lagrange λ . Essa relação é invertida no caso de valores altos de r/n . O esforço computacional requerido pelos algoritmos que utilizam fatoração QR é maior que o de seus similares utilizando equações normais.

Comparando os resultados apresentados na seção 3.5, com os desta seção, vemos que os algoritmos que utilizam janela móvel de dados requerem um esforço computacional maior que o de seus similares utilizando ponderação exponencial de dados.

De (4.35), (4.61), (4.113), (4.157), temos que a precisão dos resultados obtidos dependem principalmente de:

- AJNZ: $\text{cond}[X(i)Z]^2$

- AJQZ: $\text{cond}[X(i)Z]$
- AJNL: $\text{cond}[X(i)]^2$
- AJQL: $\text{cond}[X(i)]$

Desta forma, podemos concluir que:

- O uso de fatoração QR não altera o condicionamento original do PMQRJM e os algoritmos que a utilizam possuem condicionamento numérico muito superior aos seus similares utilizando equações normais. O uso desta última não é recomendado para problemas inerentemente mal condicionados ($\text{cond}[X(i)]$ elevado).
- o condicionamento numérico dos algoritmos utilizando o espaço nulo Z é superior ao de seus similares utilizando multiplicadores de Lagrange λ .

Devido às multiplicações por W , de (3.5) e (4.6) em geral devemos ter:

$$\text{cond}[\bar{X}(i)] \geq \text{cond}[X(i)] \quad (4.167)$$

Desta forma, dos resultados apresentados na seção 3.5 e os aqui apresentados, vemos que o condicionamento numérico dos algoritmos utilizando janela móvel de dados é em geral melhor que seus similares utilizando ponderação exponencial de dados.

Capítulo 5

Exemplo Numérico

5.1 Introdução

Neste capítulo, para ilustrar e comparar o desempenho dos algoritmos apresentados nos capítulos 3 e 4, é apresentado um exemplo numérico correspondente a um problema de síntese de um filtro de resposta impulsiva finita (FIR) para a detecção de frequências em um sinal com ruídos.

5.2 Formulação do Problema

Dado um sinal $x(i)$

$$x(i) = \text{sen}(i\pi/4) + \text{sen}(i7\pi/16) + \text{sen}(i11\pi/16) + 0.3\xi(i), \quad i \geq 1$$

onde ξ é uma variável aleatória com distribuição uniforme no intervalo $[-.5;.5]$.

Considere o problema de determinar de forma adaptativa os parâmetros $h(i) \in \mathbb{R}^n$ de um filtro FIR

$$y(i) = a(i)h(i) \tag{5.1}$$

$$a(i) = [x(i) \cdot x(i-1) \dots x(i-n+1)] \tag{5.2}$$

de forma a minimizar uma norma quadrática de $y(i)$ sujeito à restrição de que o módulo da resposta em frequência do filtro seja unitária nas frequências $f_1 = (7\pi/4)$ e $f_2 = (11\pi/16)$.

Em [10], [11] é demonstrado que a restrição sobre a resposta em frequência do filtro corresponde ao seguinte conjunto de restrições sobre os parâmetros do mesmo:

$$Ch(i) = d \quad (5.3)$$

onde

$$C = \begin{bmatrix} 1 & \cos f_1 & \cos 2f_1 & \dots & \cos(n-1)f_1 \\ 1 & \cos f_2 & \cos 2f_2 & \dots & \cos(n-1)f_2 \\ 0 & \sin f_1 & \sin 2f_1 & \dots & \sin(n-1)f_1 \\ 0 & \sin f_2 & \sin 2f_2 & \dots & \sin(n-1)f_2 \end{bmatrix} \quad (5.4)$$

$$d = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad (5.5)$$

O problema do filtro FIR com restrições (5.1) a (5.5), será formulado com os seguintes problemas de mínimos quadrados tratados neste trabalho:

- PMQRPE: mínimos quadrados com ponderação exponencial (Capítulo 3) com

$$n = 11; \quad W = .95; \quad R(0) = 10^{-5} I_{(11 \times 11)}$$

- PMQRJM: mínimos quadrados com janela móvel de dados (Capítulo 4) com

$$n = 11; \quad ld = 44; \quad R(0) = 10^{-5} I_{(11 \times 11)}$$

Para a solução do PMQRPE serão utilizados os seguintes algoritmos:

- APNZ: equações normais com espaço nulo de C (seção 3.3.1)
- APQZ: fatoração ortogonal com espaço nulo de C (seção 3.3.2)
- APNL: equações normais com multiplicadores de Lagrange (seção 3.4.1)
- APQL: fatoração ortogonal com multiplicadores de Lagrange (seção 3.4.2)

Para a solução do PMQRJM serão utilizados os seguintes algoritmos:

- AJNZ: equações normais com espaço nulo de C (seção 4.3.1)
- AJQZ: fatoração ortogonal com espaço nulo de C (seção 4.3.2)
- AJNL: equações normais com multiplicadores de Lagrange (seção 4.4.1)
- AJQL: fatoração ortogonal com multiplicadores de Lagrange (seção 4.4.2)

Sejam

$$e(i) = a(i)h(i) \quad (5.6)$$

$$er(i) = Ch(i) - d \quad (5.7)$$

o erro na equação de regressão (5.1) e nas restrições (5.3), respectivamente.

5.3 Resultados Obtidos

Nas figuras 1 a 4 são apresentados os resultados obtidos de $\|e(i)\|_2$ e $\|er(i)\|_2$ para a solução do PMQRPE utilizando os algoritmos APNZ, APQZ, APNL, APQL, respectivamente. Nas figuras 5 a 8 são apresentados os resultados obtidos de $\|e(i)\|_2$ e $\|er(i)\|_2$ para a solução do PMQRJM utilizando os algoritmos AJNZ, AJQZ, AJNL, AJQL, respectivamente.

Para melhor comparar os resultados apresentados nas figuras 1 a 8, algumas combinações dos mesmos são apresentados:

- Figura 9: APNZ $\times$ APNL
- Figura 10: APNZ $\times$ APQL
- Figura 11: APNZ $\times$ AJNL
- Figura 12: APQZ $\times$ AJQZ
- Figura 13: AJNL $\times$ AJQZ

Dos resultados obtidos, podemos ver que:

- Figura 9: os resultados obtidos com APNZ são superiores aos do APNL, mostrando que o uso do espaço nulo no lugar dos multiplicadores de Lagrange, contribui com a melhoria do condicionamento numérico do problema. Neste caso, para o APNL, vemos que inclusive as restrições de igualdade figura 9b, deixam de ser respeitadas a partir de $i > 360$, o que não ocorre com o APNZ.
- Figura 10: os resultados obtidos com APQL são superiores aos do APNZ e consequentemente aos do APNL, mostrando que o uso da fatoração ortogonal no lugar das equações normais, contribui com a melhoria do condicionamento numérico do problema.
- Figuras 11 e 12: os resultados obtidos com AJNL e AJQZ são superiores aos do APNZ e APQZ, mostrando que o uso de janela móvel no lugar de ponderação exponencial, contribui com a melhoria do condicionamento numérico do problema.
- Figura 13: os resultados obtidos com AJNL e AJQZ, mostram que neste caso, a melhoria do condicionamento numérico obtido com o uso da janela móvel de dados, foi suficiente para que não mais se percebesse as diferenças devidas ao uso de fatoração ortogonal e espaço nulo.

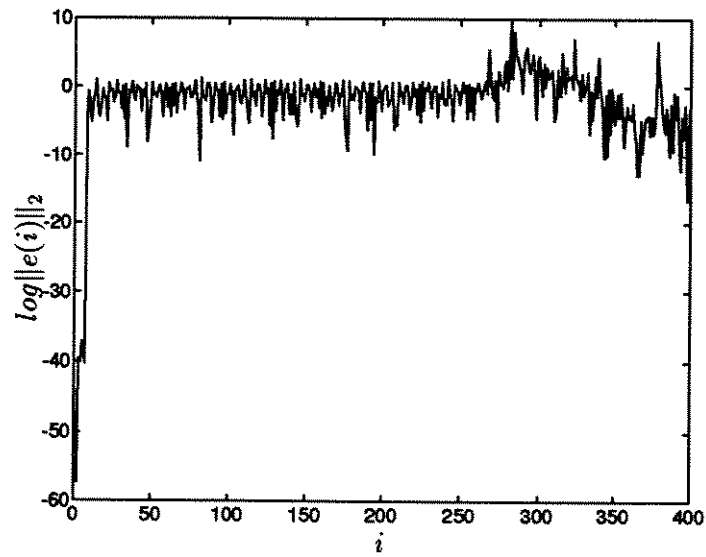


Figura 1a: APNZ - Variação de $\log\|e(i)\|_2$

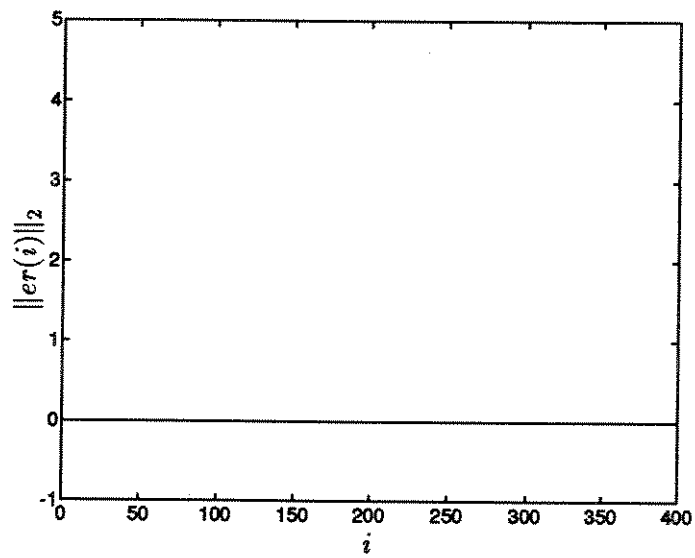


Figura 1b: APNZ - Variação de $\|er(i)\|_2$

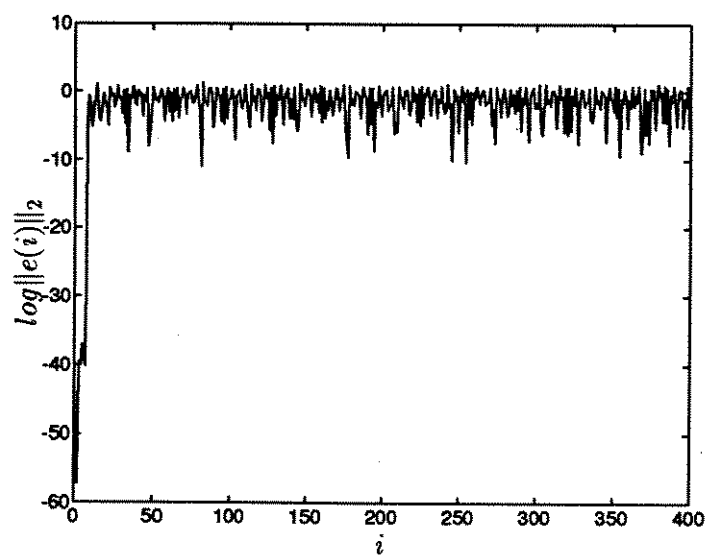


Figura 2a: APQZ - Variação de $\log\|e(i)\|_2$

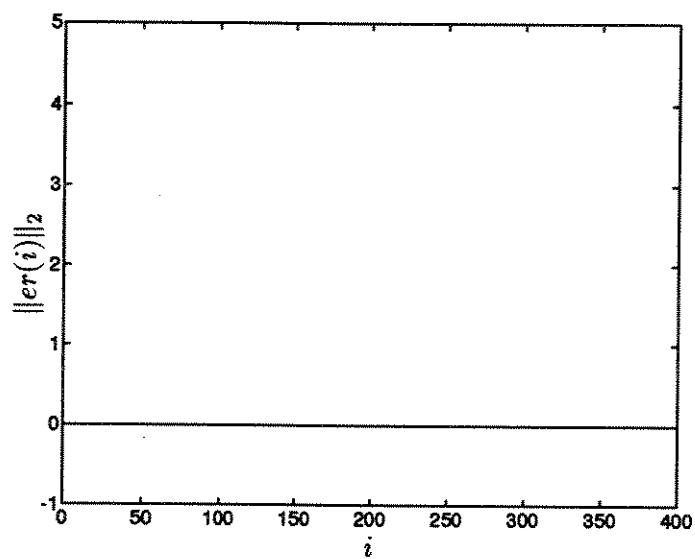


Figura 2b: APQZ - Variação de $\|er(i)\|_2$

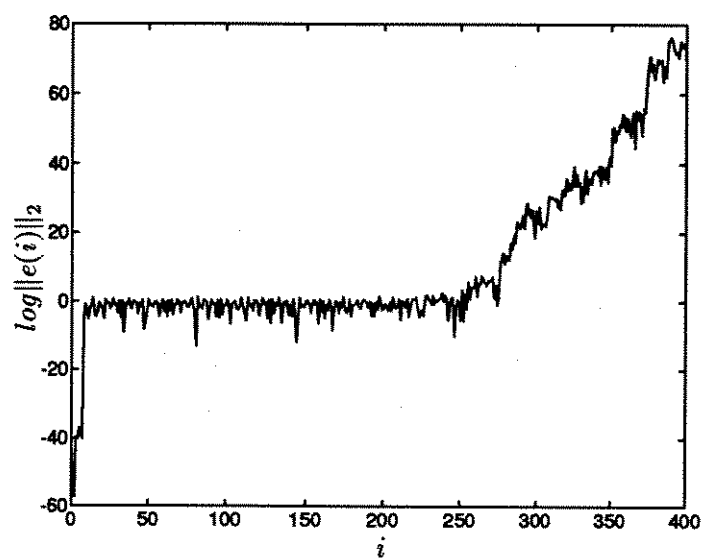


Figura 3a: APNL- Variação de $\log\|e(i)\|_2$

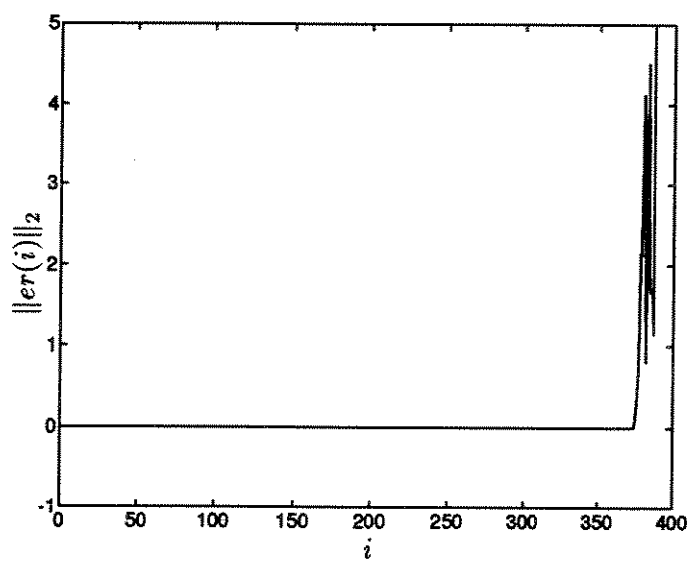


Figura 3b: APNL - Variação de $\|er(i)\|_2$

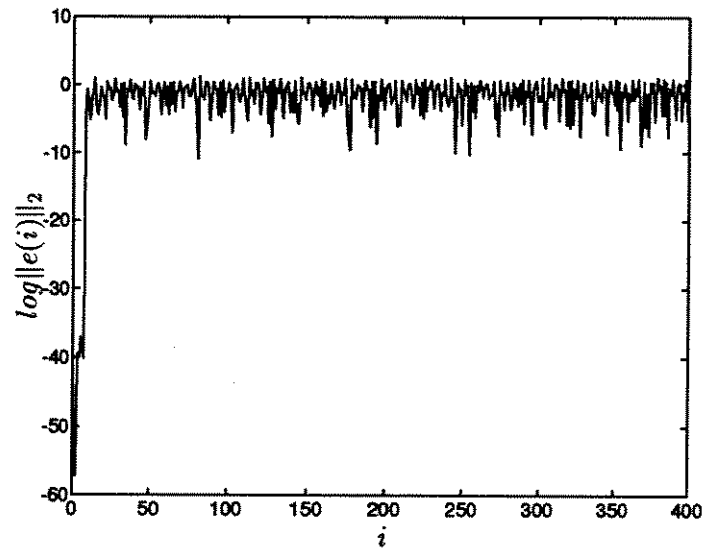


Figura 4a: APQL - Variação de $\log\|e(i)\|_2$

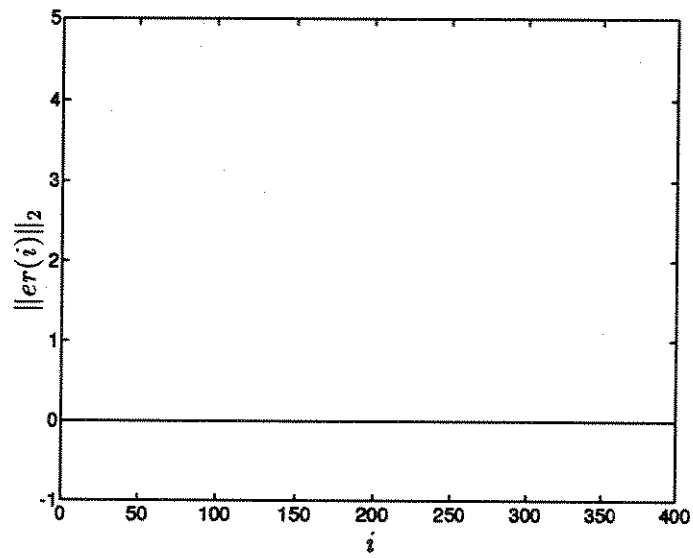


Figura 4b: APQL- Variação de $\|er(i)\|_2$

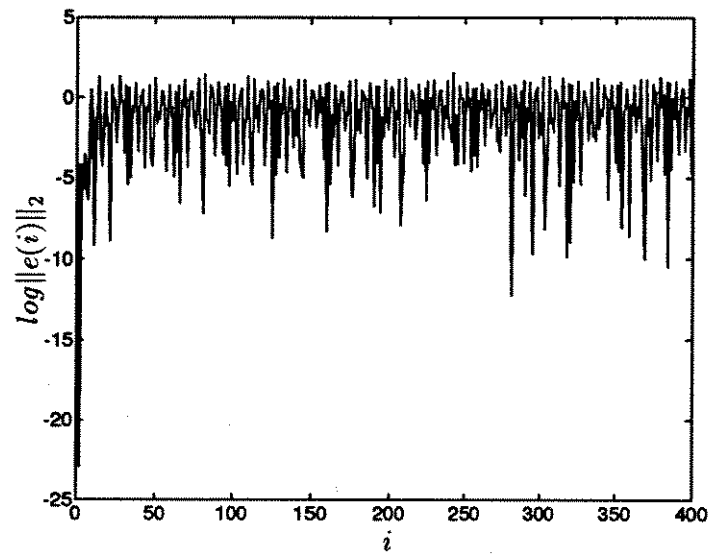


Figura 5a: AJNZ - Variação de $\log\|e(i)\|_2$

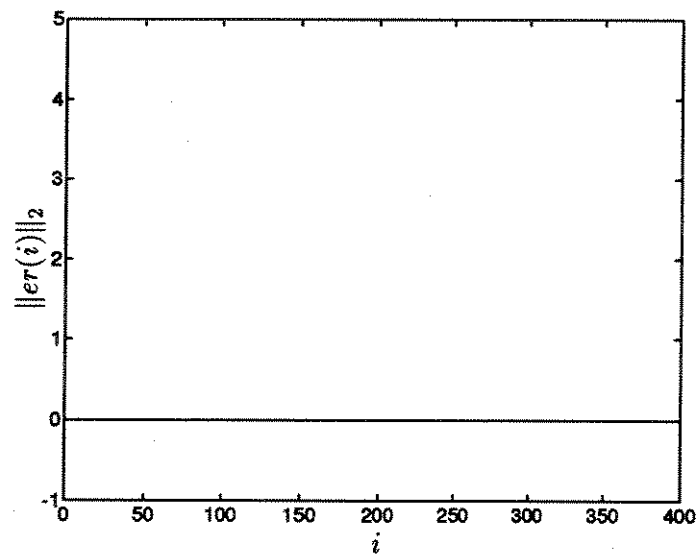
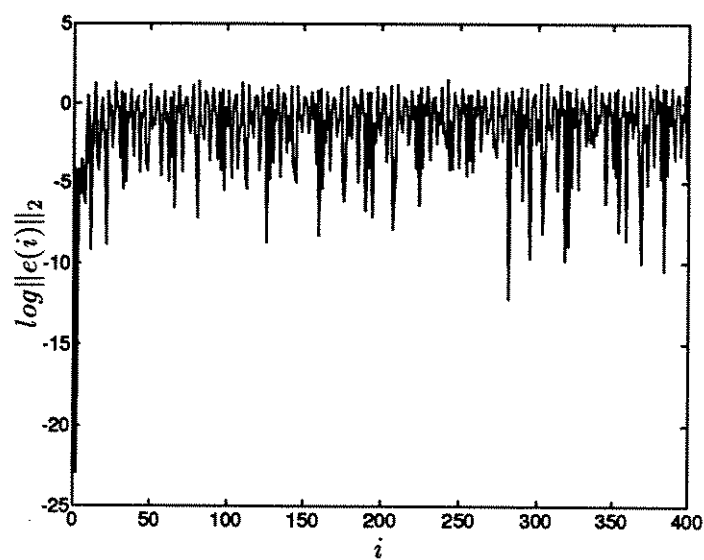
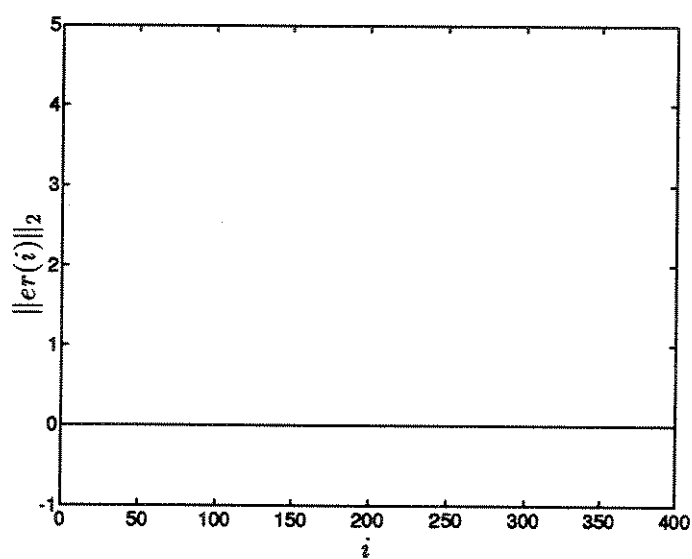


Figura 5b: AJNZ - Variação de $\|er(i)\|_2$

Figura 6a: AJQZ - Variação de $\log\|e(i)\|_2$ Figura 6b: AJQZ - Variação de $\|er(i)\|_2$

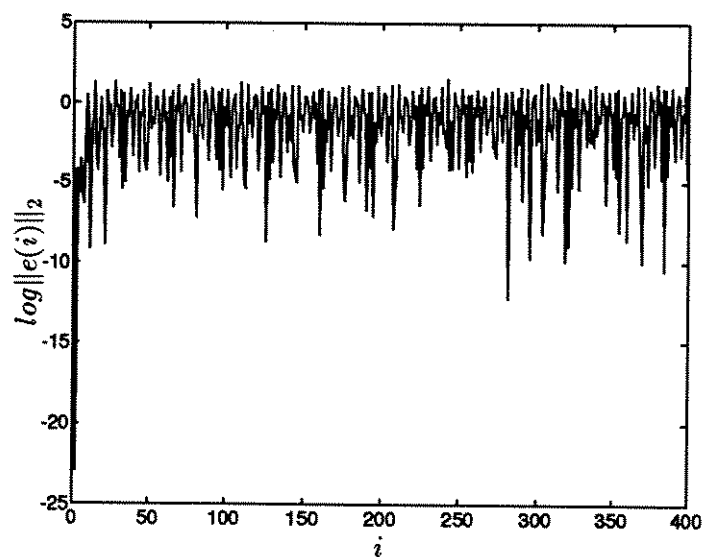


Figura 7a: AJNL - Variação de $\log\|e(i)\|_2$

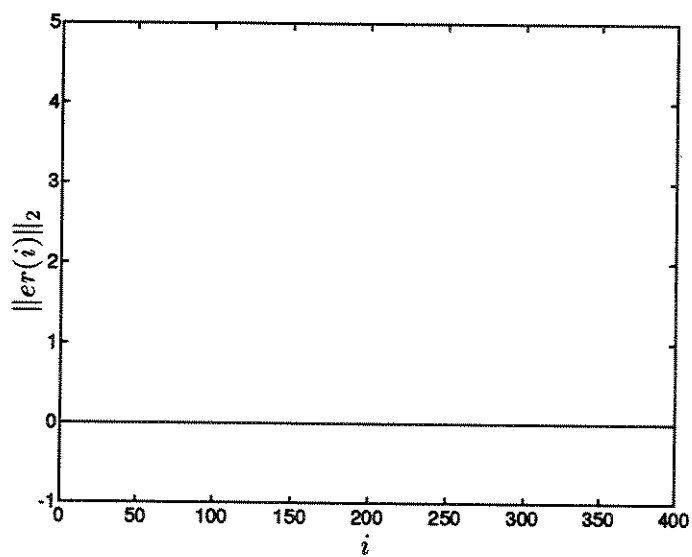


Figura 7b: AJNL - Variação de $\|er(i)\|_2$

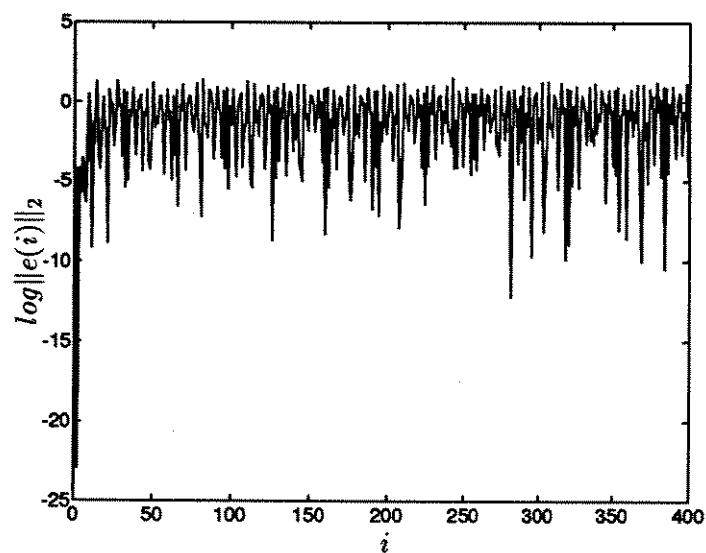


Figura 8a: AJQL - Variação de $\log\|e(i)\|_2$

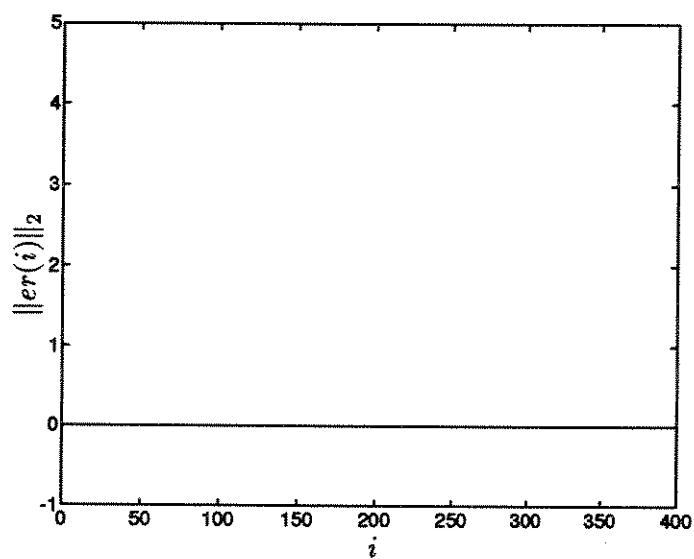


Figura 8b: AJQL - Variação de $\|er(i)\|_2$

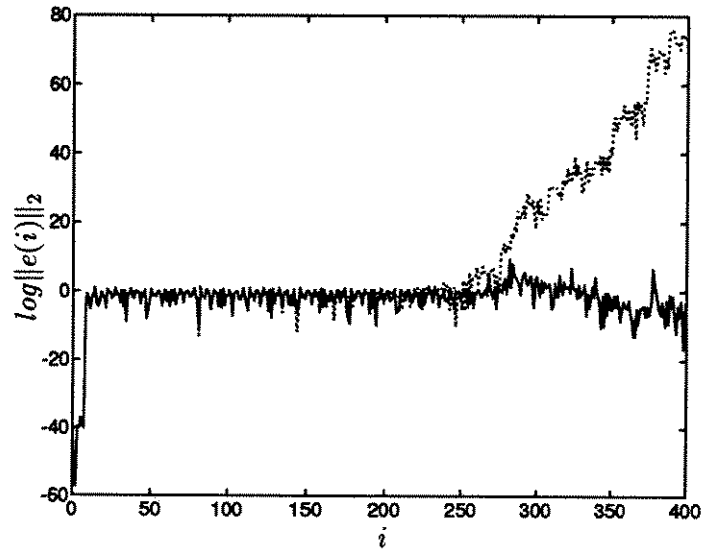


Figura 9a: Variação de $\log\|e(i)\|_2$: (—) APNZ, (....) APNL

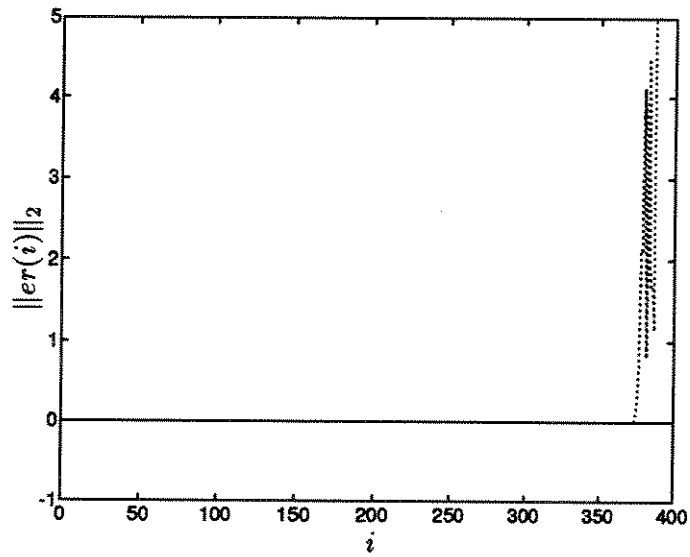


Figura 9b: Variação de $\|er(i)\|_2$: (—) APNZ, (....) APNL

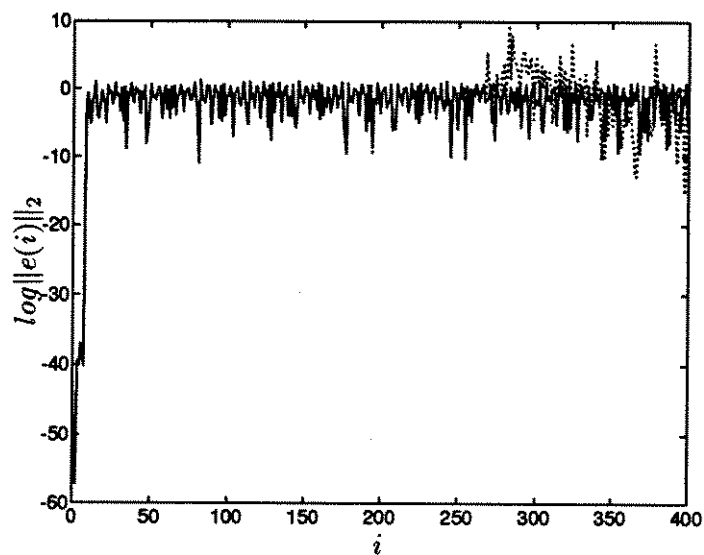


Figura 10a: Variação de $\log\|e(i)\|_2$: (....) APNZ, (—) APQL

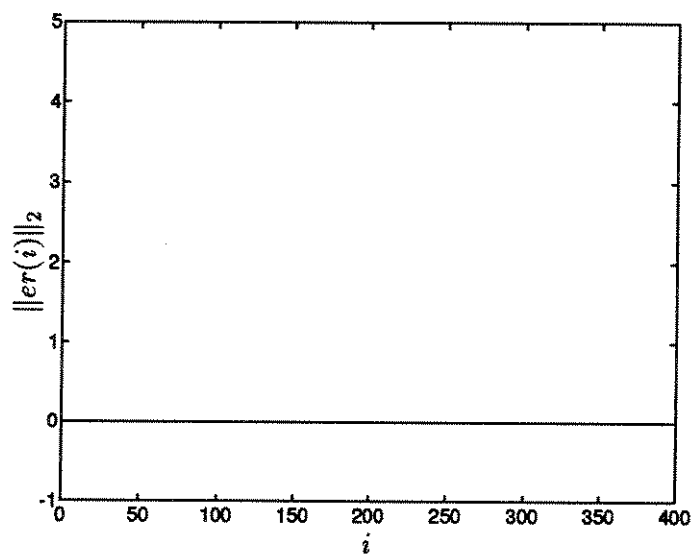


Figura 10b: Variação de $\|er(i)\|_2$: (....) APNZ, (—) APQL

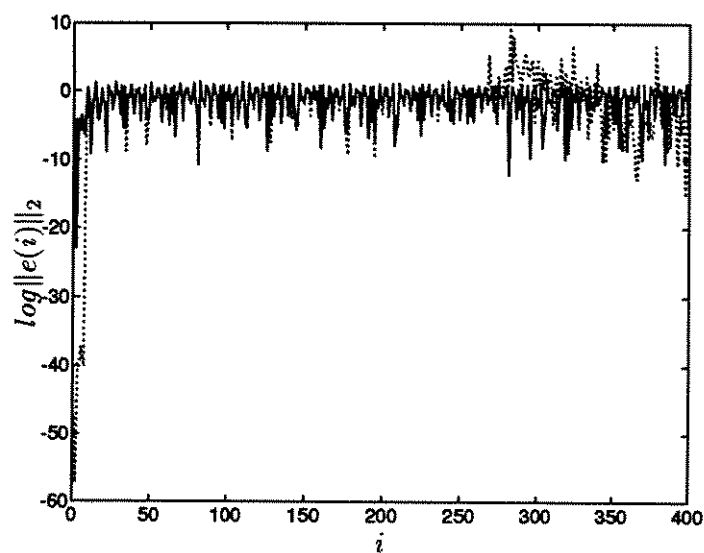


Figura 11a: Variação de $\log\|e(i)\|_2$: (....) APNZ, (—) AJNL

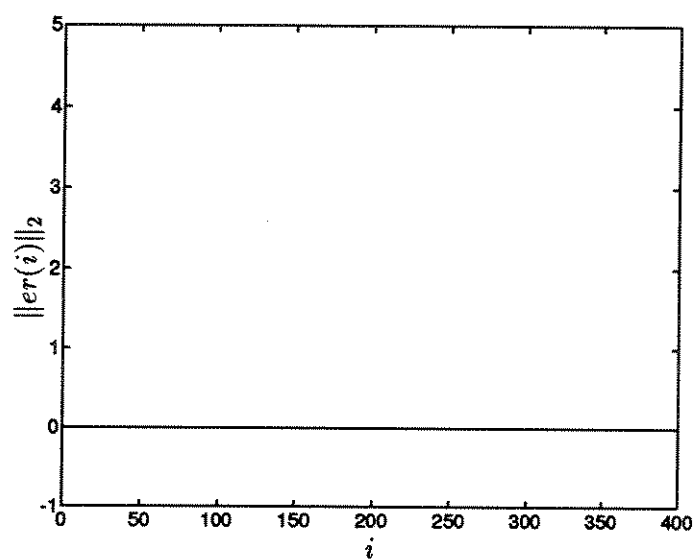


Figura 11b: Variação de $\|er(i)\|_2$: (....) APNZ, (—) AJNL

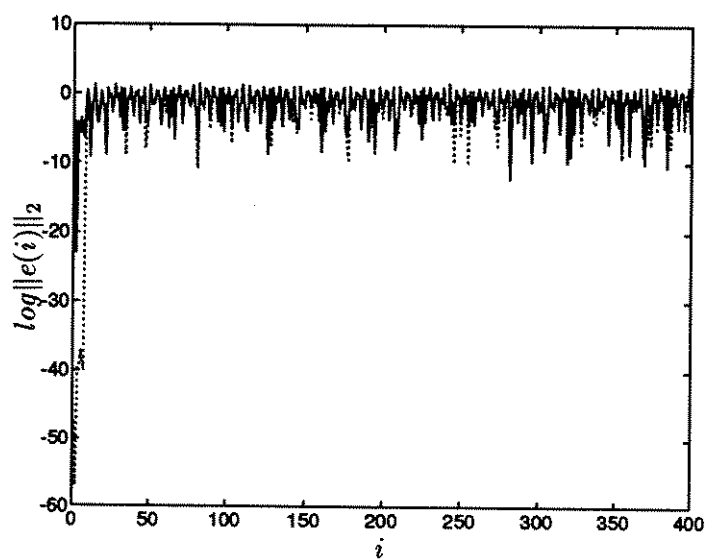


Figura 12a: Variação de $\log\|e(i)\|_2$: (....) APQZ, (—) AJQZ

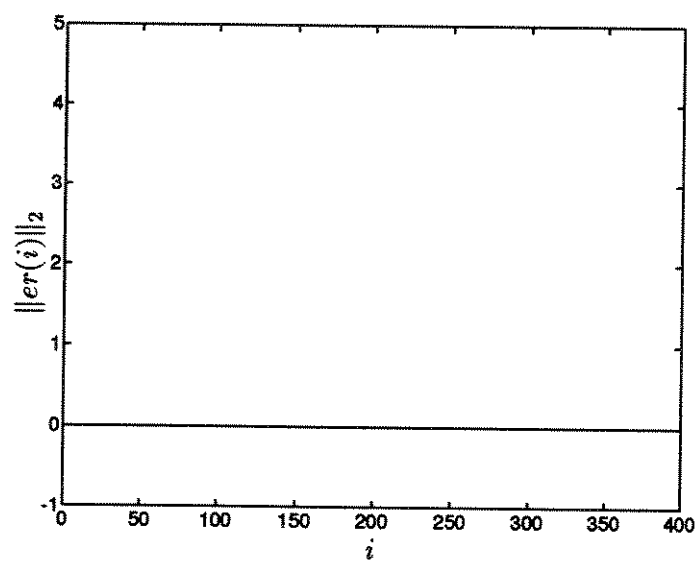


Figura 12b: Variação de $\|er(i)\|_2$: (....) APQZ, (—) AJQZ

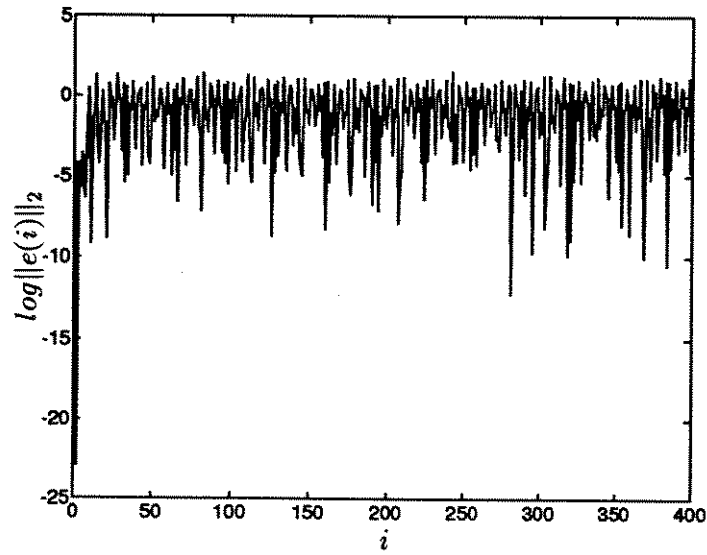


Figura 13a: Variação de $\log\|e(i)\|_2$: (....) AJNL, (—) AJQZ

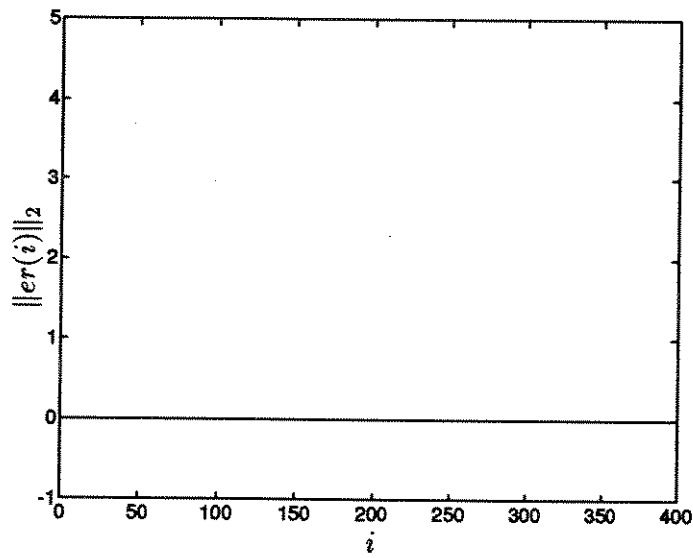


Figura 13b: Variação de $\|er(i)\|_2$: (....) AJNL, (—) AJQZ

Conclusão Geral

Este trabalho tratou do problema de estimação adaptativa de parâmetros através de mínimos quadrados lineares sujeito a restrições de igualdade lineares.

Foram apresentados algoritmos recursivos explorando as combinações cabíveis entre ponderação exponencial de dados e janela móvel de dados com o espaço nulo da matriz de restrições, multiplicadores de Lagrange, equações normais e fatoração QR.

O condicionamento numérico e o esforço computacional por passo de cada um dos algoritmos foi estimado e comparados entre si. Com os resultados obtidos pode-se ter uma visão panorâmica e crítica das características desses algoritmos. Este fato, é muito útil na seleção e aplicação dos algoritmos para a solução de problemas práticos.

Em particular, os algoritmos que utilizam janela móvel de dados com fatoração QR apresentados neste trabalho, são resultados que o autor acredita serem contribuições originais.

A continuação natural dos resultados aqui apresentados é a extensão dos mesmos a problemas com restrições lineares de desigualdade.

Apêndice A

A.1 Lema de Inversão de Matrizes

Sejam Z e S matrizes não singulares de dimensões (m, m) , Q uma matriz não singular de dimensões (i, i) , U uma matriz de dimensões (m, i) e V uma matriz de dimensões (i, m) tais que satisfaçam a equação

$$Z = S + UQV$$

Então a inversa da matriz Z é dada por :

$$Z^{-1} = (S + UQV)^{-1} = S^{-1} - S^{-1}U(VS^{-1}U + Q^{-1})^{-1}VS^{-1}$$

Se $U=V^T$, substituindo na última expressão :

$$Z^{-1} = S^{-1} - S^{-1}U(U^T S^{-1}U + Q^{-1})^{-1}U^T S^{-1}$$

Para o caso particular onde Q é um escalar, U um vetor de dimensões $(m, 1)$ denotado por u e V um vetor de dimensões $(1, m)$ denotado por v^T , temos :

$$Z = S + uv^T$$

e

$$Z^{-1} = S^{-1} - \frac{(S^{-1}u)(v^T S^{-1})}{1 + v^T S^{-1}u}$$

Ainda, neste caso, se $u = v$

$$z = S + uu^T$$

e

$$Z^{-1} = S^{-1} - \frac{(S^{-1}u)(u^T S^{-1})}{1 + u^T S^{-1}u}$$

A.2 Fatoração QR utilizando matrizes de rotação de Givens

Uma matriz de Givens $G_n(c, s, i, j)$ é uma matriz ortogonal de dimensões $(n \times n)$ cujos elementos são iguais ao da matriz unidade I_n , exceto

$$g_{ii} = g_{jj} = c \quad (\text{A.1})$$

$$g_{ij} = -g_{ji} = s \quad (\text{A.2})$$

$$c^2 + s^2 = 1 \quad (\text{A.3})$$

Multiplicando uma matriz $A(m \times n)$ por uma matriz de Givens $G_m(c, s, i, j)$, todas as linhas de A ficam inalteradas exceto as linhas i, j , que são modificadas para

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} a_{i1} & a_{i2} & \dots & a_{in} \\ a_{j1} & a_{j2} & \dots & a_{jn} \end{bmatrix} = \begin{bmatrix} \tilde{a}_{i1} & \tilde{a}_{i2} & \dots & \tilde{a}_{in} \\ \tilde{a}_{j1} & \tilde{a}_{j2} & \dots & \tilde{a}_{jn} \end{bmatrix} \quad (\text{A.4})$$

Podemos fazer $\tilde{a}_{j1} = 0$ em (A.4), utilizando a_{i1} como pivô [15], através de uma matriz $G_m(c, s, i, j)$, com:

$$c = \frac{a_{i1}}{\sqrt{a_{i1}^2 + a_{j1}^2}}; \quad s = \frac{a_{j1}}{\sqrt{a_{i1}^2 + a_{j1}^2}} \quad (\text{A.5})$$

Desta forma, dada uma matriz A de dimensões $(m \times n)$, $m \geq n$, posto $A = n$, podemos obter a fatoração ortogonal [15]:

$$QA = \begin{bmatrix} R \\ 0 \end{bmatrix} \quad (\text{A.6})$$

onde R é uma matriz triangular superior de dimensões $(n \times n)$ com posto completo e Q é uma matriz ortogonal de dimensões $(m \times m)$, dada por

$$Q = Q_1 Q_2 Q_3 \dots \quad (\text{A.7})$$

onde Q_k são matrizes de rotação de Givens, calculados de forma a anular progressivamente todos os elementos abaixo da diagonal principal de A .

Quando apenas a matriz R é necessária, a matriz Q não necessita ser construída explicitamente e R é obtida aplicando as matrizes de Givens diretamente em A .

No caso de

$$Q \begin{bmatrix} R_1 \\ a \end{bmatrix} = \begin{bmatrix} R_2 \\ 0 \end{bmatrix} \quad (\text{A.8})$$

onde R_1, R_2 são matrizes triangulares superiores de dimensões $(n \times n)$ e a é um vetor de dimensões $(1 \times n)$, a matriz Q pode ser obtida através de

$$Q = Q_1 Q_2 \dots Q_n \quad (\text{A.9})$$

onde Q_k são matrizes de Givens de dimensões $(n+1, n+1)$, calculadas de forma a anular progressivamente, da esquerda para a direita, todos os componentes do vetor a .

Consideremos o seguinte caso

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{-1} \end{bmatrix} \begin{bmatrix} x_1 & x_2 & \dots & x_n \\ y_1 & y_2 & \dots & y_n \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{-1} \end{bmatrix} \begin{bmatrix} \tilde{x}_1 & \tilde{x}_2 & \dots & \tilde{x}_n \\ 0 & \tilde{y}_2 & \dots & \tilde{y}_n \end{bmatrix} \quad (\text{A.10})$$

De (A.5), temos

$$c = \frac{x_1}{\sqrt{x_1^2 - y_1^2}}; \quad s = \frac{y_1 \sqrt{-1}}{\sqrt{x_1^2 - y_1^2}} \quad (\text{A.11})$$

Substituindo (A.11) em (A.10), podemos verificar que

$$\begin{bmatrix} \bar{c} & -\bar{s} \\ -\bar{s} & \bar{c} \end{bmatrix} \begin{bmatrix} x_1 & x_2 & \dots & x_n \\ y_1 & y_2 & \dots & y_n \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{-1} \end{bmatrix} \begin{bmatrix} \tilde{x}_1 & \tilde{x}_2 & \dots & \tilde{x}_n \\ 0 & \tilde{y}_2 & \dots & \tilde{y}_n \end{bmatrix} \quad (\text{A.12})$$

onde $\bar{c}$, $\bar{s}$ são dados por

$$\bar{c} = \frac{x_1}{\sqrt{x_1^2 - y_1^2}}; \quad \bar{s} = \frac{y_1}{\sqrt{x_1^2 - y_1^2}} \quad (\text{A.13})$$

Assim sendo, na fatoração ortogonal

$$Q \begin{bmatrix} R_1 \\ a\sqrt{-1} \end{bmatrix} = \begin{bmatrix} R_2 \\ 0 \end{bmatrix} \quad (\text{A.14})$$

a matriz R_2 pode ser obtida, fazendo

$$\bar{Q} \begin{bmatrix} R_1 \\ a \end{bmatrix} = \begin{bmatrix} R_2 \\ 0 \end{bmatrix} \quad (\text{A.15})$$

onde $\bar{Q}$ corresponde a uma sequência de matrizes com a mesma estrutura das matrizes de rotação de Givens, mas com elementos dados por

$$\bar{g}_{ii} = \bar{g}_{jj} = \bar{c} \quad (\text{A.16})$$

$$\bar{g}_{ij} = \bar{g}_{ji} = -\bar{s} \quad (\text{A.17})$$

onde $\bar{c}$, $\bar{s}$ são calculados de acordo com (A.13), de forma a anular progressivamente todos os elementos do vetor a .

Desta forma, pode ser verificado que os esforços computacionais para obtenção de R_2 em (A.8) e (A.10) são iguais e correspondentes a

$$FLOPS = 4n^2 + 4n \quad (\text{A.18})$$

Referências Bibliográficas

- [1] Bierman, G. J. (1977), Factorization Methods for discrete Sequential Equations, Academic Press, New York.
- [2] Box, G. F. P. ; Jenkins, G. M. (1976) "Time Series Analysis Forecasting and Control".
- [3] Cioffi J. M.; T. Kailath (1984) "Fast Recursive Least Squares Transversal Filters for Adaptive Filtering". IEEE Transaction ASSP, vol. ASSP. 32 304-337.
- [4] George, A. and Heath, M. T. (1980) "Solution of Sparse Linear Least Squares using Givens rotations", Linear Algebra and its Applications, 34, pp. 69-83.
- [5] Gill, P. E. ; W. Murray e M. H. Wright (1991), " Numerical Linear Algebra and Optimization". Adison Wesley Publishing Co., Redwood City.
- [6] Goodwin, G. C. ; K. S. Sin (1984) "Adaptive Filtering Prediction and Control". Prentice Hall Inc., Englewood Cliffs, New Jersey, USA.
- [7] Houacine, A. (1990) "Regularized Fast Recursive Least Squares Algorithms" , Proceedings IEEE ICASSP 90 Conference, Albuquerque, New Mexico, USA, vol.3, 1587-1590.
- [8] Lawson, C. L. and Hanson, R. J. (1974) Solving Least Squares Problems, Prentice Hall, Englewood Cliffs, New Jersey.
- [9] Reichel, L. ; Gragg, W. B. (1990) "Algorithm 686, FORTRAN Subroutines for Updating the QR Decomposition" ACM Transactions on Mathematical Software, vol. 16, no. 4, 369-377.
- [10] Resende, L. S. (1991) "Filtragem Adaptativa com restrições lineares por mínimos quadrados", Dissertação de mestrado em engenharia elétrica, UNICAMP.
- [11] Resende, L. S.; J.M.T.Romano et M.G Bellanger, " A Fast Least Squares Algorithms for Constrained Adaptativa Filtering", ICASSP'92, San Francisco USA, Mar/1992.

- [12] Sorenson, H. W. (1980) Parameter Estimation , principles and problems, Marcel DekkerInc., La Jolla, San Diego, California.
- [13] Taneguti, L. Y.; Milani, B. E. A. (1992) "Um algoritmo Utilizando Fatoração QR para Identificação Adaptativa de Sistemas através de Mínimos Quadrados com Janela Móvel de Dados" Anais do XXIV SBPO, SOBRAPO, Salvador, Bahia(a ser publicado).
- [14] Wilkinson, J. H. (1965) The Algebraic Eigenvalue Problem , Oxford University Press.
- [15] Gill, P. E. ; W. Murray e M. H. Wright (1981) "Practical Optimizati-on", Academic Press, New York.
- [16] Taneguti, L. Y.(1993) "Identificação Adaptativa de Sistemas através de Mínimos Quadrados Lineares utilizando Fatoração QR e Janela Móvel de dados", Dissertação de Mestrado em Engenharia Elétrica, UNICAMP.
- [17] Gimenez, N. D.; Milani, B. E. A.(1994) "Um Algoritmo Numericamente Estável para Identificação Adaptativa de Sistemas através de Mínimos Quadrados com Restrições de Igualdade", Anais do XVII CNMAC, Vitória - ES.