

**WILSON FADLO CURI**

**ALGORITMO DA FUNÇÃO DE PENALIDADE EXATA GLOBALMENTE  
CONVERGENTE COM APLICAÇÕES, INCLUINDO O PROBLEMA  
DE FLUXO DE CARGA ÓTIMO**

Dissertação apresentada à Coordenação dos  
Cursos de Pós-Graduação em Engenharia Elé  
trica da Universidade Federal da Paraíba,  
em cumprimento às exigências para obten  
ção do Grau de Mestre em Engenharia Elé  
trica.

**ÁREA DE CONCENTRAÇÃO: SISTEMAS DE POTÊNCIA**

**ORIENTADOR: JANUSZ STANISLAW LIPONSKI**

**CAMPINA GRANDE - PB**

**AGOSTO - 1985**



C975a Curi, Wilson Fadlo.  
Algoritmo da função de penalidade exata globalmente convergente com aplicações, incluindo o problema de fluxo de carga ótimo / Wilson Fadlo Curi. - Campina Grande, 1985. 147 f.

Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal da Paraíba, Centro de Ciências e Tecnologia, 1985.

"Orientação : Prof. Dr. Janusz Stanislaw Lipowski".  
Referências.

1. Sistemas de Potência. 2. Algoritmo - Função de Penalidade. 3. Programação Matemática. 4. Fluxo de Carga Ótimo. 5. Engenharia Elétrica - Dissertação. I. Lipowski, Janusz Stanislaw. II. Universidade Federal da Paraíba - Campina Grande (PB) III. Título

CDU 621.3.016.2(043)

### AGRADECIMENTOS

Ao orientador, JANUSZ STANISLAW LIPOWSKI, pela orientação e incentivo; ao grupo de pessoas que fazem o Núcleo de Processamento de Dados pela colaboração na execução dos trabalhos computacionais; aos professores e funcionários do Departamento de Física pelo apoio e cooperação e a minha esposa, Rosires Catão Curi, pela compreensão e estímulo durante a execução deste trabalho.

## RESUMO

Este trabalho apresenta o algoritmo da função de penalidade exata globalmente convergente, baseado nos trabalhos de Colemann & Conn (1982a) e (1982b), que se destina a resolver o problema geral de programação matemática. O trabalho contém os resultados numéricos de testes feitos com vários problemas que representam diversos níveis de dificuldade, incluindo o problema de fluxo de carga ótimo associado a sistema de potência de 23 barras.



### CONTRIBUIÇÃO DO AUTOR

Os seguintes resultados, apresentados neste trabalho, podem ser considerados como contribuição original do autor para a comunidade científica brasileira:

- i - apresentação da teoria do algoritmo da função de penalidade exata globalmente convergente [Coleman & Conn (1980 b)] em forma compacta e em idioma português;
- ii - compilação e apresentação das técnicas computacionais especiais indispensáveis para o desenvolvimento do algoritmo computacional robusto e eficiente;
- iii - desenvolvimento do algoritmo computacional, que se baseia na teoria exposta, e sua implementação no computador incorporando as técnicas computacionais especiais;
- iv - verificação de desempenho do algoritmo computacional com vários problemas de teste incluindo o pro

blema de fluxo de carga ótimo associado ao sistema de potência de 23 barras.

## LISTA DE SÍMBOLOS E RELAÇÕES BÁSICAS

### 1. Escalares:

$\mu$	- escalar positivo
$\lambda$	- pequeno escalar positivo
$\epsilon$	- pequeno escalar positivo
$\delta$	- pequeno escalar positivo
$n$	- número de variáveis
$m_1$	- número de restrições de igualdade
$m_2$	- número de restrições de desi- gualdade
$m$	- número total de restrições

$t_1^k$	- número de restrições $\epsilon$ -ativas de igualdade na $k$ -ésima iteração
$t_2^k$	- número de restrições $\epsilon$ -ativas de desigualdade na $k$ -ésima iteração
$t^k$	- número total de restrições $\epsilon$ -ativas
$\lambda_{1i}$	- multiplicador de Lagrange associado a $i$ -ésima restrição $\epsilon$ -ativa de igualdade
$\lambda_{2i}$	- multiplicador de Lagrange associado a $i$ -ésima restrição $\epsilon$ -ativa de desigualdade
$\lambda_i$	- multiplicador de Lagrange associado a $i$ -ésima restrição $\epsilon$ -ativa
$f(x)$	- função objetivo
$\phi_i(x), i = 1, \dots, m_1$	- função que define a $i$ -ésima restrição de igualdade

$\phi_i(x)$ ,  $i = m_1 + 1, \dots, m$  - função que define a  $i$ -ésima restrição de desigualdade

$$p(x, \mu) = p(x) = f(x) + \frac{1}{\mu} \sum_{i=1}^{m_1} |\phi_i(x)| - \frac{1}{\mu} \sum_{i=m_1+1}^m \min(0, \phi_i(x))$$

- função de penalidade exata

$L(x, \lambda) = f(x) - \sum_{i \in I_A} \lambda_i \phi_i(x)$  - função Lagrangeana ordinária associada ao problema geral de programação matemática

$L_{aum}(x, \lambda)$  - função Lagrangeana aumentada

$$\begin{aligned} L(h, \lambda) &= \nabla p_1(x)^T h + \frac{1}{2} h^T \nabla^2 p_1(x) h \\ &+ \sum_{i \in I_{A1}^e} \lambda_{1i} (\nabla \phi_i(x)^T h + \frac{1}{2} h^T \nabla^2 \phi_i(x) h) \\ &- \sum_{i \in I_{A2}^e} \lambda_{2i} (\nabla \phi_i(x)^T h + \frac{1}{2} h^T \nabla^2 \phi_i(x) h) \end{aligned}$$

- função Lagrangeana associado ao problema que utiliza termos de segunda ordem para minimizar  $p(x)$

## 2. Conjuntos:

$$I_{M1} = \{i | i = 1, \dots, m_1\}$$

- conjunto de índices associados as restrições de igualdade

$$I_{M2} = \{i \mid i = m_1 + 1, \dots, m_2\}$$

- conjunto de índices associados as restrições de desigualdade

$$I_M = I_{M1} \cup I_{M2}$$

- conjunto de índices associados a todas as restrições

$$I_{A1}^\epsilon = \{i \mid |\phi_i(x)| \leq \epsilon, i \in I_{M1}\}$$

- conjunto de índices associados as restrições  $\epsilon$ -ativas de igualdade

$$I_{A2}^\epsilon = \{i \mid \phi_i(x) \leq \epsilon, i \in I_{M2}\}$$

- conjunto de índices associados as restrições  $\epsilon$ -ativas de desigualdade

$$I_A^\epsilon = I_{A1}^\epsilon \cup I_{A2}^\epsilon$$

- conjunto de índices associados as todas as restrições  $\epsilon$ -ativas

$$I_{V1}^\epsilon = \{i \mid |\phi_i(x)| > \epsilon, i \in I_{M1}\}$$

- conjunto de índices associados as restrições claramente violadas de igualdade

$$I_{V2}^\epsilon = \{i \mid \phi_i(x) < -\epsilon, i \in I_{M2}\}$$

- conjunto de índices associados as restrições claramente violadas de desigualdade

$$I_V^\epsilon = I_{V1}^\epsilon \cup I_{V2}^\epsilon$$

- conjunto de índices associados a todas restrições claramente violadas

### 3. Vetores:

$$x = [x_1, \dots, x_n]^T$$

- vetor de variáveis com dimensão  $n \times 1$

$$\nabla f(x)$$

- vetor gradiente da função objetivo com dimensão  $n \times 1$

$$\nabla \phi_i(x), i \in I_M$$

- vetor gradiente da  $i$ -ésima restrição com dimensão  $n \times 1$

$$\lambda^k = [-\lambda_{1_1}, \dots, -\lambda_{1_{t_1^k}}, \lambda_{2_1}, \dots, \lambda_{2_{t_2^k}}]^T$$

- vetor dos multiplicadores de Lagrange, associados as restrições  $\epsilon$ -ativas, com dimensão  $t^k \times 1$

$$\begin{aligned} \nabla p_1(x) = \nabla f(x) + \frac{1}{\mu} \sum_{i \in I_{V1}^\epsilon} \text{sinal}(\phi_i(x)) \nabla \phi_i(x) \\ - \frac{1}{\mu} \sum_{i \in I_{V2}^\epsilon} \nabla \phi_i(x) \end{aligned}$$

- vetor pseudo-gradiente com dimensão  $n \times 1$



$$h = Z(Z^T H Z)^{-1} Z^T \nabla p_1(x)$$

- vetor com dimensão  $n \times 1$  que representa a direção de pesquisa horizontal para aproximação com polinômios de primeira ordem

$$\tilde{h} = \sigma_j Z_{k \setminus j} Z_{k \setminus j}^T \nabla \phi_j(x)$$

- vetor com dimensão  $n \times 1$  que representa a direção de pesquisa horizontal para aproximação com polinômios de primeira ordem

$$v = -A(A^T A)^{-1} \phi(x+h)$$

- vetor com dimensão  $n \times 1$  que representa a direção de pesquisa vertical

#### 4. Matrizes:

$$A_k = \left[ \nabla \phi_1(x), \dots, \nabla \phi_{t^k}(x) \right] = Q_k^T \begin{bmatrix} R_k \\ 0 \end{bmatrix}$$

- matriz dos gradientes das restrições ativas com dimensão  $n \times t^k$

H

- matriz Hessiana positiva definida de ordem  $n$

$$H_{p_k} = Z_k^T H_k Z_k = L_k D_k L_k^T$$

- matriz Hessiana projetada de ordem  $(n - t^k)$

$$Q_k = \begin{bmatrix} Q_{1k} \\ Q_{2k} \end{bmatrix}$$

$Q_{1k}$

$Q_{2k}$

$R_k$

$Z_k$

$a_{i,*}$

$a_{*,i}$

### 5. Outros:

$S_I^E$

- matriz ortogonal de ordem  $n$

- submatriz superior da matriz ortogonal com dimensão  $t^k \times n$

- submatriz inferior da matriz ortogonal com dimensão  $(n-t^k) \times n$

- matriz triangular superior de ordem  $t^k$

- matriz com dimensão  $n \times (n-t^k)$ , cujas colunas são linearmente independentes e que satisfaz a condição:  $A_k^T Z_k = 0$  e  $Z^T Z = I$

-  $i$ -ésima linha da matriz  $A = (a_{i,j})$

-  $i$ -ésima coluna da matriz  $A = (a_{i,j})$

- subespaço gerado pelos gradientes das restrições  $\epsilon$ -ativas linearmente independente

$H_u$ 

- matriz de Hessenberg superior de ordem  $h$

 $H_L$ 

- matriz de Hessenberg inferior de ordem  $h$

 $I$ 

- matriz identidade

 $\bar{I}$ 

- matriz identidade com as colunas na ordem reversa

 $L_k$ 

- matriz triangular inferior unitária de ordem  $(n-t^k)$

 $D_k$ 

- matriz diagonal de ordem  $(n-t^k)$

 $P$ 

- matriz que representa o projetor ortogonal e que satisfaz as relações:  $P^2=P$  e  $P=P^T$

 $P_j^i$ 

- matriz de Givens de ordem  $h$  embutida na matriz identidade (quando aplicada a um vetor  $z = [z_1, \dots, z_\ell]^T$ , reduz o elemento  $z_j$  a zero pela combinação deste elemento com o elemento  $z_i$  e modifica o elemento  $z_i$ ).

$(S_{\perp}^{\varepsilon})^{\perp}$

- subespaço ortogonal de  $S_{\perp}^{\varepsilon}$

\*

- Índice que denota o valor  $\underline{0}$   
timo de um escalar ou de um  
vetor

## SUMÁRIO

	PÁGII
<b>1 INTRODUÇÃO</b> .....	1
1.1 Objeto deste Trabalho .....	1
1.2 O Problema Geral de Programação Matemática .....	2
1.3 Alguns Conceitos Matemáticos .....	3
1.4 Condições Necessárias e Suficientes Para a Existência de Uma Solução do Problema Geral de Programação Matemática .....	13
1.5 Transformação do Problema Geral de Programação Matemática em Um Problema de Minimização de Uma Função de Penalidade Exata e a Relação Entre Suas Soluções .....	14

1.6 Condições Necessárias e Suficientes Para a Existência de Uma Solução do Problema de Minimização sem Restrições de uma Função de Penalidade Exata .....	16
1.7 Descrição Geral do Algoritmo da Função de Penalidade Exata Globalmente Convergente.	19
1.8 O Problema de Fluxo de Carga Ótimo .....	21
<b>2 O ALGORITMO DA FUNÇÃO DE PENALIDADE EXATA GLOBALMENTE CONVERGENTE .....</b>	<b>24</b>
2.1 A Função de Penalidade Exata .....	24
2.2 A Direção de Pesquisa Horizontal .....	26
2.2.1 Um Método de Primeira Ordem para Minimizar $p(x)$ .....	26
2.2.2 Um Método de Segunda Ordem para Minimizar $p(x)$ .....	29
2.3 Os Multiplicadores de Lagrange .....	36
2.4 A Direção de Pesquisa Vertical .....	40

	PÁGINA
2.5 Queda de Uma Restrição .....	44
2.6 Uma Estratégia da Escolha da Direção de Pesquisa .....	46
2.7 Pesquisa Linear ao Longo da Direção de Pesquisa .....	49
2.8 Atualização da Matriz Hessiana $H_k$ ou da Matriz Hessiana Projetada $H_{p_k}$ .....	56
2.8.1 Uma Região Distante do Ponto de Solução $x^*$ .....	56
2.8.2 Nas Proximidades do Ponto de Solução $x^*$ ..	58
2.9 Algumas Considerações Sobre o Algoritmo .	60
2.10 Critérios de Convergência .....	62
<b>3 OPERAÇÕES COMPUTACIONAIS RELEVANTES AO DESENVOLVIMENTO DO ALGORITMO COMPUTACIONAL .....</b>	<b>64</b>
3.1 Fatoração Ortogonal da Matriz dos Gradientes das Restrições Ativas $A_k$ .....	64



	PÁGINA
3.2 Atualização da Fatoração Ortogonal da Matriz dos Gradientes das Restrições Ativas $A_k$ .....	69
3.2.1 Inclusão de Uma Coluna na Matriz $A_k$ .....	69
3.2.2 Exclusão de Uma Coluna da Matriz $A_k$ .....	74
3.3 A Escolha da Matriz $Z_k$ .....	76
3.4 Atualização da Matriz $Z_k$ .....	79
3.4.1 Inclusão de Uma Coluna na Matriz $A_k$ .....	79
3.4.2 Exclusão de Uma Coluna da Matriz $A_k$ .....	80
3.5 Fatoração $LDL^T$ de Uma Matriz Simétrica Positiva Definida $H_p$ .....	81
3.6 Fatoração $LDL^T$ Modificada de Uma Matriz Simétrica $H_p$ .....	82
3.7 Atualização da Fatoração $LDL^T$ da Matriz Hessiana Projetada .....	87
3.7.1 Atualização da Fatoração $LDL^T$ Com Diminuição da Dimensão da Matriz $H_p$ .....	88

	PÁGINA
3.7.2 Atualização da Fatoração $LDL^T$ Com Aumento da Dimensão da Matriz $H_p$ .....	94
3.7.3 Atualização da Fatoração $LDL^T$ Sem Mudança de Dimensão da Matriz $H_p$ .....	96
3.8 Correção da Fatoração $LDL^T$ Com Uma Matriz de Posto Um .....	96
<b>4 SUMÁRIO DO ALGORITMO DA FUNÇÃO DE PENALIDADE EXATA GLOBALMENTE CONVERGENTE</b> .....	100
<b>5 RESULTADOS NUMÉRICOS</b> .....	106
<b>6 CONCLUSÕES</b> .....	113
<b>BIBLIOGRAFIA</b> .....	116
<b>APÊNDICE 1 - PROBLEMA TESTE DE ROSEN-SUZUKI [ROSEN &amp; SUZUKI (1965)] COM 4 VARIÁVEIS E 3 RESTRIÇÕES DE DESIGUALDADE</b> .....	125
<b>APÊNDICE 2 - PROBLEMA TESTE DE POWELL [POWELL (1978a)] COM 5 VARIÁVEIS E 3 RESTRIÇÕES DE IGUALDADE</b> .....	127

	PÁGINA
<b>APÊNDICE 3</b> - PROBLEMA TESTE DE WONG [ ASAADI (1973) ] COM 10 VARIÁVEIS E 8 RESTRI ÇÕES DE DESIGUALDADE .....	129
<b>APÊNDICE 4</b> - SÍMBOLOS, RELAÇÕES BÁSICAS E CIRCUITOS EQUIVALENTES RELEVANTES PARA A DESCR ÇÃO DE UM SISTEMA DE POTÊNCIA E PARA A DEFINIÇÃO DE UM PROBLEMA DE FLUXO DE CARGA ÓTIMO CORRESPONDENTE .....	131
<b>APÊNDICE 5</b> - SISTEMA DE POTÊNCIA COM 23 BARRAS, 30 LINHAS E 15 GERADORES E O RESPECTIVO PROBLEMA DE FLUXO DE CARGA ÓTIMO .....	139

## 1 INTRODUÇÃO

### 1.1 Objeto Deste Trabalho

Este trabalho apresenta o algoritmo da função de penalidade exata globalmente convergente, baseado, principalmente, nos trabalhos de Coleman & Conn (1982a) e (1982b), que pode ser utilizado para resolver o problema geral de programação matemática, ou seja, o problema de minimização de uma função objetivo não linear sujeita a restrições não lineares de igualdade e de desigualdade.

Muitas áreas de ciências e de engenharia, frequentemente enfrentam a necessidade de resolver o problema geral de programação matemática nas várias formas. Um exemplo importante de aplicação pode ser encontrado na área de sistemas de potência, onde a solução confiável e eficiente do problema de fluxo de carga ótimo, que é um problema geral de programação matemática, pode diminuir o custo da energia elétrica e o custo de operação das empresas de geração de energia elétrica. Vários pesquisadores tem aplicado várias técnicas para resolver o problema de fluxo de carga ótimo [Lipowski

(1980a), Muklerjee & Dhor (1974), Sasson (1969), Sasson et alii (1972)]. As técnicas mais recentes e populares de programação matemática não-linear são os métodos de programação quadrática recursiva [Biggs (1975), Han (1976), Han (1977a), Han (1977b), Powell (1969), Powell (1978a)] e como exemplos de aplicação desses métodos para a solução do problema de fluxo de carga ótimo, podem-se citar os trabalhos de Biggs & Laughton (1977) e Lipowski (1980b) e (1981).

A estrutura do trabalho é como segue: o capítulo 2 apresenta a teoria associada ao algoritmo da função de penalidade exata globalmente convergente; o capítulo 3 mostra os detalhes das operações computacionais importantes para o desenvolvimento do algoritmo computacional; o capítulo 4 apresenta todos os passos do algoritmo computacional; e o capítulo 5 contém os resultados de testes feitos com quatro problemas de teste que incluem o problema de fluxo de carga ótimo associado a um sistema de potência de 23 barras e no capítulo 6 se encontram as conclusões.

## 1.2 O Problema Geral de Programação Matemática

O problema geral de programação matemática pode ser expresso da seguinte forma:

$$\underset{x}{\text{minimize}} \quad f(x) \quad , \quad x \in R^n \quad (1.1)$$

$$\text{sujeito a} \quad \phi_i(x) = 0 \quad , \quad i = 1, \dots, m_1$$

$$\phi_i(x) \geq 0 \quad , \quad i = m_1+1, \dots, m$$



onde

$n$	- número de variáveis;
$m_1$	- número de restrições de igualdade;
$m$	- número total de restrições;
$x = [x_1, \dots, x_n]^T$	- vetor de variáveis de ordem $n \times 1$ ;
$f(x)$	- função objetivo;
$\phi_i(x)$	- função que define a $i$ -ésima restrição.

### 1.3 Alguns Conceitos Matemáticos

Nesta seção são apresentados alguns conceitos matemáticos básicos indispensáveis ao bom entendimento da teoria associada aos algoritmos de programação matemática não-linear.

**a) Convergência Global** - Um algoritmo iterativo é globalmente convergente quando, para qualquer ponto inicial  $x^1$ , gera uma sequência de pontos  $\{x^k\}$  que converge para um mínimo local  $x^*$ .

**b) Taxas de Convergência** - A sequência de pontos  $\{x^k\} \subseteq \mathbb{R}^n$  convergindo para um ponto  $x^*$  é dito convergir:

(i) - **linearmente**, se existe uma constante  $L$ ,  $0 < L < 1$ , e um inteiro positivo  $k_0 > 1$  tal que

$$\|x^{k+1} - x^*\| \leq L \|x^k - x^*\|, \text{ para todo } k \geq k_0 \quad (1.2)$$

(ii) - **superlinearmente**, se

$$\frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} \rightarrow 0, \text{ quando } k \rightarrow \infty, \quad (1.3)$$

(iii) - **quadraticamente**, se existe uma constante  $L$ ,  $L > 0$ , e um inteiro  $k_0 > 1$ , tal que

$$\|x^{k+1} - x^*\| \leq L \|x^k - x^*\|^2, \text{ para todos } k \geq k_0, \quad (1.4)$$

(iv) - **superlinearmente com passo 2**, se existe uma constante  $L$ ,  $L > 0$  e um inteiro  $k_0 > 1$ , tal que

$$\|x^{k+1} - x^*\| \leq L \|x^{k-1} - x^*\|^2, \text{ para todo } k \geq k_0. \quad (1.5)$$

Neste trabalho  $\|\cdot\|$  indica a norma-2, salvo indicação em contrário. Para maiores detalhes sobre taxa de convergência veja Ortega & Rheinbolt (1970).

**c) Expansão de Taylor** - Suponha que  $f(x)$  seja uma função uma vez continuamente diferenciável em  $R^n$ . Então para alguns  $x, x' \in R^n$ , a expansão de Taylor é definida como sendo

$$f(x) = f(x') + \nabla f(\xi)^T (x - x'), \quad (1.6)$$

onde

$$\xi = x' - \theta(x - x'), \quad 0 \leq \theta \leq 1,$$



que é equivalente a

$$f(x) = f(x') + \nabla f(x')^T (x - x') + o(\|x - x'\|) , \quad (1.7)$$

onde  $o(\alpha)$  está para uma expressão que satisfaça  $o(\alpha) / \alpha \rightarrow 0$ , quando  $\alpha \rightarrow 0$ .

Se  $f(x)$  é duas vezes continuamente diferenciável, então

$$f(x) = f(x') + \nabla f(x')^T (x - x') + \frac{1}{2} (x - x')^T \nabla^2 f(\xi) (x - x') , \quad (1.8)$$

onde

$$\xi = x' + \theta(x - x') , \quad 0 \leq \theta \leq 1 ,$$

que é equivalente a

$$f(x) = f(x') + \nabla f(x')^T (x - x') + \frac{1}{2} (x - x')^T \nabla^2 f(x') (x - x') + o(\|x - x'\|^2) . \quad (1.9)$$

Gill & Murray (1974a) e Zangwill (1967) são citados para maiores esclarecimentos.

**d) A Função Lagrangeana** - A função Lagrangeana ordinária (função Lagrangeana de Kuhn-Tucker) associada com o problema geral de programação matemática é:

$$L(x, \lambda) = f(x) - \sum_{i=1}^m \lambda_i \phi_i(x) \quad , \quad (1.10)$$

onde  $\lambda_i$  é o multiplicador de Lagrange associado a  $i$ -ésima restrição.

O vetor das derivadas parciais da função Lagrangeana com respeito a  $x$  é:

$$\nabla_x L(x, \lambda) = \nabla f(x) - \sum_{i=1}^m \lambda_i \nabla \phi_i(x) \quad , \quad (1.11)$$

A matriz das derivadas parciais de segunda ordem da função Lagrangeana com respeito a  $x$  é:

$$\nabla_{xx} L(x, \lambda) = \nabla^2 f(x) - \sum_{i=1}^m \lambda_i \nabla^2 \phi_i(x) \quad , \quad (1.12)$$

Existe uma extensa literatura sobre o uso de funções Lagrangeanas, por exemplo, Avriel (1976), Fiacco & McCormick (1968), Hestenes (1969), Miele et alli (1971), Miele et alli (1972), Osbourne (1972), Polyak & Tret'yakov (1973), Powell (1969), Powell (1978a) e Tapia (1974).

**e) Um Ponto Viável** - Um ponto  $x \bar{e}$  (estritamente) viável para o problema (1.1), quando satisfaz

$$\phi_i(x) = 0 \quad , \quad i = 1, \dots, m_1 \quad ,$$

$$\phi_i(x) \geq 0 \quad , \quad i = m_1 + 1, \dots, m.$$

f) Um Ponto Regular - Um ponto  $x^0$  é dito ser um ponto regular para o conjunto de restrições se:

(i) - este ponto é um ponto viável para o conjunto de restrições, e

(ii) - os gradientes das restrições de igualdade e restrições de desigualdade ativas neste ponto,

$$(\nabla \phi_i(x^0) \quad , \quad i \in \{ \phi_i(x^0) = 0 \quad , \quad 1 \leq i \leq m \} ) \quad ,$$

são linearmente independentes McCormick (1967)

Esta condição de regularidade é uma condição suficiente para que a qualificação das restrições seja satisfeita no ponto  $x^0$ .

g) Restrições Aproximadamente Ativas e Claramente Violadas - Fazendo  $I_{M1} = \{1, 2, \dots, m_2\}$  e  $I_{M2} = \{m_1 + 1, \dots, m\}$ , os conjuntos de índices das restrições de igualdade e de desigualdade ativas no ponto  $x$  são, respectivamente,

$$I_{A1} = \{i \in I_{M1} \mid \phi_i(x) = 0\} \quad ,$$

$$I_{A2} = \{i \in I_{M2} \mid \phi_i(x) = 0\} \quad .$$

Similarmente, se  $\epsilon > 0$  é um pequeno escalar positivo, o conjunto de índices de restrições aproximadamente ativas

( $\epsilon$ -ativas) para as restrições de igualdade e de desigualdade e claramente violadas ( $\epsilon$ -violadas) para as restrições de igualdade e de desigualdade, são respectivamente:

$$I_{A1}^{\epsilon} = \{i \in I_{M1} \mid |\phi_i(x)| \leq \epsilon\} , \quad (1.13)$$

$$I_{A2}^{\epsilon} = \{i \in I_{M2} \mid |\phi_i(x)| \leq \epsilon\} ,$$

$$I_{V1}^{\epsilon} = \{i \in I_{M1} \mid |\phi_i(x)| > \epsilon\} ,$$

$$I_{V2}^{\epsilon} = \{i \in I_{M2} \mid \phi_i(x) < -\epsilon\} .$$

**h) O Mínimo Global** - Denotando a região viável para o problema (1.1) por  $F$ , pode-se dizer que  $x^*$  é um mínimo global de  $f(x)$  em  $F$  se

$$f(x^*) \leq f(x) , \quad \text{para todo } x \in F ,$$

e que  $x^*$  é um mínimo global forte de  $f(x)$  em  $F$  se

$$f(x^*) < f(x) , \quad \text{para todo } x \in F .$$

**i) Um Mínimo Local** - Um ponto  $x^*$  é dito ser um mínimo local se existe um escalar positivo  $\delta$ , tal que

$$f(x^*) \leq f(x) , \quad \text{para todo } x \in N(x^*, \delta) \cap F ,$$

onde  $N(x^*, \delta)$  é uma vizinhança do ponto  $x^*$  definida como

$$N(x^*, \delta) = \{x \mid \|x - x^*\| \leq \delta, x \neq x^*\}.$$

Similarmente,  $x^*$  é um mínimo local isolado (ou forte) Fiacco & McCormick (1968) se

$$f(x^*) < f(x) \quad , \quad \text{para todo } x \in N(x^*, \delta) \quad F.$$

j) **Um Projetor Ortogonal** - Seja  $A$  uma matriz de dimensão  $n \times t$  ( $t < n$ ) com colunas linearmente independentes, então a matriz de ordem  $n$ ,

$$P = I - A(A^T A)^{-1} A^T, \quad (1.14)$$

é um projetor ortogonal, onde  $I$  denota uma matriz identidade de ordem  $n$ . Se  $w \in \mathbb{R}^n$  e  $w = A\bar{w}$  para algum  $\bar{w} \in \mathbb{R}^t$ , então  $Pw = 0$ . Alternativamente, se  $w \in \mathbb{R}^n$  e  $A^T w = 0$ , então  $Pw = w$ . Em adição  $P$  satisfaz as propriedades [ Ben & Greville (1974), Hanson (1974), Stewart (1973) e Conn (1976) ]

$$P^2 = P \quad \text{e} \quad (1.15)$$

$$P = P^T.$$

k) **Fatoração Ortogonal**  $Q^T \begin{bmatrix} R \\ \bar{0} \end{bmatrix}$  de uma Matriz Retangular  $A$  de dimensão  $n \times t$ , ( $n > t$ ), **com Colunas Linearmente Independentes, (posto  $(A) = t$ )** - A fatoração ortogonal  $Q^T \begin{bmatrix} R \\ \bar{0} \end{bmatrix}$  de uma matriz retangular  $A$  de dimensão  $n \times t$  ( $n < t$ ) com colunas linearmente independentes [ (posto  $(A) = t$ ) ] é da forma

$$Q^T \begin{vmatrix} R \\ 0 \end{vmatrix} = A \quad \text{ou} \quad (1.16)$$

$$QA = \begin{vmatrix} R \\ 0 \end{vmatrix}$$

onde

- $Q = \begin{vmatrix} Q_1 \\ Q_2 \end{vmatrix}$  - matriz ortogonal de ordem  $n$  não única;  
 $Q_1$  - submatriz superior da matriz ortogonal  $Q$  com dimensões  $t \times n$  única;  
 $Q_2$  - submatriz inferior da matriz ortogonal  $Q$  com dimensões  $(n - t) \times n$  não única;  
 $R$  - matriz triangular superior de ordem  $t$

Esta fatoração é utilizada para

- solução eficiente de um sistema hiperdeterminado de equações:  $Ax = b$ ,
- solução eficiente de um sistema hipodeterminado de equações:  $A^T x = b$ ,
- determinação da matriz  $Z \equiv Q_2^T$  cujas colunas formam uma base do subespaço nulo da matriz  $A^T$ , tal que  $A^T Z = 0$  e  $Z^T Z = I$ .

Para maiores detalhes veja Stewart (1973).

1) Fatoração por Valores Singulares  $U \Sigma V^T$  de uma Matriz  $A$  com dimensão  $n \times t$ , ( $n > t$ ), com Colunas Linearmente Dependentes ou Linearmente Independentes, (posto  $(A) \leq t$ ) -



A fatoração do valor singular  $U \Sigma V^T$  de uma matriz  $A$  de ordem  $n \times t$ , ( $n > t$ ), com colunas linearmente dependentes ou linearmente independentes ( $\text{posto}(A) \leq t$ ) é da forma

$$U \Sigma V^T = A, \quad (1.17)$$

onde

$U$  - matriz ortogonal de ordem  $n$  não única;

$V^T$  - matriz ortogonal de ordem  $t$  não única;

$\Sigma = \begin{bmatrix} D & \vdots & 0 \\ \dots & \vdots & \dots \\ 0 & \vdots & 0 \end{bmatrix}$  - matriz retangular com dimensão  $n \times t$  única;

$D = \text{diag}(\sigma_i)$  - matriz diagonal de ordem  $\text{posto}(A) \times \text{posto}(A)$  com valores singulares  $\sigma_i$  na diagonal

$\sigma_i$  - valores singulares (parâmetros positivos únicos) associados a matriz  $A$ .

Esta fatoração deve ser utilizada nas mesmas situações em que a fatoração ortogonal -  $Q^T \frac{R}{0}$ , quando a matriz  $A$  tem posto deficiente ( $\text{posto}(A) < t$ ) [Hanson (1974)].

**n) Fatoração  $LDL^T$  de uma Matriz Simétrica Positiva Definida  $H_p$**  - A fatoração  $LDL^T$  (fatoração de Cholesky) de uma matriz simétrica positiva definida  $H_p$  é da forma

$$LDL^T = H_p \quad (1.18)$$



onde

L - matriz triangular inferior unitária;

D - matriz diagonal com elementos diagonais positivos.

Esta fatoração é utilizada para a solução eficiente de um sistema determinado de equações  $H_p x = b$ , especialmente quando a solução precisa ser computada várias vezes para a mesma matriz  $H_p$  e vários vetores  $b$ .

Na seção 3.5 é mostrado o algoritmo computacional para executar esta fatoração [Martin, Peters e Wilkinson (1965)].

**o) Fatoração  $LDL^T$  Modificada de uma Matriz Simétrica Positiva Definida ou Positiva Semi-Definida ou Indefinida**

$H_p$  - A fatoração  $LDL^T$  modificada (fatoração modificada de Cholesky) de uma matriz simétrica definida ou positiva ou semi-definida ou indefinida  $H_p$  é da forma:

$$LDL^T - E = H_p, \quad (1.19)$$

onde

L - matriz triangular inferior unitária;

D - matriz diagonal com elementos diagonais positivos;

E - matriz diagonal com elementos diagonais:

- . iguais a zero, quando a matriz  $H_p$  é positiva definida;
- . não negativos (incluindo alguns elementos positivos)

quando  $H_p$  é positiva semi-definida ou indefinida.

Esta fatoração é utilizada para:

(i) - solução exata de um sistema determinado de equações,  $H_p x = b$ , quando a matriz  $A$  é simétrica positiva definida;

(ii) - solução aproximada de um sistema determinado de equações,  $H_p x = b$ , quando a matriz  $A$  é positiva semi-definida ou indefinida (neste caso se resolve o sistema  $LDL^T x = b$  no lugar do sistema original  $H_p x = b$ ).

Na seção 3.6 é mostrado o algoritmo computacional para executar esta fatoração [Gill & Murray (1974b) e Lipowski (1980b)].

#### 1.4 Condições Necessárias e Suficientes para a Existência de uma Solução do Problema Geral de Programação Matemática

a) **Condições Necessárias de Primeira Ordem** - Se  $x^*$  é um ponto de mínimo local para o problema (1.1) e um ponto regular para o conjunto de restrições, então existe um vetor dos multiplicadores de Lagrange  $\lambda^*$  tal que:

$$\nabla_x L(x^*, \lambda^*) = 0 \quad (1.20)$$

$$\lambda_i^* \phi_i(x^*) = 0, \quad i = m_1 + 1, \dots, m,$$

$$\lambda_j^* \geq 0, \quad j = m_1 + 1, \dots, m.$$

b) **Condições de Suficiência de Segunda Ordem** - Se  $x^*$  é um ponto de mínimo local forte para o problema (1.1) e um ponto regular para o conjunto de restrições, então existe um vetor dos multiplicadores de Lagrange  $\lambda^*$  tal que [Charalambous (1979)]:

(i) - as condições necessárias de primeira ordem são satisfeitas, e

(ii) -  $z^T \nabla_{XX} L(x^*, \lambda^*) z > 0$ , para qualquer  $z \neq 0$  tal que  $z \in M$ , onde

$$M = \{z \mid z^T \nabla \phi_i(x^*) = 0, \quad i = 1, \dots, m_1,\}$$

$$z^T \nabla \phi_i(x^*) = 0, \quad i \in J_1,$$

$$z^T \nabla \phi_i(x^*) \geq 0, \quad i \in J_2 \quad \} \quad (1.21)$$

$$J_1 = \{i \mid \phi_i(x^*) = 0, \quad \lambda_i^* > 0, \quad m_1 + 1 \leq i \leq m\},$$

$$J_2 = \{i \mid \phi_i(x^*) = 0, \quad \lambda_i^* = 0, \quad m_1 + 1 \leq i < m\}.$$

### 1.5 Transformação do Problema Geral de Programação Matemática em um Problema de Minimização de uma Função de Penalidade Exata e a Relação entre suas Soluções

O problema (1.1) pode ser transformado em um problema

de minimização irrestrita de uma função de penalidade exata  $p(x, \mu)$

$$\underset{x}{\text{minimize}} \quad p(x, \mu) = f(x) + \frac{1}{\mu} \sum_{i=1}^{m_1} |\phi_i(x)| - \frac{1}{\mu} \sum_{i=m_1+1}^m \min(0, \phi_i(x)) \quad ,$$

$$x \in \mathbb{R}^n \quad , \quad (1.22)$$

onde  $\mu$  é algum parâmetro positivo.

Se o parâmetro  $\mu$  é suficientemente pequeno, os mínimos locais do problema (1.22) são também as soluções do problema (1.1) [Pietrzykowski (1969)]. Este fato justifica o porquê da função de penalidade  $p(x, \mu)$  se chamar de função de penalidade exata.

Podem-se fazer as seguintes observações sobre a função de penalidade exata:

(i) - o nível de complexidade das restrições do problema (1.22) corresponde ao nível de complexidade das restrições do problema (1.1);

(ii) - não é necessário aproximar  $\mu$  de zero para se obter uma solução precisa para o problema (1.1), como seria o caso quando a função de penalidade clássica é usada.

### 1.6 Condições Necessárias e Suficientes para a Existência de uma Solução do Problema de Minimização sem Restrições de uma Função de Penalidade Exata

Definem-se os conjuntos de índices das restrições ativas de igualdade ( $I_{A1}$ ) e de desigualdade ( $I_{A2}$ ), das restrições violadas de igualdade ( $I_{V1}$ ) e de desigualdade ( $I_{V2}$ ) como segue:

$$I_{A1} = \{i \mid \phi_i(x) = 0, \quad i = 1, \dots, m_1\}, \quad (1.23)$$

$$I_{A2} = \{i \mid \phi_i(x) = 0, \quad i = m_1 + 1, \dots, m\},$$

$$I_{V1} = \{i \mid |\phi_i(x)| \neq 0, \quad i = 1, \dots, m_1\},$$

$$I_{V2} = \{i \mid \phi_i(x) < 0, \quad i = m_1 + 1, \dots, m\},$$

e

$$I_A = I_{A1} \cup I_{A2}, \quad (1.24)$$

$$I_V = I_{V1} \cup I_{V2}.$$

a) **Condições Necessárias de Primeira Ordem** - Assumindo-se que  $f(x)$  e  $\phi_i(x)$ ,  $i = 1, \dots, m$ , são continuamente diferenciáveis e que  $\{\nabla\phi_i(x) \mid i \in I_A(x)\}$  é um conjunto de gradientes linearmente independentes, então as condições necessárias de primeira ordem para que  $x^*$  seja um mínimo local de

$p(x, \mu)$  são:

(i) . existe  $\lambda$  tal que

$$\begin{aligned} \nabla p_1(x^*) &= \nabla f(x^*) + \frac{1}{\mu} \sum_{i \in I_{V1}} \text{sign}(\phi_i(x^*)) \nabla \phi_i(x^*) - \frac{1}{\mu} \sum_{i \in I_{V2}} \nabla \phi_i(x^*) \\ &= - \sum_{i \in I_{A1}} \lambda_{1i} \nabla \phi_i(x^*) + \sum_{i \in I_{A2}} \lambda_{2i} \nabla \phi_i(x^*) ; \end{aligned} \quad (1.25)$$

$$(ii) . - \frac{1}{\mu} \leq \lambda_{1i} \leq \frac{1}{\mu} , \quad i \in I_{A1} ;$$

$$(iii) . 0 \leq \lambda_{2i} \leq \frac{1}{\mu} , \quad i \in I_{A2} ;$$

onde  $\lambda_{1i}$ ,  $i \in I_{A1}$ , e  $\lambda_{2i}$ ,  $i \in I_{A2}$ , são, respectivamente os multiplicadores de Lagrange associados as restrições ativas de igualdade e de desigualdade.

**b) Condições Necessárias de Segunda Ordem** - Assumindo-se que  $(f(x^*)$  e  $\phi_i(x^*)$ ,  $i = 1, \dots, m$ , são duas vezes continuamente diferenciáveis e que  $\{\nabla \phi_i(x^*) \mid i \in I_A(x^*)\}$  é um conjunto de gradientes linearmente independentes, então as condições necessárias de segunda ordem para que  $x^*$  seja um mínimo local de  $p(x, \mu)$  são:

(i) - sejam satisfeitas as condições necessárias de primeira ordem; (1.26)

(ii) - para todo vetor  $y$  que satisfaça  $y^T \nabla \phi_i(x^*) = 0$ , onde  $i \in I_A$

$$\begin{aligned} y^T \left[ \nabla^2 f(x^*) + \frac{1}{\mu} \sum_{i \in I_{V1}} \text{sign}(\phi_i(x^*)) \nabla^2 \phi_i(x^*) - \frac{1}{\mu} \sum_{i \in I_{V2}} \nabla^2 \phi_i(x^*) + \right. \\ \left. + \frac{1}{\mu} \sum_{i \in I_{A1}} \lambda_{1i} \nabla^2 \phi_i(x^*) - \frac{1}{\mu} \sum_{i \in I_{A2}} \lambda_{2i} \nabla^2 \phi_i(x^*) \right] y \geq 0. \end{aligned}$$



c) **Condições de Suficiência de Segunda Ordem** - Assumindo-se que  $f(x^*)$  e  $\phi_i(x^*)$ ,  $i = 1, \dots, m$ , são duas vezes continuamente diferenciáveis, então as condições de suficiência de segunda ordem para que  $x^*$  seja um mínimo local isolado de  $p(x, \mu)$  são:

(i) . existe  $\lambda$  tal que

$$\begin{aligned} \nabla p_1(x^*) &= \nabla f(x^*) + \frac{1}{\mu} \sum_{i \in I_{V1}} \text{sign}(\phi_i(x^*)) \nabla \phi_i(x^*) - \frac{1}{\mu} \sum_{i \in I_{V2}} \nabla \phi_i(x^*) \\ &= - \sum_{i \in I_{A1}} \lambda_{1i} \nabla \phi_i(x^*) + \sum_{i \in I_{A2}} \lambda_{2i} \nabla \phi_i(x^*) ; \end{aligned} \quad (1.27)$$

$$(ii) . - \frac{1}{\mu} < \lambda_{1i} \leq \frac{1}{\mu} , \quad i \in I_{A1} ;$$

$$(iii) . 0 \leq \lambda_{2i} \leq \frac{1}{\mu} , \quad i \in I_{A2} ;$$

(iv) . para todo vetor  $y \neq 0$ , que satisfaça

$$(y^T \nabla \phi_i(x^*) = 0 , \quad i \in I_{A1}) \text{ e}$$

$$(y^T \nabla \phi_i(x^*) = 0 , \quad i \in I_{A2} \text{ e } \lambda_{2i} > 0), \text{ e}$$

$$(y^T \nabla \phi_i(x^*) \geq 0 , \quad i \in I_{A2} \text{ e } \lambda_{2i} = 0)$$

e a seguinte desigualdade seja satisfeita

$$\begin{aligned} y^T [ \nabla^2 f(x^*) + \frac{1}{\mu} \sum_{i \in I_{V1}} \text{sign}(\phi_i(x^*)) \nabla^2 \phi_i - \frac{1}{\mu} \sum_{i \in I_{V2}} \nabla^2 \phi_i + \\ + \frac{1}{\mu} \sum_{i \in I_{A1}} \lambda_{1i} \nabla^2 \phi_i - \frac{1}{\mu} \sum_{i \in I_{A2}} \lambda_{2i} \nabla^2 \phi_i ] y > 0 . \end{aligned}$$

Estas condições necessárias e suficientes são mostra

das com detalhes por Coleman & Conn (1978).

### 1.7 Descrição Geral do Algoritmo da Função de Penalidade Exata Globalmente Convergente

O algoritmo apresentado a seguir é numericamente estável e é baseado em um método auto-consistente e unificado para que possua as propriedades de convergência global e assintoticamente uma taxa de convergência superlinear. Este algoritmo pode ser usado como um eficiente instrumento para a solução do problema geral de programação matemática mostrado na seção (1.2).

O problema geral de programação matemática definido por (1.1) pode ser transformado em um problema de minimização irrestrita de uma função de penalidade exata  $p(x, \mu)$  dado por (1.22). O termo "exata" é usado para denotar que, no caso convexo, o mínimo global da função de penalidade  $p(x, \mu)$  coincide exatamente com o mínimo do problema (1.1) para um parâmetro de penalidade  $\mu$  finito. A função de penalidade exata  $p(x, \mu)$ , em um ponto corrente, pode ser subdividida em duas partes [Conn & Pietrzykowski (1977)]: uma parte diferenciável (incluindo a função objetivo e as restrições violadas) e uma parte não diferenciável (incluindo as restrições quase ativas).

Foram determinadas [Coleman & Conn (1978)] as condições de primeira e de segunda ordem que devem ser satisfeitas

tas por  $x$  em um mínimo local de  $p(x, \mu)$ . Para que se obtivessem direções de pesquisa descendentes para  $p(x, \mu)$ , foi discutido o comportamento do  $p(x, \mu)$  quando longe ou próximo de um mínimo local de  $p(x, \mu)$  e foi feito o uso da expansão de Taylor. Quando próximo da solução, para que se obtivesse uma taxa de convergência superlinear, foi necessária a inclusão de informações de segunda ordem sobre as restrições ativas, para isto, foi feito o uso de um método que utiliza a função de penalidade exata  $p(x, \mu)$  e a função Lagrangeana aumentada  $L_{aum}(x, \lambda)$ . Quando próximo da solução, o conjunto ativo é conhecido e os multiplicadores de Lagrange  $\{\lambda_i\}$  podem ser estimados com alguma precisão.

A versão original do algoritmo [Coleman & Conn (1982b)] foi estendida para incluir as restrições de igualdade.

O algoritmo é baseado em um processo iterativo, cada iteração contendo três passos principais:

(i) - determinação de uma direção de pesquisa ótima para um dado conjunto de restrições ativas e um dado conjunto de restrições violadas;

(ii) - determinação do novo ponto de pesquisa ao longo da direção de pesquisa ótima;

(iii) - verificação das condições necessárias e suficientes para a solução, sendo que, quando estas condições são satisfeitas no novo ponto de pesquisa, conclui-se o processo

computacional.

Dentro deste processo são identificados os conjuntos de índices das restrições violadas, ativas e inativas, é computada a fatoração ortogonal  $Q^T \begin{vmatrix} R \\ 0 \end{vmatrix}$  ou a decomposição do valor singular  $U \Sigma V^T$  da matriz dos gradientes das restrições ativas  $A$ , a fatoração  $LDL^T$  (fatoração de Cholesky) ou fatoração  $LDL^T$  modificada (fatoração modificada de Cholesky) da matriz Hessiana projetada aproximada  $H_p$ , é atualizada a matriz Hessiana projetada aproximada  $H_p$  e são modificados os parâmetros que controlam o progresso do processo computacional.

### 1.8 O Problema de Fluxo de Carga Ótimo

O problema de distribuição de cargas entre as unidades geradoras de um sistema de potência é de suma importância para a indústria de geração e distribuição de potência, porque o modo como as cargas são distribuídas afetam os custos operacionais e de geração de energia.

Um pré-requisito para um despacho ótimo de potências ativas e reativas (para condições estáticas em sistema de potência) é a solução do problema de fluxo de carga ótimo. A solução do problema de fluxo de carga ótimo define as gerações de potência ativa e potência reativa de todos geradores. Essas gerações satisfazem o critério de operação ótima escolhido a priori.



Em termos matemáticos, o problema de fluxo de carga ótimo pode ser expresso como um problema geral de programação matemática definido na seção 1.2.

A escolha da função objetivo depende do critério assumido para definir a operação ótima do sistema de potência.

Em sistemas como usinas térmicas, o custo global de geração de potência ativa é frequentemente escolhido como a função objetivo a ser minimizada. Em sistemas com unidades hidráulicas, as perdas de potência ativa nas linhas de transmissão são frequentemente escolhidas como a função objetivo a ser minimizada.

As restrições impostas a serem satisfeitas são devidas às leis que regem o sistema elétrico (restrições de igualdade para os nós do sistema), às limitações técnicas do projeto e às regras estabelecidas pelos estatutos das empresas (limites para voltagem nos nós, limites para cargas de geradores, limites para transmissão de potência nas linhas de transmissão, limites para diferenças entre ângulos de fase das voltagens, limites devido a necessidade de garantir a operação segura do sistema, etc.).

A escolha das variáveis independentes para o problema é feita entre os componentes das tensões nas barras, gerações de potência ativa e reativa e os componentes da razão de transformação dos transformadores reais ou complexos.

O problema de fluxo de carga ótimo associado a sistema de 23 barras que corresponde ao problema geral de programação matemática com 75 variáveis, 46 restrições de igualdade e 136 restrições de desigualdade [Lipowski (1980a)] é descrito nos apêndices 4 e 5.



## 2 O ALGORITMO DA FUNÇÃO DA PENALIDADE EXATA GLOBALMENTE CONVERGENTE

### 2.1 A Função de Penalidade Exata

O problema geral de programação matemática dado por (1.1) pode ser transformado em um problema de minimização sem restrições de uma função de penalidade exata:

$$\underset{x}{\text{minimize}} \quad p(x, \mu) = f(x) + \frac{1}{\mu} \sum_{i=1}^m |\phi_i(x)| - \frac{1}{\mu} \sum_{i=m_1+1}^m \min(0, \phi_i(x)), \quad (2.1)$$

onde  $\mu$  é algum parâmetro positivo. Para facilidade de notação, a função de penalidade exata  $p(x, \mu)$  será representada por  $p(x)$ .

Se o parâmetro  $\mu$  é suficientemente pequeno, os mínimos locais do problema (2.1) são também as soluções do problema (1.1) [Pietrzykowski (1969)]. Pode-se, ainda, fazer as seguintes observações sobre a função de penalidade  $p(x)$ :

(i) - o nível de complexidade das restrições que defi

nem  $p(x)$  corresponde ao nível de complexidade das restrições do problema (1.1);

(ii) - não é necessário  $\mu$  se aproximar de zero para se obter uma solução precisa para o problema (2.1), como seria o caso quando a função de penalidade clássica é usada.

Contudo, a função de penalidade exata  $p(x)$  não é diferenciável em toda sua extensão. Conn e Pietrzykowski (1977) contornam esta dificuldade dividindo a função de penalidade  $p(x)$  em duas partes: uma parte diferenciável contendo a função objetivo e as restrições violadas e uma parte não diferenciável contendo as restrições ativas.

Far-se-á o uso de um pequeno escalar positivo  $\epsilon$  para identificar os conjuntos de índices das restrições aproximadamente ativas ( $\epsilon$ -ativas) para as restrições de igualdade ( $I_{A1}^\epsilon$ ) e de desigualdade ( $I_{A2}^\epsilon$ ) e os conjuntos de índices das restrições claramente violadas ( $\epsilon$ -violadas) para as restrições de igualdade ( $I_{V1}^\epsilon$ ) e de desigualdade ( $I_{V2}^\epsilon$ ) ou seja;

$$I_{A1}^\epsilon = \{i \mid |\phi_i(x)| \leq \epsilon, i = 1, \dots, m_1\} \quad , \quad (2.2)$$

$$I_{A2}^\epsilon = \{i \mid |\phi_i(x)| \leq \epsilon, i = m_1+1, \dots, m\} \quad ,$$

$$I_{V1}^\epsilon = \{i \mid |\phi_i(x)| > \epsilon, i = 1, \dots, m_1\} \quad ,$$

$$I_{V2}^\epsilon = \{i \mid \phi_i(x) < -\epsilon, i = m_1+1, \dots, m\},$$

e para identificar  $I_A^\epsilon = I_{A1}^\epsilon \cup I_{A2}^\epsilon$  e  $I_V^\epsilon = I_{V1}^\epsilon \cup I_{V2}^\epsilon$ .

A parte diferenciável da função de penalidade  $p(x)$  pode ser representada por

$$p_1(x) = f(x) + \frac{1}{\mu} \sum_{i \in I_{V1}^\epsilon} \text{sign}(\phi_i(x)) \phi_i(x) - \frac{1}{\mu} \sum_{i \in I_{V2}^\epsilon} \phi_i(x). \quad (2.3)$$

Então, a função de penalidade exata  $p(x)$  é dada por

$$p(x) = p_1(x) + \frac{1}{\mu} \sum_{i \in I_{A1}^\epsilon} \text{sign}(\phi_i(x)) \phi_i(x) - \frac{1}{\mu} \sum_{i \in I_{A2}^\epsilon} \min(0, \phi_i(x)). \quad (2.4)$$

## 2.2 A Direção de Pesquisa Horizontal

### 2.2.1 Um Método de Primeira Ordem para Minimizar $p(x)$

Para se desenvolver um método diferenciável de primeira ordem, pode-se assumir um pequeno incremento  $h$  no valor de  $x$  e aplicar a expansão de Taylor (veja seção 1.3) na função de penalidade exata  $p(x)$  nas vizinhanças de  $x$ ,

$$p(x + h) = p(x) + \nabla p(x)^T h + \frac{1}{2} h^T \nabla^2 p(\xi) h. \quad (2.5)$$

Desprezando-se o termo de segunda ordem, tem-se

$$p(x + h) = p(x) + \nabla p(x)^T h, \quad \text{ou} \quad (2.6)$$

$$\begin{aligned}
 p(x+h) = & p_1(x) + \frac{1}{\mu} \sum_{i \in I_{A1}^\varepsilon} \text{sign}(\phi_i(x)) \phi_i(x) - \frac{1}{\mu} \sum_{i \in I_{A2}^\varepsilon} \min(0, \phi_i(x)) + \\
 & + \nabla p_1(x)^T h + \frac{1}{\mu} \sum_{i \in I_{A1}^\varepsilon} \text{sign}(\phi_i(x)) \nabla \phi_i(x)^T h - \frac{1}{\mu} \sum_{i \in I_{A2}^\varepsilon} \|\nabla \phi_i(x)^T h\|,
 \end{aligned}$$

onde as aspas são usadas para indicar que o termo entre elas é zero ou  $-\|\nabla \phi_i(x)^T h\|$ . Desprezando-se os termos constantes, pois eles alteram o valor da função mas não alteram o ponto de mínimo, tem-se

$$\begin{aligned}
 \text{minimize} \quad & \nabla p_1(x)^T h + \frac{1}{\mu} \sum_{i \in I_{A1}^\varepsilon} \text{sign}(\phi_i(x)) \nabla \phi_i(x)^T h - \frac{1}{\mu} \sum_{i \in I_{A2}^\varepsilon} \|\nabla \phi_i(x)^T h\|, \\
 & \|h\| < \beta
 \end{aligned} \tag{2.7}$$

onde  $\beta$  é um pequeno escalar positivo (utiliza-se 'h' no lugar do usual 'd' para contrastar com a direção de pesquisa vertical 'v' a ser introduzida posteriormente, satisfazendo  $h^T v = 0$ , visto que o desenvolvimento conduzirá para uma direção de pesquisa denominada direção de pesquisa horizontal).

Para simplificar a computação da direção de pesquisa  $h$  que minimiza a função de penalidade exata  $p(x)$ , porque resolver o problema (2.7) não seria trivial, providencia-se que a mudança no conjunto das restrições  $\varepsilon$ -ativas seja zero,

$$\nabla \phi_i(x)^T h = 0, \quad i \in I_A^\varepsilon \tag{2.8}$$

Logo, o problema para encontrar a direção de pesquisa restrita será

$$\begin{aligned} & \text{minimize} && \nabla p_1^T h \\ & h, \|h\| < \beta \end{aligned} \quad (2.9)$$

$$\text{sujeito a} \quad \nabla \phi_i(x)^T h = 0, \quad i \in I_A^\varepsilon.$$

Denotando-se o subespaço gerado pelo gradiente das restrições  $\varepsilon$ -ativas linearmente independentes por  $S_\perp^\varepsilon$  e se  $P_\perp^\varepsilon$  for um projetor ortogonal (veja seção 1.3) que projeta algum vetor de  $R^n$  sobre o complemento ortogonal de  $S_\perp^\varepsilon$ , denotado por  $(S_\perp^\varepsilon)^\perp$ , pode-se dizer que

$$h = -\alpha P_\perp^\varepsilon \nabla p_1(x) \quad (2.10)$$

é uma direção de pesquisa descendente para o problema (2.9), desde que

$$\|P_\perp^\varepsilon \cdot \nabla p_1(x)\| \neq 0, \quad (2.11)$$

onde  $\alpha$  é um pequeno escalar positivo e escolhido de forma que  $\|h^*\| = \beta$  e  $f(x)$  e  $\phi_i(x)$ ,  $i = 1, \dots, m$ , são duas vezes continuamente diferenciáveis.

O método de Conn e Pietrzykowski (1977) é, em geral, baseado nas deduções acima. Este método possui propriedade de convergência global, mas possui, em geral, somente uma taxa de convergência linear.



### 2.2.2 Um Método de Segunda Ordem para Minimizar $p(x)$

Pode-se fazer uma tentativa para encontrar uma direção de pesquisa que minimize  $p(x)$ , utilizando termos de segunda ordem de tal forma que as mudanças nas restrições ativas sejam nulas para os termos de primeira ordem, ou seja, considerando-se a solução do problema

$$\underset{h}{\text{minimize}} \quad \nabla p_1(x)^T h + \frac{1}{2} h^T \nabla^2 p_1(x) h \quad (2.12)$$

$$\text{sujeito a} \quad \nabla \phi_1(x)^T h = 0, \quad i \in I_A^\varepsilon$$

Seja  $A$  uma matriz com dimensão  $n \times t$  formada pelos  $t$  gradientes das restrições ativas linearmente independentes que formam a base do subespaço gerado pelos gradientes de todas as restrições ativas e seja  $Z$  uma matriz com dimensão  $n \times (n-t)$  definida como parte da fatoração ortogonal da matriz dos gradientes das restrições ativas  $A$ ,

$$A = [\nabla \phi_1(x), \dots, \nabla \phi_t(x)], \quad (2.13)$$

então  $Z$  satisfaz

$$A^T Z \equiv 0, \quad (2.14)$$

$$Z^T Z \equiv I_{(n-1) \times (n-t)}.$$

Então, todos os pontos viáveis para o problema (2.12) são da



forma

$$h = Zw \quad , \quad w \in \mathbb{R}^{n-t} \quad (2.15)$$

Substituindo a equação (2.15) no problema dado por (2.12) tem-se

$$\underset{w}{\text{minimize}} \quad \nabla p_1(x)^T Zw + \frac{1}{2} w^T Z^T \nabla^2 p_1(x) Zw, \quad (2.16)$$

onde se pode ver claramente que se  $Z^T \nabla^2 p_1(x) Z$  é positiva definida, a solução para o problema (2.16) corresponde a solução do seguinte sistema linear de equações

$$Z^T \nabla^2 p_1(x) Z w = -Z^T \nabla p_1(x). \quad (2.17)$$

Logo, a solução para o problema (2.12) será

$$h^* = Zw^*, \quad (2.18)$$

onde  $w^*$  é a solução da equação (2.17).

Apesar de se obter uma convergência global, a taxa de convergência quadrática ou superlinear não é alcançada, já que as restrições ativas são aproximadas por polinômios de primeira ordem.

Considere, então, uma direção de pesquisa que minimiza  $p(x)$ , utilizando termos de segunda ordem de tal forma que as mudanças nos termos de segunda ordem das restrições ativas

sejam nulas,

$$\begin{aligned} \underset{h}{\text{minimize}} \quad & \nabla p_1(x)^T h + \frac{1}{2} h^T \nabla^2 p_1(x) h + \frac{1}{\mu} \sum_{i \in I_{A1}^\varepsilon} \text{sign}(\phi_i(x)) \left[ \nabla \phi_i(x)^T h + \right. \\ & \left. + \frac{1}{2} h^T \nabla^2 \phi_i(x) h \right] - \frac{1}{\mu} \sum_{i \in I_{A2}^\varepsilon} \left[ \nabla \phi_i(x)^T h + \frac{1}{2} h^T \nabla^2 \phi_i(x) h \right], \quad (2.18) \end{aligned}$$

que equivale claramente a

$$\underset{h}{\text{minimize}} \quad \nabla p_1(x)^T h + \frac{1}{2} h^T \nabla^2 p_1(x) h \quad (2.19)$$

$$\text{sujeito a} \quad \nabla \phi_i(x)^T h + \frac{1}{2} h^T \nabla^2 \phi_i(x) h = 0, \quad i \in I_A^\varepsilon$$

Infelizmente, uma solução do problema (2.19) não pode ser computada em contraste com o que acontece no problema (2.12), ou seja, a determinação de  $h^T \nabla^2 \phi_i(x) h$  exige um grande esforço computacional, já que  $h$  satisfaz o problema (2.19). Pode-se, entretanto, obter uma aproximação computacional para o problema (2.19) por meio do problema Langrangeano equivalente,

$$\begin{aligned} \min_h \max_\lambda \quad & L(h, \lambda) = \nabla p_1(x)^T h + \frac{1}{2} h^T \nabla^2 p_1(x) h + \sum_{i \in I_{A1}^\varepsilon} \lambda_{1i} (\nabla \phi_i(x)^T h + \\ & + \frac{1}{2} h^T \nabla^2 \phi_i(x) h) - \sum_{i \in I_{A2}^\varepsilon} \lambda_{2i} (\nabla \phi_i(x)^T h + \frac{1}{2} h^T \nabla^2 \phi_i(x) h). \quad (2.20) \end{aligned}$$

A solução do problema (2.20) não pode ser obtida computacionalmente. Entretanto, pode-se substituir o problema (2.20) por um problema solucionável, cuja solução fornece as aproximações das variáveis  $\{\lambda_{1i}\}$  e  $\{\lambda_{2i}\}$  (conjuntos de multiplicadores

Lagrange associados às restrições aproximadamente ativas de igualdade e de desigualdade, respectivamente).

Para um ponto estacionário  $\bar{x}$  de  $p(x)$ , existem vetores  $\bar{\lambda}_{1i}$  e  $\bar{\lambda}_{2i}$  tais que

$$p_1(\bar{x}) = - \sum_{i \in I_{\bar{A}1}} \bar{\lambda}_{1i} \nabla \phi_i(\bar{x}) + \sum_{i \in I_{\bar{A}2}} \bar{\lambda}_{2i} \nabla \phi_i(\bar{x}).$$

Denotando-se a solução do sistema hiperdeterminado de equações,

$$\nabla p(x) = - \sum_{i \in I_{A1}^\epsilon} \lambda_{1i} \nabla \phi_i(x) + \sum_{i \in I_{A2}^\epsilon} \lambda_{2i} \nabla \phi_i(x), \quad (2.21)$$

por  $\tilde{\lambda}_1$  e  $\tilde{\lambda}_2$  e se  $I_A^\epsilon = I_{\bar{A}}$ , é razoável adotar  $\tilde{\lambda}_{1i}$  como a aproximação de  $\bar{\lambda}_{1i}$  e  $\tilde{\lambda}_{2i}$  como aproximação de  $\bar{\lambda}_{2i}$  (se  $x^k \rightarrow x$  e  $\epsilon$  é suficientemente pequeno, então existe um  $K$  tal que  $I_A^\epsilon(x^k) = I_{\bar{A}}(\bar{x})$  para todo  $k \geq K$ , e portanto  $\tilde{\lambda}^k \rightarrow \bar{\lambda}$ , por continuidade). Portanto, o problema (2.19) pode ser aproximado por

$$\begin{aligned} \min_h \max_\lambda \quad \tilde{L}(h, \lambda) = & \nabla p_1(x)^T h + \frac{1}{2} h^T \nabla^2 p_1(x) h + \frac{1}{2} \sum_{i \in I_{A1}^\epsilon} \tilde{\lambda}_{1i} h^T \nabla^2 \phi_i(x) h - \\ & - \frac{1}{2} \sum_{i \in I_{A2}^\epsilon} \tilde{\lambda}_{2i} h^T \nabla^2 \phi_i(x) h + \sum_{i \in I_{A1}^\epsilon} \lambda_{1i} \nabla \phi_i(x)^T h - \\ & - \sum_{i \in I_{A2}^\epsilon} \lambda_{2i} \nabla \phi_i(x)^T h. \end{aligned} \quad (2.22)$$

Juntando-se os termos Hessianos de segunda ordem, p<sub>0</sub> de-se escrever a função  $L(h, \lambda)$  como

$$\begin{aligned} \tilde{L}(h, \lambda) = & \nabla p_1(x)^T h + \frac{1}{2} h^T \left[ \nabla^2 p_1(x) + \sum_{i \in I_{A1}^\varepsilon} \tilde{\lambda}_{1i} \nabla^2 \phi_i(x) - \right. \\ & \left. - \sum_{i \in I_{A2}^\varepsilon} \tilde{\lambda}_{2i} \nabla^2 \phi_i(x) \right] h + \sum_{i \in I_{A1}^\varepsilon} \lambda_{1i} \nabla \phi_i(x)^T h - \\ & - \sum_{i \in I_{A2}^\varepsilon} \lambda_{2i} \nabla \phi_i(x)^T h. \end{aligned} \quad (2.23)$$

Supondo-se as condições

$$\frac{\partial \tilde{L}(h, \lambda)}{\partial \lambda_{1i}} = 0, \quad i \in I_{A1}^\varepsilon \quad \text{ou}$$

$$\frac{\partial L(h, \lambda)}{\partial \lambda_{2i}} = 0, \quad i \in I_{A2}^\varepsilon,$$

tem-se

$$\nabla \phi_i(x)^T h = 0, \quad i \in I_A^\varepsilon$$

e, portanto, pode-se fazer

$$h = Zw, \quad w \in R^{n-t}. \quad (2.24)$$

O problema (2.22) torna-se

$$\underset{w}{\text{minimize}} \quad \nabla p_1(x)^T Z w + \frac{1}{2} w^T Z^T H Z w, \quad (2.25)$$

onde

$$H = \nabla^2 p_1(x) + \sum_{i \in I_{A1}^\varepsilon} \tilde{\lambda}_{1i} \nabla^2 \phi_i(x) - \sum_{i \in I_{A2}^\varepsilon} \tilde{\lambda}_{2i} \nabla^2 \phi_i(x),$$

$$H_p = Z^T H Z.$$

Assumindo que  $H$  é positiva definida, a solução do problema (2.25) é obtida pela solução de

$$Z^T H Z w = -Z^T \nabla p_1(x) \quad (2.26)$$

e, por conseguinte, a solução do problema (2.22) é dada por

$$h^* = Z w^*, \quad (2.27 a)$$

onde  $w^*$  é a solução da relação (2.26).

A direção de pesquisa horizontal  $h$ , expressa pela equação (2.27 a), pode ser apresentada na forma

$$h^* = Z(Z^T H Z)^{-1} Z^T \nabla p_1(x). \quad (2.27 b)$$

Desde que  $Z^T \nabla p_1(x) \neq 0$ , a direção de pesquisa  $h^*$  é uma



direção descendente para  $p(x)$ . Em adição,  $h^*$  é o análogo de segunda ordem da direção de primeira ordem dada pela equação (3.4), usado por Conn e Pietrzykowski. Na prática, a matriz Hessiana projetada ( $Z^T H Z$ ) não precisa ser computada, mas, de preferência, aproximada por uma matriz positiva definida. Desde que as condições de suficiência para  $p(x)$  assegurem que a matriz Hessiana projetada verdadeira é positiva definida para um mínimo local de  $p(x)$ , é razoável restringir a matriz Hessiana projetada para esta classe.

Quando  $\|Z^T \nabla p_1(x)\|$  é grande ( $x$  está longe de um ponto estacionário), os  $\{\lambda_{1i}\}$ ,  $i \in I_{A1}^E$ , e  $\{\lambda_{2i}\}$ ,  $i \in I_{A2}^E$ , estimados possuem pouco significado. Quando longe de um ponto estacionário de  $p(x)$ , a função objetivo e as restrições violadas poderiam dominar a mudança da função de penalidade exata  $p(x)$ . Portanto, é razoável ignorar as mudanças nas restrições ativas correntes e  $Z^T H Z$  é uma aproximação positiva definida de  $Z^T \nabla^2 p_1(x) Z$  e o vetor  $\lambda$  não é computado (pode ser aproximado pelo vetor zero).

Pode-se interpretar  $h$  como uma aproximação do passo de Newton restrito (com respeito a  $x$ ) para a função Lagrangeana (no subespaço gerado pelas colunas de  $Z$  e contendo o ponto  $x$ ). Este passo de Newton é baseado na informação do gradiente da função Lagrangeana aproximada no ponto  $x$  e é uma tentativa do método de Newton para satisfazer  $Z^T \nabla L(x+h, \lambda) = 0$ .

Computacionalmente, a equação (2.27 b), que gera a  $di$



reção de pesquisa horizontal, pode ser resolvida com maior eficiência se a fatoração  $LDL^T$  da matriz Hessiana projetada  $H_p$ , equação (2.25), for computada primeiramente,

$$LDL^T \approx H_p. \quad (2.28)$$

Então, a nova forma da equação (2.26) será:

$$LDL^T w = -Z^T \nabla p_1(x). \quad (2.29)$$

A equação (2.29) pode ser resolvida em duas etapas. Na primeira etapa, a equação

$$Lv = -Z^T \nabla p_1(x), \quad (2.30)$$

pode ser resolvida para obter o vetor  $v$  por substituição progressiva. Na segunda etapa

$$L^T w = D^{-1} v, \quad (2.31)$$

pode ser resolvida para obter o vetor  $w$  por substituição regressiva.

### 2.3 Os Multiplicadores de Lagrange

Supondo-se que a sequência de pontos  $\{x^k\}$ ,  $k = 1, 2, \dots$ , obtidos de um processo iterativo, tende para um ponto esta

cionário  $\bar{x}$ , de  $p(x)$ , pelas condições necessárias de primeira ordem, existe um vetor  $\bar{\lambda}$  que satisfaz

$$\begin{aligned} \nabla p_1(\bar{x}) &= \nabla f(\bar{x}) + \frac{1}{\mu} \sum_{i \in I_{\bar{V}1}} \text{sign}(\phi_i(\bar{x})) \nabla \phi_i(\bar{x}) - \frac{1}{\mu} \sum_{i \in I_{\bar{V}2}} \nabla \phi_i(\bar{x}) \\ &= - \sum_{i \in I_{\bar{A}1}} \bar{\lambda}_{1i} \nabla \phi_i(\bar{x}) + \sum_{i \in I_{\bar{A}2}} \bar{\lambda}_{2i} \nabla \phi_i(\bar{x}), \end{aligned} \quad (2.32)$$

onde  $I_{\bar{A}1}$ ,  $I_{\bar{A}2}$  e  $I_{\bar{V}1}$ ,  $I_{\bar{V}2}$ , são, respectivamente, os conjuntos de índices das restrições ativas de igualdade e de desigualdade e os conjuntos de índices das restrições violadas de igualdade e de desigualdade para  $\bar{x}$ , isto é,

$$I_{\bar{A}1} = \{i \mid \phi_i(\bar{x}) = 0, i = 1, \dots, m_1\}, \quad (2.33)$$

$$I_{\bar{A}2} = \{i \mid \phi_i(\bar{x}) = 0, i = m_1+1, \dots, m\},$$

$$I_{\bar{V}1} = \{i \mid |\phi_i(\bar{x})| > 0, i = 1, \dots, m_1\},$$

$$I_{\bar{V}2} = \{i \mid \phi_i(\bar{x}) < 0, i = m_1+1, \dots, m\}.$$

Claramente, para toda iteração  $k$ ,  $k$  suficientemente grande,  $I_{\bar{A}}^\varepsilon(x^k) = I_{\bar{A}}$  e  $I_{\bar{V}}^\varepsilon(x^k) = I_{\bar{V}}$ , onde  $I_{\bar{A}} = I_{\bar{A}1} \cup I_{\bar{A}2}$  e  $I_{\bar{V}} = I_{\bar{V}1} \cup I_{\bar{V}2}$ , então  $\lambda^k \rightarrow \bar{\lambda}$ , onde  $\bar{\lambda}$  é uma solução do seguinte sistema hiperdeterminado de equações

$$\begin{aligned} \nabla p_1(x^k) &= A_k \begin{bmatrix} -\lambda_1^k \\ \lambda_2^k \end{bmatrix} = \nabla f(x^k) + \frac{1}{\mu} \sum_{i \in I_{\bar{V}1}^\varepsilon} \text{sign}(\phi_i(x^k)) \nabla \phi_i(x^k) \\ &\quad - \frac{1}{\mu} \sum_{i \in I_{\bar{V}2}^\varepsilon} \nabla \phi_i(x^k). \end{aligned} \quad (2.34)$$

onde

$t_1^k$  - número de restrições de igualdade  $\epsilon$ -ativas que formam a base do subespaço  $S_{\perp}^{\epsilon}$  na  $k$ -ésima iteração;

$t_2^k$  - número de restrições de desigualdade  $\epsilon$ -ativas que formam a base do subespaço  $S_{\perp}^{\epsilon}$  na  $k$ -ésima iteração;

$\lambda_1^k = \left[ \lambda_{11}^k, \dots, \lambda_{1t_1^k}^k \right]^T$  - vetor dos multiplicadores de Lagrange associados às restrições  $\epsilon$ -ativas de igualdade que formam a base do subespaço  $S_{\perp}^{\epsilon}$  na  $k$ -ésima iteração;

$\lambda_2^k = \left[ \lambda_{21}^k, \dots, \lambda_{2t_2^k}^k \right]^T$  - vetor dos multiplicadores de Lagrange associados às restrições  $\epsilon$ -ativas de desigualdade que formam a base do subespaço  $S_{\perp}^{\epsilon}$  na  $k$ -ésima iteração.

Se as restrições ativas foram corretamente identificadas, então  $\|Z^T \nabla p_1(x)\| \rightarrow 0$  quando  $x^k \rightarrow \bar{x}$ , onde

$$\nabla p_1(x^k) = \nabla f(x^k) + \frac{1}{\mu} \sum_{i \in I_{V1}^{\epsilon}} \text{sign}(\phi_i(x^k)) \nabla \phi_i(x^k) - \frac{1}{\mu} \sum_{i \in I_{V2}^{\epsilon}} \nabla \phi_i(x^k) \quad (2.35)$$

chama-se pseudo-gradiente de  $p(x)$  no ponto  $x^k$ .

Quando  $\|Z_k^T \nabla p_1(x^k)\|$  torna-se suficientemente pequeno, a solução da equação (2.34), usualmente, fornecerá estimativas para os multiplicadores de Lagrange  $\{\lambda_{1i}\}$ ,  $i \in I_{A1}^\varepsilon$ , e  $\{\lambda_{2i}\}$ ,  $i \in I_{A2}^\varepsilon$ , usados em (2.32). Quando  $\|Z_k^T \nabla p_1(x^k)\|$  é grande, a solução da equação (2.34), provavelmente, possui pouca similaridade com  $\bar{\lambda}_1$  e  $\bar{\lambda}_2$  (particularmente se  $I_A^\varepsilon(x^k) \neq I_{\bar{A}}$ ).

Computacionalmente, a equação (2.34), que determina os multiplicadores de Lagrange, pode ser resolvida com maior eficiência utilizando-se a fatoração ortogonal da matriz dos gradientes das restrições  $\varepsilon$ -ativas A,

$$QA = \begin{bmatrix} R \\ 0 \end{bmatrix}, \quad (2.36)$$

onde

Q - matriz ortogonal de ordem n

R - matriz triangular superior de ordem t

A matriz ortogonal Q pode ser particionada da seguinte forma

$$Q = \begin{bmatrix} Q_1 \\ \hline Q_2 \end{bmatrix}, \quad (2.37)$$

onde

$Q_1$  - submatriz superior da matriz ortogonal  $Q$  com dimensão  $t \times n$ ,

$Q_2$  - submatriz inferior da matriz ortogonal  $Q$  com dimensão  $(n - t) \times n$ .

Pré-multiplicando ambos os lados da equação (2.34) pela matriz ortogonal  $Q$  e considerando as equações (2.36) e (2.37), obtem-se:

$$\begin{bmatrix} R \\ 0 \end{bmatrix} \begin{bmatrix} -\lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \nabla p_1(x). \quad (2.38)$$

A parte da equação (2.38) que define a solução de 'mínimos quadrados', do sistema hiperdeterminado de equações, pode ser escrita na forma

$$R \begin{bmatrix} -\lambda_1 \\ \lambda_2 \end{bmatrix} = Q_1 \nabla p_1(x), \quad (2.39)$$

e pode ser resolvida para obter o vetor  $\lambda$  por substituição regressiva.

## 2.4 A Direção de Pesquisa Vertical

A direção de pesquisa horizontal dada pela equação (2.27b) faz decrescer  $p(x)$  enquanto tenta manter as restrições  $\epsilon$ -ativas constantes. Supondo-se que  $x^*$  satisfaça as condições de suficiência de segunda ordem para o problema (2.19), que  $\epsilon$  e  $\|x^k - x^*\|$  são suficientemente pequenos tal que



$I_A^\varepsilon(x^k) = I_A^*$  e  $I_V^\varepsilon(x^k) = I_V^*$ , onde as restrições no ponto  $x^k$  não são precisamente nulas, e que o passo horizontal mudará os valores das restrições  $\varepsilon$ -ativas de alguns 'graus', parece razoável tentar satisfazer as restrições  $\varepsilon$ -ativas mais precisamente, quando  $x^k$  é próximo de  $x^*$ . Pode-se fazer isto introduzindo uma direção de pesquisa vertical  $v$ .

Supondo-se que  $I_A^\varepsilon(x^k) = \{i_1, i_2, \dots, i_{t^k}\}$ , onde  $i_1, i_2, \dots, i_{t^k}$  são os índices das restrições  $\varepsilon$ -ativas na  $k$ -ésima iteração, cujos gradientes são linearmente independentes,  $t^k \leq n$ , e que

$$\phi(x^k) = \left[ \phi_{i_1}(x^k), \dots, \phi_{i_{t^k}}(x^k) \right]^T, \quad (2.40)$$

$$A_k = \left[ \nabla \phi_{i_1}(x^k), \dots, \nabla \phi_{i_{t^k}}(x^k) \right],$$

obtêm-se, usando a expansão de Taylor,

$$\phi(x^k + v) = \phi(x^k) + A_k^T v + O(\|v\|^2). \quad (2.41)$$

Ignorando-se o termo de segunda ordem e igualando esta expressão a zero, tem-se

$$A_k^T v = -\phi(x^k). \quad (2.42)$$

Este sistema de equações é hipodeterminado e a solução exata da equação (2.42) é dada por



$$v = -A_k (A_k^T A_k)^{-1} \phi(x^k), \quad (2.43)$$

se as colunas da matriz  $A_k$  são linearmente independentes.

Conn e Pietrzykowski mostraram que nas vizinhanças da solução  $p(x^k + v) < p(x^k)$ , portanto, a direção de pesquisa vertical  $v$  é usada juntamente com a direção de pesquisa horizontal  $h$  nas vizinhanças de um mínimo.

Se  $v^k$  é somado após  $h^k$ , então  $v^k$  é uma tentativa para resolver  $\phi(x^k + h^k + v^k) = 0$ . Parece razoável que  $v^k$  seria melhor avaliado se as atualizações das matrizes e de suas fatorações fossem feitas após determinar  $x^k + h^k$ . Tal computação Coleman & Conn (1982a) seria injustificavelmente cara. Entretanto considerações de convergência global Coleman & Conn (1982b) exigem que as restrições  $\epsilon$ -ativas  $\phi_i(x^k)$ ,  $i \in I_A^\epsilon$ , sejam avaliadas após determinar  $x^k + h^k$ , embora este procedimento não destrua a superlinearidade com a execução do passo 2. Logo, define-se

$$\bar{v}^k = -A_k (A_k^T A_k)^{-1} \phi(x^k + h^k) \quad (2.44)$$

e

$$x^{k+1} = x^k + h^k + \bar{v}^k.$$

Somente os valores das funções  $\phi(x^k + h^k)$  são obtidos

para  $x^k + h^k$  e a matriz  $A_k$  não é atualizada para os valores de  $x^k + h^k$ , mas  $A_k$  contém a informação precisa sobre os gradientes das restrições  $\epsilon$ -ativas linearmente independentes no ponto  $x^k$ .

Computacionalmente, a equação (2.42), que gera a direção de pesquisa vertical  $v$ , pode ser resolvida com maior eficiência se a fatoração ortogonal da matriz dos gradientes das restrições  $\epsilon$ -ativas for computada primeiramente.

Das equações (2.36) e (2.37), obtem-se

$$A^T = R^T Q_1, \quad (2.45)$$

Logo, a equação (2.42) torna-se

$$R^T Q_1 v = -\phi(x + h). \quad (2.46)$$

A relação (2.46) pode ser resolvida em duas etapas. Na primeira etapa, a equação

$$R^T w = -\phi(x + h), \quad (2.47)$$

pode ser resolvida por substituição regressiva para obter o vetor  $w$ . Na segunda etapa, calcula-se

$$v = Q_1^T w. \quad (2.48)$$

## 2.5 Queda de Uma Restrição

Supondo-se que  $\{x^k\} \rightarrow x^*$ ,  $x^*$  é um ponto de mínimo para  $p(x)$ , e que  $\|Z_k^T \nabla p_1(x^k)\|$  seja pequeno, então os multiplicadores de Lagrange  $\lambda$  são computados e satisfazem a equação (2.34). Supondo-se que  $\lambda_{1j} \notin \left[-\frac{1}{\mu}, \frac{1}{\mu}\right]$  ou  $\lambda_{2j} \notin \left[0, \frac{1}{\mu}\right]$ ,  $j \in I_A^\epsilon$ , e levando-se em consideração as condições necessárias de primeira ordem para um mínimo local de  $p(x)$ , deve-se tentar encontrar uma direção de pesquisa de primeira ordem que faça decrescer  $p(x)$  em um subespaço gerado pelos gradientes das restrições  $\epsilon$ -ativas que não contenham a restrição  $j$ , isto é, o subespaço gerado por  $\{\nabla \phi_i(x^k) \mid i \in I_A^\epsilon - \{j\}\}$ .

Supondo que seja um projetor ortogonal no subespaço gerado por  $\{\nabla \phi_i(x) \mid i \in I_A^\epsilon - \{j\}\}$  e considerando-se que  $\lambda_j$  exista, o multiplicador de Lagrange associado à restrição que não pertence ao intervalo definido pelas condições necessárias de primeira ordem, a direção de pesquisa horizontal

$$\hat{h} = \sigma_j P^j \nabla \phi_j(x), \quad (2.49)$$

onde

$$\sigma_j = -\text{sinal}(\lambda_j)$$

é uma direção de pesquisa descendente para  $p(x)$ . Escolhendo-se o projetor ortogonal  $P^j$  como sendo  $Z_{k \setminus j}^T Z_{k \setminus j}$ , onde  $Z_{k \setminus j}$  é determinado pela fatoração ortogonal da matriz  $A$  cujas co

funas são os gradientes  $\{\nabla\phi_i(x) \mid i \in I_A^E - \{j\}\}$ , tem-se

$$\tilde{h} = \sigma_j Z_{k \setminus j}^T Z_{k \setminus j} \nabla\phi_j(x^k). \quad (2.50)$$

Para verificar que esta direção de pesquisa decresce  $p(x)$  [Coleman & Conn (1982b)], nota-se que

$$\nabla\phi_i(x^k)^T \tilde{h} = \sigma_j \nabla\phi_i(x^k)^T Z_{k \setminus j}^T Z_{k \setminus j} \nabla\phi_j(x^k) = 0, \quad i \in I_A^E - \{j\}. \quad (2.51)$$

Por outro lado, tem-se

$$p(x^k + \alpha \tilde{h}) \leq p(x^k) + \alpha g(x^k)^T \tilde{h} + o(\alpha^2), \quad (2.52)$$

onde

$$g(x^k) = \nabla p_1(x^k) + \frac{1}{\mu} \min(0, \sigma_j) \cdot \nabla\phi_j(x^k).$$

Considerando a equação (2.50), tem-se

$$\phi_i(x^k + \alpha \tilde{h}) = \phi_i(x^k) + o(\alpha^2), \quad i \in I_A^E - \{j\}. \quad (2.53)$$

Logo, podem-se fazer as seguintes observações:

CASO 1: supondo-se que  $\lambda_j < 0$ , então  $\sigma_j = 1$  e

$$g(x^k)^T \tilde{h} = \nabla p_1(x^k)^T \tilde{h}, \quad \text{, usando a equação (2.53)}$$

$$= (A_k \lambda^k)^T \tilde{h}, \quad \text{, usando a equação (2.34)}$$

$$\begin{aligned}
 &= \lambda_j^k \nabla \phi_j(x^k)^T \tilde{h} \quad , \text{ usando a equação (2.51)} \\
 &= \lambda_j^k \| Z_{k \setminus j}^T \nabla \phi_j(x^k) \| < 0 ;
 \end{aligned}$$

CASO 2: supondo-se que  $\lambda_j > \frac{1}{\mu}$ , então  $\sigma_j = -1$  e

$$\begin{aligned}
 g(x^k)^T \tilde{h} &= \nabla p_1(x^k) - \frac{1}{\mu} \nabla \phi_j(x^k)^T \tilde{h}, \text{ usando a equação (2.53)} \\
 &= (\lambda_j^k - \frac{1}{\mu}) \nabla \phi_j(x^k)^T \tilde{h} \quad , \text{ usando a equação (2.51)} \\
 &= (\frac{1}{\mu} - \lambda_j^k) \| Z_{k \setminus j}^T \nabla \phi_j(x^k) \|^2 < 0.
 \end{aligned}$$

## 2.6 Uma Estratégia da Escolha da Direção de Pesquisa

Foram definidas três direções de pesquisa

$$h^k = Z_k (Z_k^T H_k Z_k)^{-1} Z_k \nabla p_1(x^k) , \quad (2.54)$$

$$v^k = -A_k (A_k^T A_k)^{-1} \phi(x^k + h^k) , \quad (2.55)$$

$$\tilde{h}^k = \sigma_j Z_{k \setminus j} Z_{k \setminus j}^T \nabla \phi_j(x^k) . \quad (2.56)$$

Pode-se sugerir uma estratégia simples e razoável para o bom desempenho do algoritmo, embora estratégias mais complicadas possam ser utilizadas sem a violação das propriedades de convergência.

Quando  $\| Z_k^T \nabla p_1(x^k) \|$  é grande, isto é, quando

$\|Z_k^T \nabla p_1(x^k)\| > \underline{\mu}$  pode-se escolher a direção de pesquisa  $h^*$ , onde  $Z_k^T H_k Z_k$  é uma aproximação de

$$Z_k^T \left[ \nabla^2 f(x^k) + \frac{1}{\mu} \sum_{i \in I_{V1}} \text{sinal}(\phi_i(x^k)) \nabla^2 \phi_i(x^k) - \frac{1}{\mu} \sum_{i \in I_{V2}} \nabla^2 \phi_i(x^k) \right] Z_k.$$

Quando  $\|Z_k \nabla p_1(x^k)\| \leq \underline{\mu}$  os multiplicadores de Lagrange tornam-se importantes. A direção de pesquisa  $\tilde{h}^k$  é tentada quando os multiplicadores de Lagrange não estão dentro dos domínios definidos pelas condições necessárias de primeira ordem, ou seja,  $\lambda_{1i} \notin \left[-\frac{1}{\mu}, \frac{1}{\mu}\right]$  ou  $\lambda_{2i} \notin \left[0, \frac{1}{\mu}\right]$ . Neste caso, a direção de pesquisa  $\tilde{h}$  é aceita somente se um decréscimo suficiente da função de penalidade exata  $p(x)$  pode ser garantido, isto é, usa-se  $\tilde{h}$  se

$$\nabla p_1(x^k) + \min(0, \sigma_j^k) \cdot \tilde{h}^k < -\delta \quad (2.57)$$

para algum  $\delta > 0$  e

$$\sigma_j^k = -\text{sinal}(\lambda_{1j}^k) \quad \text{se} \quad \lambda_{1j}^k \in \left[-\frac{1}{\mu}, \frac{1}{\mu}\right], \quad \text{ou} \quad (2.58)$$

$$\sigma_j^k = -\text{sinal}(\lambda_{2j}^k) \quad \text{se} \quad \lambda_{2j}^k \in \left[0, \frac{1}{\mu}\right].$$

Se ocorrer outra situação, pode-se determinar um comprimento do passo  $\alpha_k$ , obtido pela pesquisa linear ao longo da direção de pesquisa, e adotar-se-á

$$x^{k+1} = x^k + \alpha_k h^k \quad \text{ou} \quad (2.59)$$

$$x^{k+1} = x^k + \alpha_k \tilde{h}^k.$$



O segundo uso dos multiplicadores de Lagrange serve para incluir informações de curvatura das restrições na matriz Hessiana aproximada, isto é, quando  $\|Z_k^T \nabla p_1(x^k)\| \leq \mathcal{L}$ , então  $Z_k^T H_k Z_k$  é uma aproximação de

$$Z_k^T \left[ \nabla^2 f(x^k) + \frac{1}{\mu} \sum_{i \in I_{V1}^\epsilon} \text{sign}(\phi_i(x^k)) \nabla^2 \phi_i(x^k) - \frac{1}{\mu} \sum_{i \in I_{V2}^\epsilon} \nabla^2 \phi_i(x^k) + \sum_{i \in I_{A1}^\epsilon} \lambda_{1i} \nabla^2 \phi_i(x^k) - \sum_{i \in I_{A2}^\epsilon} \lambda_{2i} \nabla^2 \phi_i(x^k) \right] Z_k.$$

Se os multiplicadores estão em seus domínios  $\lambda_{1i} \in \left[-\frac{1}{\mu}, \frac{1}{\mu}\right]$  e  $\lambda_{2i} \in \left[0, \frac{1}{\mu}\right]$ , a direção de pesquisa torna-se

$$d^k = h^k + v^k. \quad (2.60)$$

Aqui, tenta-se um comprimento de passo  $\alpha_k$  igual a 1; se um suficiente decréscimo não é observado, então a pesquisa linear é feita ao longo da direção de pesquisa  $h^k$ .

Se a direção de pesquisa determinada pela equação (2.56) ou o passo de Newton determinada pela equação (2.60) são tentados sem sucesso, então reduz-se  $\epsilon$  e  $\mathcal{L}$ . A escolha de  $\epsilon$  e  $\mathcal{L}$  é arbitrária desde que  $\epsilon > 0$  e  $\mathcal{L} > 0$ .

A escolha da direção de pesquisa acima é arbitrária. É possível que através de experimentos numéricos mais especializados poderá surgir uma estratégia seletiva alternativa.

## 2.7 Pesquisa Linear ao Longo da Direção de Pesquisa

Para um dado ponto  $x^k$  e uma direção de pesquisa  $d^k$  (por exemplo,  $h^k$  dada pela equação (2.54)), a minimização da função de penalidade exata  $p(x^k + \alpha d^k)$ , ao longo da direção de pesquisa  $d^k$ , pode ser feita aplicando algum algoritmo apropriado para a minimização da função univariável não diferenciável  $\psi(\alpha)$ , onde

$$\psi(\alpha) = p(x^k + \alpha d^k). \quad (2.61)$$

Quando se está em uma vizinhança de  $x^*$ , num ponto que satisfaça as condições de suficiência de segunda ordem, o comprimento do passo  $\alpha$  adotado será uma unidade ( $\alpha = 1$ ). Espera-se que o procedimento da pesquisa linear seja usado quando um ponto está longe do ponto de solução  $x^*$ .

Se as funções  $f(x^k)$  e  $\phi_j(x^k)$ ,  $i = 1, \dots, m$ , são todas lineares, duas situações podem-se perceber: (i) a função de penalidade exata  $p(x^k + \alpha d^k)$  não tem um limite finito na direção de pesquisa  $d^k$ , ou (ii) um mínimo ao longo da direção de pesquisa  $d^k$  ocorre num ponto de descontinuidade  $\tilde{x}$  tal que  $\phi_j(\tilde{x}) = 0$  para pelo menos um  $j \in I_A(x^k)$ . Desde que está sendo descrita uma situação linear, pode-se assumir que  $\epsilon = 0$ . Então, para este caso, um algoritmo simples para a pesquisa linear é:

(i) - determine todos os pontos de descontinuidade  $\tilde{x}_i$  tais que  $\phi_j(\tilde{x}) = 0$  para todo  $j \in I_A(x^k)$  e

(ii) - encontre o valor m nimo da fun o de penalidade exata  $p(x)$  para estes pontos.

Usando a expans o de Taylor de  $\phi_j(x^k + \alpha d^k)$  e desprezando-se os termos de segunda ordem, pode-se obter o comprimento do passo  $\alpha$ , ao longo da dire o de pesquisa  $d^k$ , para os pontos de descontinuidade,  $\phi_j(x^k + \alpha d^k) = 0$ , por

$$\alpha_j = \frac{-\phi_j(x^k)}{\nabla\phi_j(x^k)^T d^k}, \quad i \in I_M - I_A(x^k), \quad (2.62)$$

onde

$$I_M = \{i \mid i = 1, \dots, m\}.$$

Desde que  $d^k$  seja uma dire o de pesquisa descendente para  $p(x)$  em  $x^k$ , segue-se que  $\nabla p_1(x^k)^T \cdot d^k < 0$ . Supondo que  $\tilde{x}_1 = x^k + \alpha_1 d^k$  seja o primeiro ponto de descontinuidade ao longo da dire o de pesquisa  $d^k$  e correspondendo a fun o  $\phi_j$ , tem-se que  $\phi_j(\tilde{x}_1) = 0$  e  $\phi_j(x^k) \neq 0$ . O vetor  $\nabla p_1(x)$  n o mudar  at  que  $\tilde{x}_1$  seja alcan ado, por conseguinte,  $p(x)$  continuar  a decrescer,

Ap s um incremento  $\Delta d$  em  $\tilde{x}_1$ , o vetor  $\nabla p_1(\alpha)$  torna-se

$$\nabla p_1(\tilde{x}_1 + \Delta d) = \nabla p_1(x) + \frac{1}{\mu} \text{sign}(\nabla\phi_j(\tilde{x}_1)^T d^k) \nabla\phi_j(x_1), \quad (2.63)$$

onde  $\Delta d$  é alguma pequena quantidade positiva. Assume-se, para a finalidade de descrição que os pontos de descontinuidade sejam distintos. Se

$$\nabla p_1(\tilde{x} + \Delta d)^T d^k = \nabla p_1(\tilde{x}_1)^T d^k + \frac{1}{\mu} \sin \alpha (\nabla \phi_j(\tilde{x}_1)^T d^k) \nabla \phi_j(\tilde{x}_1) d^k < 0 \quad (2.64)$$

significa que  $f(x)$  continua a decrescer ao longo da direção de pesquisa  $d^k$ . Logo, determina-se o próximo ponto de descontinuidade e o argumento é repetido.

Se progredir através de uma lista distribuída de pontos de descontinuidade e  $p(x)$  continuar a decrescer, depois do último ponto de descontinuidade, o procedimento retorna indicando que a função de penalidade exata não tem limite inferior finito ao longo da direção de pesquisa  $d^k$  para o valor dado ao parâmetro  $\mu$ .

No caso não linear, onde o mínimo de  $p(x)$  ao longo de  $d^k$  não precisa ser um ponto de descontinuidade, se depois do último ponto de descontinuidade  $p(x)$  continuar a decrescer, não significa que  $p(x)$  seja ilimitado. Neste caso, toma-se o menor comprimento do passo  $\alpha$  obtido pela equação (2.62) e verifica se um suficiente decréscimo da função de penalidade exata  $p(x)$  é garantido. Se um suficiente decréscimo de  $p(x)$  não é assegurado ou se não existirem pontos de descontinuidade, tenta-se, então, uma pesquisa linear com interpolação quadrática.



Para esta pesquisa linear, o seguinte algoritmo pode ser utilizado

Passo 1

$$a_0 \leftarrow \nabla p_1(x^k)^T d^k < 0 \quad (\text{por suposição})$$

$$\alpha_i \leftarrow -\phi_i(x^k) / \nabla \phi_i(x^k)^T d^k$$

$$I_0 \leftarrow \{i \in I_M - I_A \mid \alpha_i > 0\}$$

$$k \leftarrow 0$$

$$\ell \leftarrow 0$$

$$\alpha_0 \leftarrow 0$$

Passo 2

Se  $I_k = \emptyset$  e  $k = 0$ ,  $\alpha_{\ell_k} \leftarrow 0$ , vá para o passo 5

Se  $I_k = \emptyset$  e  $k > 0$ ,  $\alpha_{\ell_k} \leftarrow \alpha_{\ell_{k-1}}$ , vá para o passo 5

Passo 3

determine  $\ell_k$  tal que  $\alpha_{\ell_k} = \alpha_i$ , para todo  $i \in I_k$

$a_{k+1} \leftarrow a_k + \alpha_{\ell_k} \nabla \phi_{\ell_k}(x^k)^T d^k$ , onde  $\alpha_{\ell_k} \leftarrow \text{sign}(\nabla \phi_{\ell_k}(x^k)^T d^k)$

Passo 4

Se  $a_{k+1} \geq 0$ , vá para o passo 5

$$I_{k+1} \leftarrow I_k - \{\ell_k\}$$

$k \leftarrow k + 1$  , vá para o passo 2

Passo 5

$$x^{k+1} = x^k + \alpha_{\ell_k} d^k$$

Se  $p(x^{k+1}) < p(x^k) - \delta$ , onde  $\delta$  é um pequeno escalar positivo, retorne;

faça a pesquisa linear com interpolação quadrática minimizando entre  $x^k$  e  $x^k + \tau d^k$ , onde  $\tau > 0$ , retorne.

Para maiores esclarecimentos veja Murray & Overton (1970).

Como o algoritmo proposto, que apresenta baixo custo computacional, necessita de uma pesquisa linear auxiliar, propõe-se, aqui, a pesquisa linear com interpolação quadrática.

Será apresentado a seguir um algoritmo computacional eficiente de pesquisa linear com interpolação quadrática, baseado nos trabalhos apresentados por Gill & Murray (1974c), Armijo (1966) e Powell (1977).

Desde que  $\nabla p_1(x^k + \tau d^k)^T d^k$  e  $p_1(x^k + \tau d^k)$  são continuamente diferenciáveis com respeito a  $\tau$  no intervalo  $[x^k, x^k + \tau d^k]$ ,  $\tau > 0$ , uma rotina quadrática pode ser usada



com sucesso.

Seja a função\*

$$\gamma(\alpha) = \gamma(x^k + \alpha d^k). \quad (2.65)$$

Esta função pode ser aproximada pela expansão de Taylor, até termos de segunda ordem, nas vizinhanças do ponto  $\alpha=0$ , por

$$\gamma(\alpha) \approx \gamma(0) + \nabla\gamma(0)\alpha + \frac{1}{2} \nabla^2\gamma(0)\alpha^2. \quad (2.66)$$

O vetor gradiente da função  $\gamma(\alpha)$ , da pesquisa linear com interpolação quadrática, pode ser expresso como

$$\nabla\gamma(\alpha) \approx \nabla\gamma(0) + \nabla^2\gamma(0)\alpha. \quad (2.67)$$

O mínimo da função  $\gamma(\alpha)$ , da pesquisa linear, ao longo da direção de pesquisa  $d^k$  ocorre no ponto determinado pelo valor ótimo do comprimento do passo  $\alpha_{0t}$ . Neste ponto, a seguinte equação é satisfeita

$$\nabla\gamma(\alpha_{0t}) = \nabla\gamma(0) + \nabla^2\gamma(0)\alpha_{0t} = 0. \quad (2.68)$$

Após substituir  $\nabla^2\gamma(0)$  da equação (2.67) na equação (2.66), obtém-se

$$\nabla\gamma(\alpha) = 2 \frac{\gamma(\alpha) - \gamma(0)}{\alpha} - \nabla\gamma(0). \quad (2.69)$$

Se  $\gamma(0)$ ,  $\nabla\gamma(0)$ ,  $\gamma(\alpha)$  e  $\alpha$  são conhecidos, então das equações (2.67), (2.68) e (2.69) pode-se obter a seguinte fórmula para se determinar o valor ótimo do comprimento do passo  $\alpha$  ao longo da direção de pesquisa  $d^k$ :

$$\alpha_{0t} = R(\alpha)\alpha, \quad (2.70)$$

onde

$$R(\alpha) = \frac{1}{2} \frac{\nabla\gamma(0)}{\nabla\gamma(0) - \frac{\gamma(\alpha) - \gamma(0)}{\alpha}}.$$

Para se obter um término do processo computacional eficiente da pesquisa linear com interpolação quadrática, as seguintes condições devem ser satisfeitas:

$$\sigma_1 \nabla\gamma(0) < \frac{\gamma(\alpha) - \gamma(0)}{\alpha} < \sigma_2 \nabla\gamma(0), \quad (2.71)$$

onde

$\sigma_1$  - constante que determina o valor mínimo aceitável para o comprimento do passo  $\alpha$  (usualmente se escolhe:  $\sigma_2 < \sigma_1 \leq 1$ ),

$\sigma_2$  - constante que determina o valor máximo aceitável para o comprimento do passo  $\alpha$  (usualmente se escolhe:  $0 < \sigma_2 \leq 0.5$ ).

O parâmetro  $\sigma_2$  não pode ser maior do que 0.5 porque se a função  $p(\alpha)$  é uma função quadrática, o mínimo somente pode

ser encontrado se  $\sigma_2 \leq 0.5$ .

Na prática, a pesquisa linear com interpolação quadrática é feita iterativamente usando a equação (2.70) para consecutivas reduções do comprimento do passo  $\alpha$ :

$$\alpha^1 = 1$$

$$\alpha^{j+1} = R(\alpha^j) \alpha^j, \quad j = 1, \dots, j_{\max},$$

atê que as condições dadas em (2.71) sejam satisfeitas.

## 2.8 Atualização da Matriz Hessiana $H_k$ ou da Matriz Hessiana na Projetada $H_{p_k}$

Usam-se geralmente, duas estratégias para se obter uma sequência de aproximações da matriz Hessiana projetada  $H_{p_k}$  que tende para  $Z^T \nabla^2 L(x^*, \lambda^*) Z$ , ou seja, uma estratégia para um ponto que está em uma região distante do ponto de solução  $x^*$  e outra estratégia para um ponto que está nas proximidades do ponto de solução  $x^*$ .

### 2.8.1 Uma Região Distante do Ponto de Solução $x^*$

Quando em uma região distante do ponto de solução  $x^*$ , desde que a matriz Hessiana projetada permaneça positiva definida e a propriedade de convergência global seja mantida,

pode-se escolher uma entre duas das técnicas que não utilizam a in formação associada aos multiplicadores de Lagrange e que não sejam computacionalmente caras.

a - Técnica 1

Atualiza-se a matriz Hessiana  $H_k$  de acordo com a fórmula BFGS (Broyden-Fletcher-Goldfarb-Shanno) [Powell (1977)]:

$$H_{k+1} = H_k + \frac{\eta^k \eta^k}{\delta^{kT} \eta^k} - \frac{H_k \delta^k \delta^{kT} H_k}{\delta^{kT} H_k \delta^k}, \quad (2.72 a)$$

onde

$$\delta^k = x^{k+1} - x^k,$$

$$\gamma^k = \nabla p_1(x^{k+1}) - \nabla p_1(x^k),$$

$$\xi = \frac{\delta^{kT} \delta^k}{\delta^{kT} H_k \delta^k},$$

$$\theta = \frac{0.8}{1 - \xi},$$

$$\eta^k = \begin{cases} \gamma^k & \text{se } \xi \geq 0.2, \\ \theta \gamma^k + (1 - \theta) H_k \delta^k & \text{se } \xi < 0.2, \end{cases}$$

em seguida calcula-se a matriz Hessiana projetada atualizada

$H_{p_{k+1}}$  de acordo com a equação

$$H_{p_{k+1}} = Z_{k+1}^T H_{k+1} Z_{k+1}. \quad (2.72 b)$$

b - Técnica 2

Quando no novo ponto de pesquisa  $x^{k+1}$ , o conjunto de restrições  $\epsilon$ -ativas  $I_A^\epsilon(x^{k+1})$  é diferente que o conjunto de restrições  $\epsilon$ -ativas no ponto  $x^k$ , ou seja,  $I_A^\epsilon(x^{k+1}) \neq I_A^\epsilon(x^k)$ , aproxima-se a matriz Hessiana projetada  $H_{p_{k+1}}$  de acordo com a seguinte fórmula:

$$H_{p_{k+1}} = Z_{k+1}^T Z_{k+1} \quad (2.73)$$

Quando  $I_A^\epsilon(x^{k+1}) = I_A^\epsilon(x^k)$ , atualiza-se a matriz Hessiana projetada  $H_{p_k}$  de acordo com a seguinte fórmula BFGS.

Coleman & Conn (1984) :

$$H_{p_k} = H_{p_k} + \frac{y^k y^k T}{y^k T s^k} - \frac{H_{p_k} s^k s^k T H_{p_k}}{s^k T H_{p_k} s^k} \quad (2.74)$$

onde

$$s^k = Z_k^T (x^{k+1} - x^k) \quad ,$$

$$y^k = Z_k^T \nabla p_1(x^{k+1}) - \nabla p_1(x^k) \quad .$$

### 2.8.2 Nas Proximidades do Ponto de Solução $x^*$

Quando nas proximidades do ponto de solução  $x^*$ , haverá, em geral, convergência superlinear se a matriz Hessiana projetada  $H_{p_k}$  tenda para  $Z^T \nabla^2 L(x^*, \lambda^*) Z$ , para que isto ocorra,

pode-se escolher uma entre duas das técnicas que utilizam a informação associada aos multiplicadores de Lagrange.

a - Técnica 1

Atualiza-se a matriz Hessiana projetada  $H_{pk}$  de acordo com a seguinte fórmula que utiliza as diferenças dos gradientes ao longo das colunas da matriz  $Z_k$  Gill & Murray (1974a) :

$$H_{p_{k+1}} = \frac{1}{2}(Z_{k+1}^T V_{k+1} + V_{k+1}^T Z_{k+1}), \quad (2.75)$$

onde

$V_{k+1}$  matriz cujas colunas são formadas por

$$\frac{\nabla p_1(x^{k+1} + \xi z_j) - \nabla p_1(x^{k+1})}{\xi},$$

$z_j$  - j-ésima coluna da matriz  $Z_{k+1}$  ,

$\xi$  - pequeno escalar positivo.

b - Técnica 2

Utilizando-se a informação associada aos multiplicadores de Lagrange, atualiza-se a matriz Hessiana projetada  $H_{pk}$  de acordo com a seguinte fórmula BFGS Coleman & Conn(1984) :

$$H_{p_{k+1}} = H_{pk} + \frac{y^k y^k T}{y^k T s^k} - \frac{H_k s^k s^k T H_k T}{s^k T H_k s^k}, \quad (2.76)$$



onde

$x^{k+1} = x^k + h^k$  - ponto intermediário quando o passo horizontal é feito

$$s^k = Z_k^T (x^{k+1} - x^k) ,$$

$$y^k = Z_k^T \left[ \nabla p_1(x^k) - A_{k+1} \lambda^k - \nabla p_1(x^k) \right] ,$$

$(x^{k+1} = x^{k+1} + v^k$  - o novo ponto quando o passo vertical é feito).

Esta técnica é menos cara computacionalmente do que a técnica que utiliza as diferenças dos gradientes, porém, a apresenta menor precisão.

## 2.9 Algumas Considerações Sobre o Algoritmo

O algoritmo da função de penalidade exata globalmente convergente é destinado a minimizar a função de penalidade exata com  $\mu$  fixo. O usuário, entretanto, precisa fornecer os valores iniciais para  $x^0$ ,  $\epsilon^0$ ,  $\lambda^0$ ,  $\delta$  e  $\mu^0$  que são

arbitrários. As propriedades de convergência global não são afetadas por estas atribuições de valores iniciais. Entretanto, a eficiência do algoritmo pode ser adversamente afetada por uma escolha imprópria. Na seção que aborda como tema os resultados numéricos, serão fornecidos os valores u

utilizados para o ponto inicial de pesquisa e os parâmetros do algoritmo.

No algoritmo, o parâmetro  $\delta$  é mantido constante durante todo o processo iterativo e com valor suficientemente pequeno, entretanto, os valores de  $\epsilon$ ,  $\lambda$  e  $\mu$  podem mudar de iteração para iteração.

A estratégia de reduzir  $\epsilon$  e  $\lambda$  quando próximo da solução é arbitrária e com o intuito de reduzir o intervalo para as restrições  $\epsilon$ -ativas.

Em todos os casos existe um valor limite para  $\mu$ , representado por  $\mu_0$ , tal que, se  $\mu \leq \mu_0$ , um mínimo local para a função de penalidade exata  $p(x)$  é também uma solução local para o problema (1.1). Este valor não é conhecido inicialmente, pois é uma função dos multiplicadores de Lagrange. O valor inicial de  $\mu$  escolhido pode ser maior que  $\mu_0$  e gerar uma solução ótima para  $p(x)$  que é inviável para (1.1), ou ainda,  $p(x)$  pode ser ilimitado abaixo. Todavia, a experiência sugere que, quando se reconhece que  $\mu$  é grande durante a minimização de  $p(x)$ , parece adequado fazer uma simples redução de  $\mu$ ,  $\mu \leftarrow \mu/10$ . Um número finito de reduções de  $\mu$  não afeta a propriedade de convergência global do algoritmo.

No algoritmo, pode-se verificar, periodicamente, duas possibilidades:

- (i)  $\{x^k\}$  é convergente para um ponto viável, e
- (ii)  $\{\phi_j(x^k)\} \rightarrow -\infty$  para algum  $j$ .

Para a primeira possibilidade, basta reduzir  $\mu$  e continuar a decrescer  $p(x)$ . Para a segunda possibilidade, reduz-se  $\mu$  e recomeça o procedimento do ponto inicial original. Uma estratégia alternativa para contornar a segunda possibilidade é de primeiro encontrar um ponto viável para o problema (1.1), através da aplicação do próprio algoritmo, substituindo-se a função objetivo por uma função constante. Se um ponto viável é encontrado, então, pode-se escolher este ponto como ponto inicial para o problema (1.22). Estas estratégias não garantem a eficiência do algoritmo, entretanto, elas são razoáveis e tem sido utilizadas com sucesso pelos autores.

## 2.10 Critérios de Convergência

Na prática, pode-se considerar que um mínimo local do problema (1.1) foi alcançado quando as seguintes condições são satisfeitas:

$$(i) \quad \lambda_{1i} \in \left[ -\frac{1}{\mu}, \frac{1}{\mu} \right], \text{ para todo } i \in I_{A1}^{\epsilon} \quad (2.77)$$

$$(ii) \quad \lambda_{2i} \in \left[ 0, \frac{1}{\mu} \right], \text{ para todo } i \in I_{A2}^{\epsilon} \quad (2.78)$$

$$(iii) \quad \| Z_k^T \nabla p_1(x^k) \| \leq \text{tol} \quad (2.79)$$



$$(iv) \quad \|\Phi(x^k)\| \leq \text{tol} \quad (2.80)$$

Powell [Powell (1977)] tem utilizado a seguinte medida de convergência:

$$c = |\nabla p_1(x^k)^T d^k| + \sum_{i=1}^m |\lambda_i \phi_i(x^k)| \leq \text{tol} \quad (2.81)$$

Esta medida de convergência pode ser interpretada da seguinte forma: se o primeiro termo é suficientemente pequeno, significa que a função objetivo não pode decrescer significativamente ao longo da direção de pesquisa  $d^k$ ; se o segundo termo é suficientemente pequeno, significa que as restrições são satisfeitas o suficiente no ponto  $x^k$ .

No algoritmo da função de penalidade exata globalmente convergente pode-se fazer o uso de uma estratégia que verifica, logo após o decréscimo da função de penalidade exata  $p(x^k)$  ao longo de uma direção de pesquisa vertical, se as condições dadas pelas equações (2.79) e (2.80) são satisfeitas e, opcionalmente, se a condição dada pela equação (2.81) é satisfeita.

### 3 OPERAÇÕES COMPUTACIONAIS RELEVANTES AO DESENVOLVIMENTO DO ALGORITMO COMPUTACIONAL

#### 3.1 Fatoração Ortogonal da Matriz dos Gradientes das Restrições Ativas $A_k$

A fatoração ortogonal da matriz dos gradientes das restrições ativas  $A_k$  é utilizada na determinação da direção de pesquisa vertical resolvendo o sistema hipodeterminado de equações dado por (2.55) e na determinação dos multiplicadores de Lagrange resolvendo o sistema hiperdeterminado de equações dado por (2.34). Em adição, será mostrado que a matriz  $Z_k$  é obtida da matriz ortogonal  $Q_k$ , consequentemente, as direções de pesquisa horizontais dadas pelas equações (2.54) e (2.56) a utilizam em sua determinação.

A fatoração ortogonal da matriz  $A_k$  dada pela equação (2.36) pode ser encontrada em duas etapas computacionais distintas. Na primeira etapa, efetua-se a fatoração ortogonal da primeira coluna da matriz  $A_k$ . Na segunda etapa, adiciona-se, consecutivamente, coluna por coluna, as demais colunas da matriz  $A_k$ , à primeira coluna e atualiza-se a fatoração ortogonal de acordo com o esquema apresentado na seção 3.2.1.

A fatoração ortogonal da primeira coluna da matriz  $A_k$  é da forma

$$Q_0 a_1 = \begin{bmatrix} r_0 \\ 0 \end{bmatrix}$$

onde

$$Q_0 = \begin{bmatrix} H_u & \vdots & 0 \\ \cdots & \vdots & \cdots \\ 0 & \vdots & I \end{bmatrix} \quad - \text{ matriz ortogonal de ordem } n$$

$H_u$  - matriz de Hessenberg superior de ordem  $h$

$I$  - matriz identidade de ordem  $\ell$

$r_0$  - matriz triangular superior de ordem 1

$a_1$  - primeira coluna da matriz  $A_k$ ;

$$\ell = n - h \geq 0.$$

A dimensão  $\ell$  é maior que zero somente se a parte inferior do vetor  $a_1$  for um subvetor nulo, então  $\ell$  é a dimensão deste subvetor. A matriz de Hessenberg superior  $H_u$  é derivada do produto de uma sequência de matrizes de Givens embutidas em matrizes identidades  $P_{i+1}^i$ ,  $i = 1, \dots, h-1$ ,

$$H_u = P_2^1 \cdot P_3^2 \cdot P_4^3 \cdots P_h^{h-1}. \quad (3.1)$$

As matrizes de Givens embutidas nas matrizes identida





$$\bar{z}_i = \Delta \quad e \quad (3.5)$$

$$\bar{z}_{i+1} = 0.$$

A transformação feita por uma matriz de Givens corresponde a uma rotação no plano seguido por uma reflexão sobre um eixo. Se todos os elementos do vetor  $s$  (um dos vetores que definem as matrizes  $P_{i+1}^i$ ,  $i = 1, \dots, h - 1$ ) são diferentes de zero, pode-se tirar vantagem de uma estrutura especial da matriz de Hessenberg superior  $H_u$ , dada pela equação (3.1), e esta matriz pode ser apresentada como uma matriz de Hessenberg superior na forma especial  $H_u(\rho, \beta, \gamma)$  definida pelos vetores  $\rho$ ,  $\beta$  e  $\gamma$

$$H_u = H_u(\rho, \beta, \gamma) = \begin{bmatrix} \rho_1 \beta_1 & \rho_2 \beta_1 & \cdots & \rho_{h-1} \beta_1 & \rho_h \beta_1 \\ \gamma_1 & \rho_2 \beta_2 & \cdots & \rho_{h-1} \beta_2 & \rho_h \beta_2 \\ & \gamma_2 & \cdots & \rho_{h-1} \beta_3 & \rho_h \beta_3 \\ & & & \vdots & \\ & 0 & & \rho_{h-1} \beta_{h-1} & \rho_h \beta_{h-1} \\ & & & \gamma_{n-1} & \rho_h \beta_h \end{bmatrix} \quad (3.6)$$

Para uma dada sequência de matrizes  $P_{i+1}^i$ ,  $i = 1, \dots, h - 1$ , os vetores  $\rho$ ,  $\beta$  e  $\gamma$  são definidos por qualquer uma

das seguintes relações de recorrência:

### Recorrência progressiva

Passo 1, faça:  $\gamma_1 \leftarrow s_1$

$$\rho_1 \leftarrow c_1$$

$$\beta_1 \leftarrow 1$$

$$\eta_1 \leftarrow s_1$$

Passo 2, calcule:

$$\left. \begin{array}{l} \gamma_i \leftarrow s_i \\ \rho_i \leftarrow c_i \eta \\ \beta_i \leftarrow -c_{i-1} / \eta \\ \eta \leftarrow s_i \eta \end{array} \right\} i = 2, \dots, h-1$$

$$\rho_h \leftarrow \eta$$

$$\beta_h \leftarrow -c_{h-1} / \rho_h$$

### Recorrência regressiva

Passo 1, faça:  $\gamma_{h-1} \leftarrow s_{h-1}$

$$\rho_h \leftarrow 1$$

$$\beta_h \leftarrow -c_{h-1}$$

$$\eta \leftarrow s_{h-1}$$

$$\begin{array}{l}
 \text{Passo 2, calcule:} \\
 \left. \begin{array}{l}
 \gamma_{i-1} \leftarrow s_{i-1} \\
 \rho_i \leftarrow c_i/\eta \\
 \beta_i \leftarrow -c_{i-1}/\eta \\
 \eta \leftarrow s_{i-1}\eta
 \end{array} \right\} i = h-1, \dots, 2 \\
 \\
 \beta_1 \leftarrow \eta \\
 \rho_1 \leftarrow c_1/\beta_1
 \end{array}$$

### 3.2 Atualização da Fatoração Ortogonal da Matriz dos Gradientes das Restrições Ativas $A_k$

Uma mudança de base do subespaço definido pelos gradientes das restrições ativas implica na inclusão ou exclusão de uma coluna da matriz dos gradientes das restrições ativas  $A_k$ , conseqüentemente a fatoração ortogonal da matriz  $A_k$  deverá ser atualizada [Gill & Murray (1978)].

#### 3.2.1 Inclusão de uma Coluna na Matriz $A_k$

Após incluir uma nova restrição no conjunto das restrições ativas, cujo gradiente seja linearmente independente dos outros gradientes das restrições ativas que formam a base do subespaço  $S_{\perp}^E$ , a nova matriz dos gradientes e sua fatoração ortogonal são:

$$A_{k+1} = [A_k, a_L] = Q_{k+1}^T \begin{bmatrix} R_{k+1} \\ \dots \\ 0 \end{bmatrix}, \quad (3.7)$$

onde

$$A_k = Q_k^T \begin{bmatrix} R_k \\ \dots \\ 0 \end{bmatrix} \quad - \quad \text{antiga matriz dos gradientes das restrições} \\ \text{ativas com dimensão } n \times t^k \text{ e sua conhecida fatoração ortogonal,}$$

$a_L$             - vetor gradiente da nova restrição ativa com dimensão  $n \times 1$ .

Pré-multiplicando ambos os membros da equação (3.7) pela matriz ortogonal  $Q_k$ , obtem-se

$$Q_k A_{k+1} = [Q_k A_k, Q_k a_L] = \begin{bmatrix} R_k & , & Q_{1k} & a_L \\ \dots & \dots & \dots & \dots \\ 0 & , & Q_{2k} & a_L \end{bmatrix}, \quad (3.8)$$

onde

- $Q_{1k}$  - submatriz superior com dimensão  $t^k \times n$  da matriz ortogonal  $Q_k$ ,
- $Q_{2k}$  - submatriz inferior com dimensão  $(n - t^k) \times n$  da matriz ortogonal  $Q_k$ .

Ambos os membros da equação (3.8) podem ser pré-multiplicados pela matriz ortogonal  $\tilde{Q}_k$  que reduz o vetor  $Q_{2k} a_L$  para um múltiplo da primeira coluna da matriz identidade de ordem  $(n - t^k)$  dando

$$Q_{k+1} A_{k+1} = \tilde{Q}_k Q_k A_{k+1} = \begin{bmatrix} R_k & \vdots & Q_{1k} a_L \\ \dots & \dots & \dots \\ 0 & \vdots & u \\ \dots & \dots & \dots \\ 0 & \vdots & 0 \end{bmatrix} = \begin{bmatrix} R_{k+1} \\ \dots \\ 0 \end{bmatrix}, \quad (3.9)$$

onde

$$Q_{k+1} = \tilde{Q}_k Q_k - \text{matriz ortogonal de ordem } n$$

$$R_{k+1} = \begin{bmatrix} R_k & \vdots & Q_{1k} a_L \\ \dots & \dots & \dots \\ 0 & \vdots & u \end{bmatrix} - \text{matriz triangular superior de ordem } t^{k+1}$$

O gradiente da nova restrição  $\epsilon$ -ativa não é incluído na base do subespaço definido pelos gradientes das restrições ativas quando

$$u \leq \theta r_{\max},$$

onde

$\theta$  - pequeno escalar positivo,

$$r_{\max} = \max_i r_{ii}, \quad 1 \leq i \leq t^k,$$

$r_{ii}$  - elemento diagonal da matriz triangular superior  $R_k$ .



Neste caso, assume-se que o gradiente da nova restrição ativa  $\bar{e}$  é uma combinação linear dos gradientes das restrições ativas que formam a base do subespaço  $S_{\underline{1}}^E$ .

A matriz ortogonal transformada  $\tilde{Q}_k$  é da forma,

$$\tilde{Q}_k = \begin{bmatrix} \tilde{Q}_{1k} & \vdots & 0 \\ \dots & \dots & \dots \\ 0 & \vdots & \tilde{Q}_{2k} \end{bmatrix}, \quad (3.10)$$

onde

$$\tilde{Q}_{1k} = I_1,$$

$$\tilde{Q}_{2k} = \begin{bmatrix} H_u & \vdots & 0 \\ \dots & \dots & \dots \\ 0 & \vdots & I_2 \end{bmatrix},$$

$I_1$  - matriz identidade de ordem  $t^k$

$H_u$  - matriz de Hessenberg superior de ordem  $h$

$I_2$  - matriz identidade de ordem  $\ell$

$$\ell = n - t^k - h \geq 0.$$

A dimensão  $\ell$  é maior que zero somente se a parte inferior do vetor  $Q_{2k} a_L$  for um subvetor nulo, então  $\ell$  é a dimensão deste subvetor.

A matriz de Hessenberg superior  $H_u$  é o produto de uma

sequência de matrizes de Givens embutidas nas matrizes iden-  
tidades dada pelas equações (3.1) e (3.2). Se todos os ele-  
mentos do vetor  $s$  (um dos dois vetores que definem as matri-  
zes  $P_{i+1}^i$ ,  $i = 1, \dots, h - 1$ ) são diferentes de zero, a ma-  
triz de Hessenberg superior pode ser apresentada como uma ma-  
triz de Hessenberg superior na forma especial  
 $H_U(\rho, \beta, \gamma)$ , definida pelos vetores  $\rho$ ,  $\beta$ , e  $\gamma$  (dada pela equa-  
ção (3.6)).

Sendo a matriz de Hessenberg superior dada em sua for-  
ma especial  $H_U(\rho, \beta, \gamma)$ , então a matriz atualizada  $Q_{k+1} =$   
 $= \tilde{Q}_k Q_k$  pode ser calculada eficientemente usando o processo  
de recorrência regressiva similar a proposta na seção 3.1.

Para calcular eficientemente a matriz  $B = H_U(\rho, \beta, \gamma)A$   
com dimensão  $h \times t$ , onde  $A$  é uma matriz geral com dimensão  $h \times t$ ,  
podem ser usados as seguintes relações de recorrência:

#### Recorrência progressiva

$$\begin{aligned} \text{Calcule } w &\leftarrow A^T \rho \\ b_{1,*} &\leftarrow \beta_1 w^T \\ \left. \begin{aligned} w_i &\leftarrow w_i - \rho_{j-1} a_{j-1,i} \\ b_{j,i} &\leftarrow \gamma_{j-1} a_{j-1,i} + \beta_j w_i \end{aligned} \right\} \begin{aligned} i &= 1, \dots, t \\ j &= 2, \dots, h \end{aligned} \end{aligned}$$

#### Recorrência regressiva

$$\text{Calcule } w^T \leftarrow \rho_h a_{h,*}$$

$$\left. \begin{aligned} b_{j,i} &\leftarrow \gamma_{j-1} a_{j-1,i} + \beta_j w_i \\ w_i &\leftarrow w_i + \rho_{j-1} a_{j-1,i} \\ b_{1,*} &\leftarrow \beta_1 w^T \end{aligned} \right\} \begin{array}{l} i = 1, \dots, t \\ j = h, \dots, 2 \end{array}$$

### 3.2.2 Exclusão de uma Coluna da Matriz $A^k$

A exclusão da  $j$ -ésima coluna ( $a_d$ ) da matriz dos gradientes  $A_k$  equivale a remoção da  $j$ -ésima restrição do conjunto das restrições ativas e a remoção da  $j$ -ésima coluna ( $r_d$ ) da matriz triangular superior  $R_k$  (uma das matrizes componentes da fatoração ortogonal da matriz  $A_k$ ).

As matrizes  $A_k$  e  $R_k$  podem ser particionadas da seguinte forma:

$$A_k = \begin{bmatrix} j-1 & 1 & t^{k-j} \\ A_{1_k} & a_d & A_{2_k} \end{bmatrix},$$

$$R_k = \begin{bmatrix} j-1 & 1 & t^{k-j} \\ R_{1_k} & r_d & R_{2_k} \end{bmatrix}.$$

$$\text{Se } Q_k A_k = \begin{bmatrix} R_k \\ \dots \\ 0 \end{bmatrix}, \quad (3.11)$$

então

$$Q_k A_{k+1} = \begin{bmatrix} \bar{R}_{k+1} \\ \dots \\ 0 \end{bmatrix} \quad (3.12)$$

onde

$$A_{k+1} = \begin{bmatrix} A_{1k} & \vdots & A_{2k} \end{bmatrix} \quad e$$

$$\bar{R}_{k+1} = \begin{bmatrix} R_{1k} & \vdots & R_{2k} \end{bmatrix} = \begin{array}{c} \begin{array}{|c|} \hline \text{[Matrix Diagram]} \\ \hline \end{array} \\ \left. \begin{array}{l} \text{] } j-1 \\ \text{] } 1 \\ \text{] } t^{k-1} \end{array} \right\} \end{array}$$

Ambos os membros da equação (3.12) podem ser pré-multiplicados por uma matriz ortogonal de transformação  $\tilde{Q}_k$  que transforma a matriz  $\bar{R}_{k+1}$  em uma matriz triangular superior  $R_{k+1}$  pela redução dos seus elementos subdiagonais a zero. Como resultado, é obtida a fatoração ortogonal da matriz dos gradientes atualizada  $A_{k+1}$ .

$$Q_{k+1} A_{k+1} = \tilde{Q}_k \tilde{Q}_k A_{k+1} = \tilde{Q}_k \begin{bmatrix} \bar{R}_{k+1} \\ \dots \\ 0 \end{bmatrix} = \begin{bmatrix} R_{k+1} \\ 0 \end{bmatrix} \quad (3.13)$$

A matriz ortogonal  $\tilde{Q}_k$  é da forma



$$\tilde{Q}_k = \begin{bmatrix} I_1 & \vdots & 0 & \vdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \vdots & H_L & \vdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \vdots & 0 & \vdots & I_2 \end{bmatrix} \quad (3.14)$$

onde

$I_1$  - matriz identidade de ordem  $(j - 1)$

$H_L$  - matriz de Hessenberg inferior de ordem  $h$

$I_2$  - matriz identidade de ordem  $(n - t^k)$

$$h = t^k - j + 1$$

A matriz de Hessenberg inferior é o produto de uma sequência das matrizes de Givens embutidas nas matrizes identidades  $P_{i+1}^i$ ,  $i = j, \dots, t^k - 1$ , (veja a equação (3.2)),

$$H_L = P_{t^k}^{t^k-1} \cdots P_{j+2}^{j+1} P_{j+1}^j. \quad (3.15)$$

Quando todos os elementos do vetor  $s$  (um dos dois vetores que definem as matrizes  $P_{i+1}^i$ ,  $i = j, \dots, t^k - 1$ ) são diferentes de zero, a matriz de Hessenberg inferior pode ser dada em sua forma especial  $H_L(\rho, \beta, \gamma)$ , onde  $H_L(\rho, \beta, \gamma) = H_U(\rho, \beta, \gamma)^T$ . (veja a equação (3.6)).

Para uma dada sequência de matrizes  $P_{i+1}^i$ ,  $i = j, \dots,$

$t^k - 1$ , os vetores  $\rho$ ,  $\beta$  e  $\gamma$  são definidos pela relação de recorrência mostrada na seção 3.1.

A matriz ortogonal  $Q_{k+1}$  tem a forma

$$Q_{k+1} = \begin{bmatrix} Q_{1_{k+1}} \\ \dots \\ Q_{2_{k+1}} \end{bmatrix} = \tilde{Q}_k Q_k = \begin{bmatrix} Q_{11_k} \\ \dots \\ H_L Q_{12_k} \\ \dots \\ Q_{2_k} \end{bmatrix}, \quad (3.16)$$

onde

$Q_{1_{k+1}}$  - submatriz da matriz ortogonal  $Q_{k+1}$  com dimensão  $t^{k+1} \times n$ ,

$Q_{2_{k+1}}$  - submatriz da matriz ortogonal  $Q_{k+1}$  com dimensão  $(n - t^{k+1}) \times n$ ,

$Q_{11_k}$  - submatriz da matriz ortogonal  $Q_k$  com dimensão  $(j - 1) \times n$ ,

$Q_{12_k}$  - submatriz da matriz ortogonal  $Q_k$  com dimensão  $h \times n$ ,

$Q_{2_k}$  - submatriz da matriz ortogonal  $Q_k$  com dimensão  $(n - t^k) \times n$ .

Se as linhas da submatriz  $Q_{2_{k+1}}$  são comparadas com as linhas equivalentes da matriz  $Q_k$  (veja equação (3.16)), observa-se que somente a primeira linha da submatriz  $Q_{2_{k+1}}$  é tro



cada (esta linha é a última linha da submatriz  $H_L Q_{12_k}$ ); se esta linha é denotada por  $z^T$ , então

$$Q_{2_{k+1}} = \begin{bmatrix} z^T \\ \dots \\ Q_{2_k} \end{bmatrix} \quad (3.17)$$

### 3.3 A Escolha da Matriz $Z_k$

Da fatoração ortogonal da matriz dos gradientes das restrições ativas  $A_k$  obtem-se a matriz  $Z_k$ .

Considerando-se a equação (2.36) e a partição da matriz  $Q$  dada pela equação (2.37), tem-se

$$A_k = R_k^T Q_{1_k} \quad (3.18)$$

Pós-multiplicando membros os lados desta equação por  $Q_{2_k}^T$ , obtêm-se

$$A_k^T Q_{2_k}^T = R_k^T Q_{1_k} Q_{2_k}^T = 0. \quad (3.19)$$

Este resultado mostra que as colunas da submatriz  $Q_{2_k}^T$  são ortogonais aos vetores normais das restrições ativas e podem ser usados como as colunas da matriz  $Z_k$ . As colunas da

matriz  $Z_k$  não precisam estar na mesma ordem que as colunas da submatriz  $Q_{2k}^T$ . Então, a matriz pode ser escolhida como sendo

$$Z_k = Q_{2k}^T \tilde{I}, \quad (3.20)$$

onde

$\tilde{I}$  - matriz identidade de ordem  $n - t^k$  com as colunas na ordem reversa.

As colunas da matriz  $Z_k$  são escolhidas na ordem reversa das colunas da submatriz  $Q_{2k}^T$  para facilitar o processo de atualização da fatoração  $LDL^T$  da matriz Hessiana projetada  $H_{p_k}$  quando for mudada a base do subespaço gerado pelos gradientes das restrições ativas.

### 3.4 Atualização da Matriz $Z_k$

Uma mudança da base do subespaço definido pelos gradientes das restrições ativas implica em uma atualização da matriz dos gradientes das restrições ativas  $A_k$  e, conseqüentemente, deve-se atualizar a matriz  $Z^k$ .

#### 3.4.1 Inclusão de uma Coluna na Matriz $A_k$

As  $(n - t^{k+1})$  últimas linhas da matriz ortogonal atua

lizada  $Q_{k+1}$ , na ordem reversa, constituem as colunas da matriz  $Z_{k+1}$ , ou seja,

$$Z_{k+1} = Q_{2_{k+1}}^T \tilde{I}, \quad (3.21)$$

onde

- $Q_{2_{k+1}}$  - submatriz inferior da matriz ortogonal  $Q_{k+1}$  com dimensão  $(n - t^{k+1}) \times n$ ,
- $\tilde{I}$  - matriz identidade de ordem  $n - t^{k+1}$  com suas colunas na ordem reversa.

### 3.4.2 Exclusão de uma Coluna da Matriz $A_k$

As últimas  $(n - t^{k+1})$  linhas da matriz ortogonal atualizada  $Q_{k+1}$ , na ordem reversa, constituem as colunas da matriz atualizada  $Z_{k+1}$ , ou seja,

$$Z_{k+1} = Q_{2_{k+1}}^T \tilde{I}, \quad (3.22)$$

onde

- $Q_{2_{k+1}}$  - submatriz inferior da matriz ortogonal  $Q_{k+1}$  com dimensão  $(n - t^{k+1}) \times n$ ,
- $\tilde{I}$  - matriz identidade de ordem  $n - t^{k+1}$  com suas colunas na ordem reversa.

Se a equação (3.17) é considerada, então a atualização da matriz  $Z^{k+1}$  dada pela equação (3.22) torna-se

$$Z_{k+1} = \begin{bmatrix} z^T \\ \dots \\ Q_{2k} \end{bmatrix} \tilde{I} = \begin{bmatrix} z \\ \vdots \\ Q_{2k}^T \end{bmatrix} \tilde{I} = \begin{bmatrix} Z_k & : & z \end{bmatrix}. \quad (3.23)$$

### 3.5 Fatoração $LDL^T$ de uma Matriz Simétrica Positiva Definida $H_p$

A fatoração  $LDL^T$  (fatoração de Cholesky) de uma matriz simétrica positiva definida  $H_n$  é da forma

$$LDL^T \leftarrow H_p \quad (3.24)$$

O algoritmo computacional para efetuar esta fatoração [Martin et alli (1965)] é mostrada a seguir.

#### Notação

$H_p(N, N)$  - matriz simétrica positiva definida

$L(N, N)$  - matriz triangular inferior e unitária

$D(N)$  - vetor dos elementos diagonais positivos da matriz diagonal  $D(N, N)$



AUX(N,N) - matriz auxiliar

Algoritmo

1. Faça  $D(1)=H_p(1,1)$

$$\left. \begin{array}{l} \text{se } (I < N), L(I,J)=0 \quad J=I+1, \dots, N \\ \\ L(I,I)=1 \end{array} \right\} I=1, \dots, N$$

2. Calcule  $AUX(I,1)=H_p(I,1)$

$$\left. \begin{array}{l} \\ \\ L(I,1)=AUX(I,1)/D(1) \end{array} \right\} I=2, \dots, N$$

$$\left. \begin{array}{l} \text{se } I=N, \text{ PARE} \\ \\ D(I)=H_p(I,I) - \sum_{K=1}^{I-1} AUX(I,K)*L(I,K) \\ \\ \left. \begin{array}{l} AUX(J,I)=H_p(J,I) - \sum_{K=1}^{I-1} AUX(J,K)*L(I,K) \\ \\ L(J,I)=AUX(J,I)/D(I) \end{array} \right\} J=I+1, \dots, N \end{array} \right\} I=2, \dots, N$$

### 3.6 Fatoração $LDL^T$ Modificada de uma Matriz Simétrica $H_p$

A fatoração  $LDL^T$  modificada (fatoração de Cholesky mo  
dificada) de uma matriz simétrica  $H_p$  é da forma

$$LDL^T - E \leftarrow H_p \quad (3.25)$$



O algoritmo computacional para efetuar esta fatoração é baseado no trabalho apresentado por Gill & Murray (1984c). O algoritmo é idêntico ao algoritmo da fatoração  $LDL^T$  (fatoração de Cholesky) exceto o fato de limitar o valor dos elementos fora da diagonal da matriz  $LD^{1/2}$  quando a matriz  $H_p$  não for positiva definida e de assegurar que todos os elementos diagonais da matriz  $D$  sejam positivos. Cada elemento fora da diagonal da matriz  $LD^{1/2}$  é menor em módulo que um parâmetro BETA. A matriz  $E$  é uma matriz diagonal que é zero quando a matriz  $H_p$  é suficientemente positiva definida. Um parâmetro DELTA é introduzido para otimizar o processo de fatoração quando a matriz  $H_p$  é positiva definida mas muito mal-condicionada.

#### Notação

$H_p(N, N)$	- matriz simétrica
$L(N, N)$	- matriz triangular inferior e unitária
$C(N, N)$	- matriz auxiliar
$D(N, N)$	- vetor dos elementos diagonais da matriz diagonal positiva $D(N, N)$

- E(N) - vetor dos elementos diagonais da matriz diagonal não-negativa E(N, N)
- BETAS = max(GAMMA, LAMBDA, TAU)
- BETA = BETAS<sup>1/2</sup> - parâmetro que limita o valor dos elementos diagonais da matriz LD<sup>1/2</sup>
- GAMMA - máximo valor (em módulo) dos elementos diagonais da matriz H<sub>p</sub>
- KSI - máximo valor (em módulo) dos elementos fora da diagonal da matriz A
- LAMBDA = KSI/N
- ANORM - norma-1 da matriz H<sub>p</sub>
- TAU - precisão relativa da máquina (computador) definido como sendo o menor número TAU tal que  $1 + \text{TAU} > 1$

Algoritmo

1. Faça:  $\text{TAU} = 2.2 \cdot 10^{-6}$  (no caso de equipamentos IBM)

2. Calcule:  $\text{GAMMA} = \max(|H_p(I,I)|)$  ,  $I = 1, \dots, N$

$\text{KSI} = \max(|H_p(I,J)|)$  ,  $J = 1, \dots, I-1$  e  $I = 2, \dots, N$

$\text{LAMBDA} = \text{KSI}/N$

$\text{BETAS} = \max(\text{GAMMA}, \text{LAMBDA}, \text{TAU})$

$\text{ANORM} = \max(\sum_{I=1}^N |H_p(I,J)|)$  ,  $J = 1, \dots, N$

$\text{DELTA} = \max(\text{TAU} \cdot \text{ANORM}, \text{TAU})$

se  $I = N$ , vá para 3

$L(I,J) = 0$  ,  $J = I + 1, \dots, N$  ,  $I = 1, \dots, N$

3.  $L(I,I) = 1$  ,  $I = 1, \dots, 1$

$\text{PHI}(1) = H_p(1,1)$

$\text{PHIAB} = |\text{PHI}(1)|$

se  $N = 1$ , vá para 5

$$\left. \begin{aligned} C(I,1) &= H_p(I,1) \\ \text{THETAS} &= (\max(|C(I,1)|))^2 \end{aligned} \right\} I = 2, \dots, N$$

$$R = \text{THETAS}/\text{BETAS}$$

$$D(1) = \max(\text{DELTA}, \text{PHIAB}, R)$$

$$E(1) = D(1) - \text{PHI}(1)$$

$$L(I,1) = C(I,1)/D(1), \quad I = 2, \dots, N$$

4. Calcule:  $\text{PHI}(I) = H_p(I,I) - \sum_{K=1}^{I-1} C(I,K)*L(I,K)$

$$\text{PHIAB} = |\text{PHI}(I)|$$

se  $I = N$ , vá para 5

$$\left. \begin{aligned} C(J,I) &= H_p(J,I) - \\ &- \sum_{K=1}^{I-1} C(J,K)*L(I,K), \quad J = I+1, \dots, N \end{aligned} \right\} I = 2, \dots, N$$

$$\text{THETAS} = (\max(|C(J,I)|))^2$$

$$R = \text{THETAS}/\text{BETAS}$$

$$D(I) = \max(\text{DELTA}, \text{PHIAB}, R)$$

$$E(I) = D(I) - \text{PHI}(I)$$

$$L(J,I) = C(J,I)/D(I), \quad J = I+1, \dots, N$$

5. Calcule:  $D(N) = \max(\text{DELTA}, \text{PHIAB})$

$$E(N) = D(N) - \text{PHI}(N)$$

### 3.7 Atualização da Fatoração $LDL^T$ da matriz Hessiana Projetada

A fatoração  $LDL^T$  da matriz Hessiana projetada simétrica positiva definida  $H_p$  é utilizada na determinação da direção de pesquisa horizontal dada pela equação (2.54) e pode ser expressa por

$$LDL^T \leftarrow H_p, \quad \text{onde } H_p = Z^T H Z. \quad (3.26)$$

A cada iteração do processo computacional, os fatores da matriz triangular inferior unitária  $L$  e da matriz diagonal  $D$  devem ser atualizados para incluírem a nova informação sobre a mudança nos gradientes (veja a seção 2.8) e a nova informação que reflete a mudança na matriz  $Z$ . Para incluir esta nova informação que reflete a mudança na matriz  $Z$  deve-se considerar três subproblemas:

(i) a inclusão dos gradientes das restrições ativas à base



do subespaço  $S_{\perp}^E$  que implica na diminuição da dimensão da matriz Hessiana projetada  $H_p$ ;

(ii) a exclusão dos gradientes das restrições ativas da base do subespaço  $S_{\perp}^E$  que implica no aumento da dimensão da matriz Hessiana projetada  $H_p$ ; e

(iii) a não mudança da dimensão do subespaço  $S_{\perp}^E$  que implica na manutenção da dimensão da matriz Hessiana projetada.

### 3.7.1 Atualização da Fatoração $LDL^T$ com Diminuição da Dimensão da Matriz $H_p$

A antiga matriz Hessiana projetada é

$$H_{p_k} = Z_k^T H Z_k = L_k D_k L_k^T, \quad (3.27)$$

e a nova matriz Hessiana projetada é

$$H_{p_{k+1}} = Z_{k+1}^T H Z_{k+1} = L_{k+1} D_{k+1} L_{k+1}^T \quad (3.28)$$

A seguinte equação é satisfeita para uma matriz  $\bar{Z}_{k+1}$  (matriz onde as primeiras  $(n-t^{k+1})$  colunas são idênticas as colunas da matriz  $Z_{k+1}$ ):

$$\bar{Z}_{k+1} = \begin{bmatrix} Z_{k+1} \\ \vdots \\ z \end{bmatrix} \quad (3.29 a)$$

$$\begin{aligned}
&= (\tilde{Q}_{2_k} Q_{2_k})^T \tilde{I} \quad , \quad \text{da equação (3.23)} \\
&= Q_{2_k}^T \tilde{Q}_{2_k}^T \tilde{I} \quad , \\
&= Q_{2_k}^T \tilde{I} \tilde{I} \tilde{Q}_{2_k}^T \tilde{I} \quad , \\
&= Z_k \tilde{I} Q_{2_k}^T \tilde{I} \quad , \quad \text{da equação (3.20)} \quad , \quad (3.29 \text{ b})
\end{aligned}$$

onde a matriz  $\tilde{I} Q_{2_k}^T \tilde{I}$  pode ser encontrada da seguinte forma:

$$\begin{aligned}
\tilde{I} Q_{2_k}^T \tilde{I} &= \begin{bmatrix} 0 & \vdots & \tilde{I}_2 \\ \dots & \dots & \dots \\ \tilde{I}_1 & \vdots & 0 \end{bmatrix} \begin{bmatrix} H_u & \vdots & 0 \\ \dots & \dots & \dots \\ 0 & \vdots & I_2 \end{bmatrix} \begin{bmatrix} 0 & \vdots & \tilde{I}_1 \\ \dots & \dots & \dots \\ \tilde{I}_2 & \vdots & 0 \end{bmatrix} \quad (3.30) \\
&= \begin{bmatrix} \tilde{I}_2 I_2 \tilde{I}_2 & \vdots & 0 \\ \dots & \dots & \dots \\ 0 & \vdots & \tilde{I}_1 H_u \tilde{I}_1 \end{bmatrix} = \begin{bmatrix} I_2 & \vdots & 0 \\ \dots & \dots & \dots \\ 0 & \vdots & \tilde{H}_L \end{bmatrix}
\end{aligned}$$

Se a matriz de Hessenberg superior  $H_u$  é dada em sua forma especial  $H_u(\rho, \beta, \gamma)$ , então a matriz de Hessenberg inferior  $\tilde{H}_L$  pode ser dada em sua forma especial  $\tilde{H}_L(\bar{\rho}, \bar{\beta}, \bar{\gamma})$ , ou seja,  $\tilde{H}_L(\bar{\rho}, \bar{\beta}, \bar{\gamma}) = \tilde{H}_u(\bar{\rho}, \bar{\beta}, \bar{\gamma})^T$  (veja equação (3.1)) ,

$$\tilde{H}_L(\bar{\rho}, \bar{\beta}, \bar{\gamma}) = \tilde{I}_1 H_u(\bar{\rho}, \bar{\beta}, \bar{\gamma}) \tilde{I}_1 \quad , \quad (3.31)$$

onde

$$\bar{\rho} = \tilde{I}_1 \beta \quad ,$$

$$\bar{\beta} = \tilde{I}_1 \rho \quad ,$$

$$\bar{\gamma} = \tilde{I}_1 \gamma \quad .$$

A nova matriz Hessiana projetada  $H_{p_{k+1}}$  é uma submatriz da matriz  $\bar{H}_{p_{k+1}}$  definida como

$$\bar{H}_{p_{k+1}} = \bar{Z}_{k+1}^T H \bar{Z}_{k+1} \quad (3.32 a)$$

$$= \begin{bmatrix} Z_{k+1}^T \\ \dots \\ z^T \end{bmatrix} H \begin{bmatrix} Z_{k+1} \\ \dots \\ z \end{bmatrix}$$

$$= \begin{bmatrix} H_{p_{k+1}} & \vdots & Z_{k+1}^T H z \\ \dots & \dots & \dots \\ z^T H Z_{k+1} & \vdots & z^T H z \end{bmatrix} \quad (3.32 b)$$

Considerando-se as equações (3.32 a) e (3.29 b) a matriz  $\bar{H}_{p_{k+1}}$  é dada por

$$\begin{aligned} \bar{H}_{p_{k+1}} &= \tilde{I} Q_{2k} \tilde{I} Z_k^T H Z_k \tilde{I} Q_{2k}^T \tilde{I} \\ &= \tilde{I} Q_{2k} \tilde{I} L_k D_k^{1/2} D_k^{1/2} L_k^T \tilde{I} Q_{2k}^T \tilde{I} \end{aligned}$$







A fatoração  $LDL^T$  atualizada da matriz Hessiana projetada atualizada pode ser obtida se a transformação ortogonal for aplicada a matriz  $C$ , que resulta em uma matriz triangular inferior  $L_{k+1} D_{k+1}^{1/2}$ , dada por

$$H_{P_{k+1}} = CC^T = C\tilde{P}^T C^T = \begin{bmatrix} L_{k+1} D_{k+1}^{1/2} & 0 \end{bmatrix} \begin{bmatrix} D_{k+1}^{1/2} & L_{k+1}^T \\ \dots & \dots \\ 0 & \dots \end{bmatrix}, \quad (3.38)$$

$$= L_{k+1} D_{k+1} L_{k+1}^T,$$

onde

$$\tilde{P}^T C^T = \tilde{P}^T \begin{array}{c} \begin{array}{|c|} \hline \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \\ \hline \end{array} \begin{array}{c} l+h-1 \\ l \\ 1 \\ h-1 \end{array} = \begin{array}{c} \begin{array}{|c|} \hline \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \\ \hline \end{array} \begin{array}{c} t+h-1 \\ t+h-1 \\ 1 \end{array} = \begin{bmatrix} D_{k+1}^{1/2} L_{k+1}^T \\ \dots \\ 0 \end{bmatrix}$$

$$\text{e } \tilde{P}^T = \begin{array}{c} p_{l+h-1}^{l+h-1} \dots p_{l+3}^{l+2} p_{l+2}^{l+1} \\ p_{l+h} \dots p_{l+3} p_{l+2} \end{array}$$

A matriz de transformação  $\tilde{P}$  dada pela equação (3.2) é um produto das matrizes de Givens embutidas nas matrizes identidade,  $P_{i+1}^i$ ,  $i = l+1, \dots, l+h-1$ .

A matriz  $\tilde{P}$  é utilizada exclusivamente para transformar a matriz  $C$  e, conseqüentemente, não precisa ser armazenada em sua forma explícita. A matriz  $\tilde{P}$  não é utilizada em cálculos posteriores.

### 3.7.2 Atualização da Fatoração $LDL^T$ com Aumento da Dimensão da Matriz $H_p$

Considerando a equação (3.23), a matriz Hessiana projetada atualizada  $H_{p_{k+1}}$  é dada por

$$H_{p_{k+1}} = Z_{k+1}^T H_k Z_{k+1} = \begin{bmatrix} Z_k^T \\ \vdots \\ z^T \end{bmatrix} H_k \begin{bmatrix} Z_k \\ \vdots \\ z \end{bmatrix} = \begin{bmatrix} H_{p_k} & \vdots & Z_k^T H_k z \\ \vdots & \ddots & \vdots \\ z^T H_k Z_k & \vdots & z^T H_k z \end{bmatrix} \quad (3.39)$$

Pode-se ver pela equação, que a atualização da matriz Hessiana projetada simétrica  $H_{p_k}$  é equivalente a inclusão de uma linha e uma coluna adicionais a esta matriz.

A matriz Hessiana projetada atualizada  $H_{p_{k+1}}$  pode ser expressa, também, em termos da atualização da fatoração  $LDL^T$

$$H_{p_{k+1}} = L_{k+1} D_{k+1} L_{k+1}^T = \begin{bmatrix} L_k & \vdots & 0 \\ \vdots & \ddots & \vdots \\ u^T & \vdots & 1 \end{bmatrix} \begin{bmatrix} D_k & 0 \\ \vdots & \sigma \end{bmatrix} \begin{bmatrix} L_k^T & \vdots & u \\ \vdots & \ddots & \vdots \\ 0 & \vdots & 1 \end{bmatrix} \\ = \begin{bmatrix} H_{p_k} & \vdots & L_k^T D_k u \\ \vdots & \ddots & \vdots \\ u^T D_k L_k^T & \vdots & u^T D_k u + \sigma \end{bmatrix} \quad (3.40)$$

O vetor  $u$  e o escalar  $\sigma$  podem ser determinados pela comparação entre as equações (3.40) e (3.39), ou seja, pela solução das equações

$$L_k D_k \mu = Z_k^T H_k z \quad (3.41)$$

$$\sigma = z^T H z - u^T$$

Para uma matriz Hessiana positiva definida  $H_k$ , o elemento  $\sigma$  será sempre positivo.

Caso não se tenha a matriz Hessiana  $H_k$  ou a matriz Hessiana projetada (tem-se a fatoração  $L_k D_k L_k^T$ ), pode-se estimar a última linha ou coluna. Pode-se estimar esta última linha ou coluna pelo vetor  $[0, \dots, 0, 1]^T$  e obtém-se

$$L_{k+1} = \begin{bmatrix} L_k & \vdots & 0 \\ \dots & \dots & \dots \\ 0 & \vdots & 1 \end{bmatrix} \quad \text{e} \quad (3.42)$$

$$D_{k+1} = \begin{bmatrix} D_k & \vdots & 0 \\ \dots & \dots & \dots \\ 0 & \vdots & 1 \end{bmatrix} .$$

Este procedimento é algo arbitrário. Entretanto, nota-se que a direção de pesquisa horizontal  $\tilde{h}^k$  obtida pela queda de uma restrição é feita ao longo da nova coluna de  $Z_{k+1}$ . Se a queda de uma restrição é chamada pelo algoritmo, na próxima iteração, a última linha de  $L_{k+1}$  refletirá com maior precisão a informação da verdadeira matriz Hessiana projetada.

### 3.7.3 Atualização da Fatoração $LDL^T$ sem Mudança de Dimensão da Matriz $H_p$

Quando não ocorrer uma mudança da dimensão do subespaço  $S^E$ , ou seja,  $I_A^E(x^{k+1}) = I_A^E(x^k)$ , a atualização da fatoração  $LDL^T$  da matriz Hessiana projetada  $H_p$  inclui a nova informação associada a matriz  $Z_{k+1}$ . Esta estratégia é conseguida em duas etapas: na primeira etapa, atualiza-se a matriz Hessiana projetada (veja a seção 2.8) utilizando a matriz  $Z_{k+1}$  e, na segunda etapa, efetua-se a fatoração  $LDL^T$  da matriz Hessiana projetada atualizada. Entretanto, pode-se atualizar a fatoração  $LDL^T$  da matriz Hessiana projetada  $H_p$  fazendo a correção da fatoração  $LDL^T$  antiga com uma matriz de posto um aplicando as atualizações da seção 2.8.

### 3.8 Correção da Fatoração $LDL^T$ com uma Matriz de Posto Um

O algoritmo computacional (apresentado a seguir) para efetuar a correção da fatoração  $LDL^T$  aplicada na atualização da matriz Hessiana projetada, expressa por

$$L_{k+1} D_{k+1} L_{k+1}^T = L_k D_k L_k^T + \pi y y^T - \delta s s^T \quad (3.43)$$

é baseado no trabalho apresentado em Gill et alii (1975).



Notação

LA(N, N) - antiga matriz triangular inferior e unitária

LN(N, N) - nova matriz triangular inferior e unitária

DA(N) - antigo vetor dos elementos diagonais da matriz diagonal

DN(N) - novo vetor dos elementos diagonais da matriz diagonal

V(N) - vetor que define a correção da fatoração  $LDL^T$

VA(N) - vetor auxiliar

VN(N) - vetor auxiliar

a. Correção da fatoração  $LDL^T$  com uma matriz de posto um (+):

$$L_{k+1} D_{k+1} L_{k+1}^T \leftarrow L_k D_k L_k^T + vv^T$$

1. Faça:  $TA = 1$

2. Faça:  $VA(I) = DA(I)$  ,  $I = 1, \dots, N$

$$TN = TA + A*VA(I)$$



$$B = TN/TA$$

$$DN(I) = B*DA(I)$$

$$C = A/TN$$

se  $I = N$  , vá para 4

$$VN(K) = VA(K) - VA(I)*LA(K,I)$$

se  $B \leq 4$   $LN(K,I) = LA(K,I) + C*VN(K)$

se  $B > 4$   $LN(K,I) = LA(K,I)/B + C*VA(K)$

$$VA(K) = VN(K), \quad K = I+1, \dots, N$$

$$TA = TN$$

4. Faça:  $LN(I,I) = 1$  ,  $I = 1, \dots, N$

b. Correção da fatoração  $LDL^T$  com uma matriz de posto um(1):

$$L_{k+} D_{k+} L_{k+}^T \leftarrow L_k D_k L_k^T - vv^T$$

1. Faça:  $VN(1) = V(1)$

se  $N = 1$  , vá para 3

2. Calcule:  $VN(I) = V(I) - \sum_{K=1}^N LA(I,K)*VN(K)$  ,  $I = 2, \dots, N$

$$3. \text{ Calcule: } TA = 1 - \frac{\sum_{K=1}^N VN(K) * VN(K)}{DA(K)}$$

$$\text{se } TA \leq 0, \quad TA = 2.2 * 10^{-14}$$

$$4. \text{ Calcule: } j = N - I + 1$$

$$A = VN(J) / DA(J)$$

$$TN = TA + A * VN(J)$$

$$B = TA / TN$$

$$DN(J) = B * DA(J)$$

$$C = -A / TA$$

$$\text{se } J = N \quad \text{vã para 5}$$

$$LN(K, J) = LA(K, J) + C * VN(K)$$

$$VN(K) \leftarrow VN(K) + VN(J) * LA(K, J)$$

$$\left. \begin{array}{l} LN(K, J) = LA(K, J) + C * VN(K) \\ VN(K) \leftarrow VN(K) + VN(J) * LA(K, J) \end{array} \right\} K = J + 1, \dots, N$$

$$\left. \begin{array}{l} DN(J) = B * DA(J) \\ \left. \begin{array}{l} LN(K, J) = LA(K, J) + C * VN(K) \\ VN(K) \leftarrow VN(K) + VN(J) * LA(K, J) \end{array} \right\} K = J + 1, \dots, N \end{array} \right\} I = 1, N$$

$$5. \text{ Faça: } TA = TN$$

$$6. \text{ Faça: } LN(I, I) = 1, \quad I = 1, \dots, N$$

#### 4 SUMÁRIO DO ALGORITMO DA FUNÇÃO DE PENALIDADE EXATA GLOBALMENTE CONVERGENTE

A seguir será apresentado o sumário do algoritmo da função de penalidade exata globalmente convergente que se destina a minimizar uma função de penalidade exata com o parâmetro  $\mu > 0$  fixo.

##### Sumário do algoritmo

##### Início:

- escolha  $x^0, \epsilon^0, \delta, \lambda^0, \mu^0$
- escolha  $H_0$  (aproximação inicial da matriz Hessiana)
- faça  $k = 0$  (contador de iterações)

##### Processo iterativo:

- 1 - faça  $k \leftarrow k + 1$

- identifique  $I_A^\epsilon$  e  $I_V^\epsilon$  (conjunto de índices das restrições  $\epsilon$ -ativas e  $\epsilon$ -violadas)
- determine a matriz  $A_k$  (matriz dos gradientes das restrições ativas no ponto  $x^k$ )
- determine a fatoração  $Q^T \begin{bmatrix} R \\ \bar{0} \end{bmatrix}$  (ou a fatoração  $U\Sigma V^T$ ) da matriz  $A_k$
- identifique a matriz  $Z_k$  tal que  $A_k^T Z_k = 0$  e  $Z_k^T Z_k = I$
- calcule o pseudo-gradiente  $\nabla p_1(x^k)$
- calcule o pseudo-gradiente projetado  $Z_k^T \nabla p_1(x^k)$
- se  $\|Z_k^T \nabla p_1(x^k)\| \leq \underline{\lambda}$ , vá para 4

c.... computação da direção de pesquisa horizontal

$$h^k = -Z_k (Z_k^T H_k Z_k)^{-1} Z_k^T \nabla p_1(x^k)$$

- 2 - determine a fatoração  $L_k D_k L_k^T$  da matriz  $H_{p_k} = Z_k^T H_k Z_k$ 
  - resolva  $L_k \bar{w} = -Z_k^T \nabla p_1(x^k)$ , para  $\bar{w}$  por substituição progressiva
  - resolva  $L_k^T w = D_k^{-1} \bar{w}$ , para  $w$  por substituição regressiva

- faça  $h^k = Z_k w$

c.... pesquisa linear ao longo da direção de pesquisa horizontal  $h^k$

- se  $p(x^k + \alpha_k h^k)$  não tem um limite inferior finito para  $\alpha_k > 0$ , vá para 7

3 - encontre  $x^{k+1} = x^k + \alpha h^k$ , tal que  $p(x^{k+1}) < p(x^k)$  (pesquisa linear aproximada) ou  $p(x^{k+1}) = p(x^k + \alpha h^k) = \min_{\alpha} p(x^k + \alpha h^k)$  (pesquisa linear exata)

- atualize a matriz Hessiana projetada  $H_p$

- vá para 1

c.... computação dos multiplicadores de Lagrange (solução do sistema sobredeterminado de equações  $A\lambda = \nabla p_1(x)$ )

4 - resolva  $R_k \begin{bmatrix} -\lambda_1^k \\ \lambda_2^k \end{bmatrix} = Q_k \nabla p_1(x^k)$ , para o vetor  $\begin{bmatrix} -\lambda_1^k \\ \lambda_2^k \end{bmatrix}$  por substituição regressiva

- se  $\lambda_{1j}^k \notin \left[-\frac{1}{\mu}, \frac{1}{\mu}\right]$  para qualquer  $j \in I_{A_1}^E$  ou

$\lambda_{2j}^k \notin \left[0, \frac{1}{\mu}\right]$  para qualquer  $j \in I_{A_2}^E$ , vá para 5

c.... computação da direção de pesquisa horizontal

$$h^k = -Z_k (Z_k^T H_k Z_k)^{-1} Z_k^T \nabla p_1(x^k)$$



- determine a fatora $\tilde{c}$ o  $L_k D_k L_k^T$  da matriz  $H_{p_k} = Z_k^T H_k Z_k$
- resolva  $L_k \bar{w} = -Z_k^T \nabla p_1(x^k)$ , para o vetor  $\bar{w}$  por substitui $\tilde{c}$ o progressiva
- resolva  $L_k^T w = D_k^{-1} \bar{w}$ , para o vetor  $w$  por substitui $\tilde{c}$ o re $\tilde{c}$ gressiva
- fa $\tilde{c}$ a  $h^k = Z_k w$

c.... computa $\tilde{c}$ o da dire $\tilde{c}$ o de pesquisa vertical (solu $\tilde{c}$ o do sistema hipodeterminado de equa $\tilde{c}$ o $\tilde{c}$ es  $A_k^T v^k = \phi(x^k + h^k)$ , onde  $\phi(x^k + h^k)$   $\tilde{e}$  o vetor das fun $\tilde{c}$ o $\tilde{c}$ es das restri $\tilde{c}$ o $\tilde{c}$ es ati $\tilde{c}$ vas em  $x^k$ , mas avaliadas em  $x^k + h^k$ )

- resolva  $R^T \bar{v} = -\phi(x^k + h^k)$ , para o vetor  $\bar{v}$  por substitui $\tilde{c}$ o progressiva
- determine  $v^k = Q_1 \bar{v}$
- se  $p(x^k + h^k + v^k) \geq p(x^k + h^k) - \delta( \|Z_k^T \nabla p_1(x^k)\| + \sum_{i \in I_A^c} |\phi_i(x^k)| )$ , v $\tilde{a}$  para  $\delta$

c.... determina $\tilde{c}$ o do novo ponto de pesquisa (nas proximida $\tilde{c}$ es da solu $\tilde{c}$ o)

- fa $\tilde{c}$ a  $x^{k+1} \leftarrow x^k + h^k + v^k$

- faça o teste de convergência

- se  $\|Z^T \nabla p_1(x^{k+1})\|_2 \leq \text{tol}$ , pare

(ou, se  $\left[ \|\nabla p_1(x^{k+1})\| + \sum_{i \in I_{A1}} \lambda_{1i} \nabla \phi_i(x^{k+1}) - \right.$

$\left. - \sum_{i \in I_{A2}} \lambda_{2i} \nabla \phi_i(x^{k+1}) \right] \leq \text{tol}$ , pare)

- atualize a matriz Hessiana projetada  $H_p$

- vá para 4

c.... computação da direção de pesquisa horizontal  $\tilde{h}^k$  (queda de uma restrição)

5 - faça  $\sigma_j^k = -\text{sinal}(\lambda_j)$

- determine  $\tilde{h}^k = \sigma_j^k Z_{k \setminus j} Z_{k \setminus j}^T \nabla \phi_j(x^k)$

- se  $\nabla p_1(x^k) + \frac{1}{\mu} \min(0, \sigma_j^k) \nabla \phi_j(x^k) \cdot \tilde{h}^k < -\delta$ , vá para 3

6 - faça:  $\epsilon \leftarrow \epsilon/2$

$\Lambda \leftarrow \Lambda/2$

- identifique  $I_A^\epsilon$  e  $I_V^\epsilon$

- atualize a matriz  $A_k$

- determine a fatoração  $Q^T \begin{vmatrix} R \\ 0 \end{vmatrix}$  (ou a fatoração  $U \Sigma V$ ) da matriz  $A_k$

- identifique a matriz  $Z_k$
- calcule o pseudo-gradiente  $\nabla p_1(x^k)$
- calcule o pseudo-gradiente projetado  $Z_k^T \nabla p_1(x^k)$
- se  $\|Z_k^T \nabla p_1(x^k)\| = 0$ , vá para 6
- vá para 2

c.... redução do parâmetro  $\mu$

7 - faça  $\mu \leftarrow \mu/10$

- vá para 1

## 5 RESULTADOS NUMÉRICOS

O algoritmo computacional da função de penalidade exata globalmente convergente, apresentado no capítulo 4, foi desenvolvido e implementado em linguagem Fortran com precisão estendida no computador IBM-4341 da Universidade Federal da Paraíba.

O desempenho do algoritmo foi verificado utilizando quatro problemas de teste:

- o problema de teste de Rosen-Suzuki com 4 variáveis e 3 restrições de desigualdade, apresentado no apêndice 1;
- o problema de teste de Powell com 5 variáveis e 3 restrições de igualdade, apresentado no apêndice 2 e;
- o problema de teste de Wong com 10 variáveis e 8 restrições de desigualdade, apresentado no apêndice 3;
- o problema de fluxo de carga ótimo associado a um sistema de potência com 23 barras, 30 linhas e 15 geradores

que corresponde a um problema geral de programação matemática com 75 variáveis, 46 restrições de igualdade e 136 restrições de desigualdade, apresentado no apêndice 5.

Nos apêndices 1, 2 e 3 são apresentados, para cada problema de teste, a função objetiva, as restrições de igualdade e de desigualdade, o ponto inicial de pesquisa, o valor da função objetiva no ponto inicial de pesquisa, o ponto de solução e o valor da função objetiva no ponto de solução. No apêndice 4 são apresentados os símbolos, as relações básicas e os circuitos equivalentes para a descrição de um sistema de potência e para a definição de um problema de fluxo de carga ótimo. No apêndice 5 são apresentados os dados, a configuração de um sistema de potência com 23 barras, 30 linhas e 15 geradores e a formulação do problema de fluxo de carga ótimo associado a este sistema.

Os resultados numéricos para os problemas de testes são apresentados em tabelas que contenham os valores utilizados para os parâmetros  $\mu$ ,  $\epsilon$ ,  $\lambda$  e  $\delta$ , o número de iterações, o valor da função objetiva no ponto de solução, a norma do vetor das restrições ativas (CNORM) no ponto de solução e o número de passos de Newton utilizados em cada teste.

Assume-se, para todos os casos, a aproximação da matriz Hessiana inicial como sendo igual a matriz identidade.



**a) Resultados numéricos para o problema de teste de Rosen-Suzuki, apresentado no apêndice 1**

PARÂMETROS				NÚMERO DE ITERAÇÕES	FUNÇÃO OBJETIVA	CNORM	PASSOS DE NEWTON
1/μ	ε	λ	δ				
10	0.05	5	0.00001	12	-44	0.22E-11	6
5	0.01	5	0.00001	14	-44	0.57E-11	4
5	0.05	5	0.00001	13	-44	0.17E-12	4

Ponto de solução encontrado:

$$x^* = [-0.00001, 0.99997, 2.00002, -0.99998]^T$$

Valor das restrições de desigualdade no ponto de solução encontrado:

$$\phi_1 = 0.2420 \text{ E-13}$$

$$\phi_2 = 0.2178 \text{ E-11}$$

$$\phi_3 = 0.1000 \text{ E 01}$$

**b) Resultados numéricos para o problema de teste de Powell, apresentado no apêndice 2**

PARÂMETROS				NÚMERO DE ITERAÇÕES	FUNÇÃO OBJETIVA	CNORM	PASSOS DE NEWTON
1/μ	ε	λ	δ				
10	0.01	5	0.1	30	-2.9195	0.31E-09	03

Ponto de solução encontrado:

$$x^* = [-1.721725, 1.601006, 1.818720, -0.763157, -0.763030]^T$$

Valor das restrições de igualdade no ponto de solução encontrado:

$$\phi_1 = 0.6039 \text{ E-13}$$

$$\phi_2 = 0.1230 \text{ E-10}$$

$$\phi_3 = 0.3012 \text{ E-09}$$

**c) Resultados numéricos para o problema de teste de Wong, apresentado no apêndice 3**

PARÂMETROS				NÚMERO DE ITERAÇÕES	FUNÇÃO OBJETIVA	CNORM	PASSOS DE NEWTON
1/μ	ε	Λ	δ				
5	0.001	2	0.00001	24	24.3	0.83E-12	2
8	0.001	1	0.00001	32	24.3	0.47E-12	2
5	0.001	1	0.00001	22	24.3	0.22E-12	2
4	0.001	1	0.00001	22	24.3	0.25E-11	2

Ponto de solução encontrado:

$$x^* = [2.17199, 2.36368, 8.77392, 5.09598, 0.99065, 1.43057, 1.32164, 9.82872, 8.28009, 8.37592]^T$$

Valor das restrições de desigualdade no ponto de solução en  
contrado:

$$\phi_1 = 0.1776 \text{ E-13}$$

$$\phi_2 = 0.9592 \text{ E-13}$$

$$\phi_3 = 0.2131 \text{ E-13}$$

$$\phi_4 = 0.5115 \text{ E-12}$$

$$\phi_5 = 0.1847 \text{ E-12}$$

$$\phi_6 = 6.1485$$

$$\phi_7 = 0.1776 \text{ E-14}$$

$$\phi_8 = 50.0239$$

d) Resultados numéricos do problema de fluxo de carga ótimo  
associado a um sistema de potência com 23 barras, 30 li  
nhas e 15 geradores, apresentado no apêndice 5

PARÂMETROS				NÚMERO DE ITERAÇÕES	FUNÇÃO OBJETIVA	CNORM	PASSOS DE NEWTON
1/μ	ε	λ	δ				
300	0.001	10	0.00001	153	3843	0.01	0

Valores ótimos encontrados para as tensões nas barras e para  
as gerações de potência.

BARRA	TENSÃO NA BARRA		GEPADOR	POTÊNCIA GERADA	
	VALOR ABSOLUTO	ÂNGULO DE FASE		POT. ATIVA	POT. REATIVA
01	0.964468	0.000000	1	0.4500	0.3791
02	0.978687	0.040202	2	0.1500	0.1857
			3	1.2200	0.4161
03	0.960536	0.008314			
04	0.950152	-0.001778			
05	0.963921	0.044938			
06	0.951952	0.029477			
07	0.950012	0.013066			
08	0.980823	0.081948			
09	0.959625	0.027870			
10	1.001786	-0.027218			
11	1.095920	-0.011737			
			4	0.5750	0.4816
			5	0.5850	0.4695
			6	0.5850	0.4695
			7	0.5850	0.4695
12	0.970518	0.138389			
13	0.975888	0.172158			
14	1.005182	0.155012	8	5.0030	3.2540
15	0.953124	0.171668			
16	0.973075	0.307383			
17	0.963796	0.273885			
18	0.969315	0.263090			
19	0.972621	0.266304			
20	1.050156	0.403644	9	1.1200	0.9978
			10	3.3400	1.6150
			11	3.5700	1.6150
21	1.009727	0.335320			
22	0.998846	0.323843			
23	1.006427	0.394548	12	1.7110	0.6554
			13	1.1200	0.6554
			14	3.3400	0.6554
			15	3.5700	0.6554

Observação: - Este problema de teste foi executado uma única vez, portanto, é possível obter melhores resultados com menor número de iterações pela manipulação dos parâmetros  $\mu$ ,  $\epsilon$ ,  $\lambda$  e  $\delta$ .

- O problema físico em questão apresenta pequenas variações de iteração para iteração (quinta casa decimal) que está além da capacidade de controle

dos sistemas usuais. Portanto, o ponto encontrado pode ser considerado um ponto ótimo de solução para o sistema.



## 6 CONCLUSÕES

Neste trabalho, foi apresentado o algoritmo da função de penalidade exata globalmente convergente, destinado a resolver o problema geral de programação matemática, e os resultados dos testes feitos com quatro problemas de teste, incluindo o problema de fluxo de carga ótimo associado a um sistema de potência com 23 barras, 30 linhas e 15 geradores. Pelos resultados destes testes pode-se verificar que:

- em todos os casos, o algoritmo atingiu a solução com a precisão exigida, mostrando a propriedade de convergência global;
- o algoritmo apresentou a taxa de convergência superlinear com passo 2, quando próximo da solução;
- a convergência global, durante o processo iterativo, não depende da escolha dos parâmetros  $\mu$ ,  $\epsilon$ ,  $\lambda$  e  $\delta$ ;
- a eficiência do algoritmo depende da escolha dos parâmetros  $\mu$ ,  $\epsilon$ ,  $\lambda$  e  $\delta$ ;

- o número de passos de Newton,  $h^k + v^k$ , é pequeno em comparação com o número total de iterações requeridas para resolver os problemas de teste e, portanto, o uso do método das diferenças dos gradientes na atualização da matriz Hessiana projetada não é, computacionalmente, proibitivamente dispendioso;
- a eficiência do algoritmo é mantida com a utilização da matriz  $Z^T Z$  como aproximação da matriz Hessiana projetada, quando longe da solução;
- a dependência linear entre os gradientes das restrições ativas afetam o desempenho do algoritmo, mas não é claro como contornar este problema de maneira ótima;
- o algoritmo pode ser um instrumento útil na solução de problemas gerais de programação matemática, incluindo o problema de fluxo de carga ótimo, pois pode gerar resultados com eficiência e precisão requerida.

Pode-se fazer as seguintes sugestões para um futuro trabalho:

- incluir a dependência linear dos gradientes das restrições ativas aplicando, por exemplo, a teoria desenvolvida por Senad Bussovaca (1984) no algoritmo;

- otimizar as atualizações de  $\mu$ ,  $\epsilon$ ,  $\Lambda$  e  $\delta$ , durante o processo iterativo, para que a eficiência do algoritmo seja menos dependente da escolha inicial destes parâmetros;
- verificar se a normalização das variáveis podem influir no desempenho do algoritmo;
- verificar se outras técnicas de atualizações (fatoração ortogonal, fatoração  $LDL^T$ , matriz Hessiana ou matriz Hessiana projetada) podem melhorar o desempenho do algoritmo;
- incluir a informação de segunda ordem na direção de pesquisa que é utilizada quando ocorre a queda de uma restrição;
- verificar a dependência entre a eficiência do algoritmo e a pesquisa linear (fazer o uso de pesquisa linear exata ou aproximada ao longo da direção de pesquisa).

## BIBLIOGRAFIA

- | 1 | Armijo, L.  
"Minimization of Functions Having Lipschitz Continuous  
First Partial Derivatives"  
Pacific Journal of Mathematics, 1966, Vol. 16,  
Pages 1-3.
- | 2 | Asaadi, J.  
"A Computational Comparison of Some Non-Linear  
Programs"  
Math. Progr., 1973, Vol. 4, Pages 144-154.
- | 3 | Avriel, M.  
"Nonlinear Programming"  
Prentice Hall, Englewood Cliffs, New Jersey, 1976.
- | 4 | Ben, A.; Greville, T.  
"Generalized Inverses, Theory and Applications"  
John Wiley & Sons, New York, 1974.

- | 5 | Biggs, M.C.  
"Constrained Minimization Using Recursive Quadratic Programming",  
Towards Global Optimization,  
L.C.W Dixon and G.P.Szego, editors, North-Holland,  
Amsterdam, 1975.
- | 6 | Biggs, M.C.; Laughton, M.A.  
"Optimal Electric Power Scheduling: A Large Nonlinear  
Programming Test Problem Solved by Recursive Quadratic  
Programming"  
Math. Progr., 1977, Vol. 13, Pages 167-182
- | 7 | Busovaca, Senad  
"Handling Degeneracy in a Nonlinear  $\ell_1$  Algorithm"  
Ph.D. Thesis, University of Waterloo, Ontario, 1984.
- | 8 | Charalambous, C.  
"On the Conditions for Optimality of the Nonlinear  
Problems"  
Math. Progr., 1979.
- | 9 | Coleman, T.F.; Conn, A.R.  
"Nonlinear Programming Via an Exact Penalty Function:  
Asymptotic Analysis"  
Math. Progr., 1982a, Vol. 24, Pages 123-136.



- [10] Coleman, T.F.; Conn, A.R.  
"Nonlinear Programming Via an Exact Penalty Function:  
Global Analysis"  
Math. Progr., 1982b, Vol. 24, Pages 137-161.
- [11] Coleman, T.F.; Conn, A.R.  
"On the Local Convergence of a Quasi-Newton Method for  
the Nonlinear Programming Problem"  
SIAM J. NUMER. ANAL., August, 1984, Vol. 4,  
Pages 755-769.
- [12] Coleman, T.F.; Conn, A.R.  
"Second-Order Condition for an Exact Penalty Function  
and Applications"  
Technical Report CORR 78-27, Dept. of Combinatorics  
and Optimization, University of Waterloo, 1978.
- [13] Conn, A.R.  
"Projection Matrices - A Fundamental Concept in  
Optimization"  
Modeling and Simulation, Vol. 7, Part 1, edited by  
Voght and Mickle, 1976, Pages 599-605.
- [14] Conn, A.R.; Pietrzykowski, T.  
"A Penalty Function Method Converging Directly to a  
Constrained Optimum"  
SIAM J. NUMER. ANAL., 1977, Vol. 14.

- [15] Fiacco, A.V.; McCormick, G.P.  
"Nonlinear Programming"  
John Wiley & Sons, New York, 1968
- [16] Gill, P.; Murray, W.  
"Newton - Type Methods for Unconstrained and Linearly  
Constrained Optimization"  
Math. Progr., 1974, Vol. 7, Pages 311-350.
- [17] Gill, P.; Murray, W.  
"Numerical Methods for Constrained Optimization"  
Academic Press, London, 1974a.
- [18] Gill, P.; Murray, W.  
"Numerically Stable Methods for Quadratic Programming"  
Math. Progr., 1978, Vol. 14, Pages 349-372.
- [19] Gill, P.; Golub, G.H.; Murray, W.; Saunders, M.  
"Methods for Modifying Matrix Factorizations"  
Math. Comput., 1974b, Vol. 28, Pages 505-535.
- [20] Gill, P.; Murray, W.; Saunders, M.  
"Methods for Computing and Modifying the LDV Factors  
of a Matrix"  
Math. Comput., 1975, Vol. 29, Pages 1051-1077.

- [21] Han, S.P.  
"A Globally Convergent Method for Nonlinear Programming"  
JOTA, 1977a, Vol. 22, No 4.
- [22] Han, S.P.  
"Dual Variable Metric Methods for Constrained Optimization Problems"  
SIAM J. CONTROL and Opt, 1977b, Vol. 15, No 4.
- [23] Han, S.P.  
"Superlinealy Convergent Variable Metric Algoritms for General Nonlinear Programming Problems"  
Math. Progr., 1976, Vol. 11, Pages 263-282.
- [24] Hanson, R.J.; Lawson, C.L.  
"Solving Least Square Problems"  
Prentice Hall, Englewood Cliffs, N.J., 1974.
- [25] Hestenes, M.R.  
"Multiplier and Gradient Methods"  
JOTA, 1969, Vol. 4, Pages 303-320
- [26] Lipowski, J.S.  
"Modified Recursive Quadratic Programming Method and the Optimal Load Flow Problem"  
Technical Report No 80-4, University of Waterloo, 1980.

- [27] Lipowski, J.S.  
"Some new Computational Algorithms for the Solution of  
the Optimal Load Flow Problem"  
Ph.D. Thesis, University of Waterloo, Ontario, 1980.
- [28] Lipowski, J.S.  
"Solution of Optimal Load Flow Problem by Modified  
Recursive Quadratic - Programming Method"  
IEE Proc., 1981, Vol. 128, Pages 288-294.
- [29] Martin, R.S.; Peters, G.; Wilkinsons, J.H.  
"Symmetric Decomposition of a Positive Definite  
Matrix"  
Numerische Mathematik, 1965, Vol. 7, Pages 362-383.
- [30] McCormick, G.P.  
"Second Order Conditions for Constrained Minima"  
SIAM J. Appl. Math., 1967, Vol. 15, Pages 641-652.
- [31] Miele, A.; Cragg, E.; Levy, A.  
"Use of the Augmented Penalty Function in Mathematical  
Programming Problems"  
JOTA, 1971, Vol. 8, Pages 115-153.
- [32] Miele, A.; Moseley, P.; Levy, A.; Coggins, G.  
"On the Method of Multipliers for Mathematical  
Programming Problems"  
JOTA, 1972, Vol. 10, Pages 1-33.



- [33] Muklerjee, P.K.; Dhor, R.N.  
"Optimal Load Flow Solution by Reduced-Gradient Method"  
IEE Proc. - June, 1974, Vol. 121.
- [34] Murray, W.; Overton, M.  
"Steplength Algorithms of Nonlinear Equations in Several Variables"  
Academic Press, New York, 1970.
- [35] Ortega, J.M.; Rheinboldt, W.C.  
"Iterative Solution of Non-linear Equation in Several Variables"  
Academic Press, New York, 1970.
- [36] Osbourne, M.R.; Ryan, D.M.  
"A Hibrid Algorithm for Nonlinear Programming",  
Numerical Methods for Non-linear Optimization, F.A.,  
Lootsma, editors, Academic Press, London, 1972.
- [37] Pietrzykowski, T.  
"An Exact Potential Method for Constrained Maxima"  
SIAM J. NUMER. ANAL., 1969, Vol. 6, Pages 299-304.



- [38] Polyak, V.T., Tret'yakov  
"The Method of Penalty Estimates for Conditional  
Extremum Problems"  
Zh. Vyehil. Mat. mat. Fiz., 1973, Vol. 13, Pages 34-46.
- [39] Powell, M.J.D.  
"Algoritms for Nonlinear Constraints that Use  
Lagrangian Functions"  
Math. Progr., 1978, Vol. 14, Pages 224-248.
- [40] Powell, M.J.D.  
"A Fast Algorithm for Non-linearly Constrained  
Optimization Calculation"  
DAMPT, University of Cambridge, 1977, Report NAZ.
- [41] Powell, M.J.D.  
"A Method for Non-linear Constraints in Minimization  
Problems"  
Optimization, R. Fletcher, editor, Academic Press,  
London, 1969.
- [42] Rosen, J.B., Suzuki, S.  
"Construction of Nonlinear Programming Test Problems"  
Comm. ACM, 1965, Vol. 8.

- [43] Sasson, A.M.  
"Non-linear Programming Solution for Load-Flow,  
Minimum-Loss and Economic Dispatching Problems"  
IEEE PAS, Abril, 1969, Vol. 88.
- [44] Sasson, A.; Vilorio, F.; Aboytes, F.  
"Optimal Load Flow Solution Using the Hessian Matrix"  
IEEE PAS, Março/Abril, 1972.
- [45] Stewart, G.W.  
"Introduction to Matrix Computation"  
Academic Press, N.Y., 1973.
- [46] Tapia, R.  
"Newton's Method for Optimization Problems with  
Equality Constraints"  
SIAM J. NUMER. ANAL., 1974, Vol. 11, Pages 174-196.
- [47] Zangwill, W.I.  
"Non-linear Programming"  
Prentice Hall, N.J., 1967.

## APÊNDICE 1

PROBLEMA    TESTE DE ROSEN-SUZUKI [ROSEN & SUZUKI (1965)]  
COM 4 VARIÁVEIS E 3 RESTRIÇÕES DE DESIGUALDADE

Função objetivo:

$$f(x) = x_1^2 + x_2^2 + 2x_3^2 + x_4^2 - 5x_1 - 5x_2 - 21x_3 - 7x_4, \quad x \in \mathbb{R}^4.$$

Restrições de desigualdade:

$$\phi_1(x) = -2x_1^2 - x_2^2 - x_3^2 - 2x_1 + x_2 + x_4 + x_5 \geq 0,$$

$$\phi_2(x) = -x_1^2 - x_2^2 - x_3^2 - x_4^2 - x_1 - x_2 - x_3 + x_4 + 8 \geq 0,$$

$$\phi_3(x) = -x_1^2 - 2x_2^2 - x_3^2 - 2x_4^2 + x_1 + x_4 + 10 \geq 0.$$

Ponto de solução:

$$x^* = [0, 1, 2, -1]^T,$$

$$f(x^*) = -44.$$

Ponto inicial viável:

$$x^0 = [0, 0, 0, 0]^T,$$

$$f(x^0) = 0.$$

## APÊNDICE 2

### PROBLEMA TESTE DE POWELL [POWELL (1978a)] COM 5 VARIÁVEIS E 3 RESTRIÇÕES DE IGUALDADE

Função objetivo:

$$f(x) = \exp(x_1 x_2 x_3 x_4 x_5) \quad , \quad x \in \mathbb{R}^5$$

Restrições de igualdade:

$$\phi_1(x) = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 - 10 = 0 \quad ,$$

$$\phi_2(x) = x_2 x_3 - 5 x_4 x_5 = 0 \quad ,$$

$$\phi_3(x) = x_1^3 + x_2^3 + 1 = 0 \quad .$$

Ponto de solução:

$$x^* = [-1.717144, 1.595709, 1.827245, - .763643, - .763643]^T$$



$$f(x^*) = -2.9197004$$

Ponto inicial

$$x^0 = [-2, 2, 2, -1, -1]^T$$

$$f(x^0) = -8$$

### APÊNDICE 3

#### PROBLEMA TESTE DE WONG [ASAADI (1973)] COM 10 VARIÁVEIS E 8 RESTRIÇÕES DE DESIGUALDADE

Função objetivo:

$$\begin{aligned} f(x) = & x_1^2 + x_2^2 + x_1x_2 - 14x_1 - 16x_2 + (x_3 - 10)^2 + \\ & + 4(x_4 - 5)^2 + (x_5 - 3)^2 + 2(x_6 - 1)^2 + 5x_7^2 \\ & + 7(x_8 - 11)^2 + 2(x_9 - 10)^2 + (x_{10} - 7)^2 + 45, x \in \mathbb{R}^{10}. \end{aligned}$$

Restrições de desigualdade:

$$\phi_1(x) = -3(x_1 - 2)^2 - 4(x_2 - 3)^2 - 2x_3 + 7x_4 + 120 \geq 0, \quad ,$$

$$\phi_2(x) = -5x_1^2 - 8x_2 - (x_3 - 6)^2 + 2x_4 + 40 \geq 0, \quad ,$$

$$\phi_3(x) = -\frac{1}{2}(x_1 - 8)^2 - 2(x_2 - 4)^2 - 3x_5^2 + x_6 + 30 \geq 0, \quad ,$$

$$\phi_4(x) = -x_1^2 - 2(x_2 - 2) + 2x_1x_2 - 14x_5 + 6x_6 \geq 0, \quad ,$$

$$\phi_5(x) = -4x_1 - 5x_2 + 3x_7 - 9x_8 + 105 \geq 0 \quad ,$$

$$\phi_6(x) = -10x_1 + 8x_2 + 17x_7 - 2x_8 \geq 0 \quad ,$$

$$\phi_7(x) = 3x_1 - 6x_2 - 12(x_9 - 8)^2 + 7x_{10} \geq 0 \quad ,$$

$$\phi_8(x) = 8x_1 - 2x_2 - 5x_9 + 2x_{10} + 12 \geq 0 \quad .$$

Ponto de solução:

$$x^* = [2.17, 2.36, 8.77, 5.09, 0.99, 1.43, 1.32, 9.82, 8.27, 8.37]^T$$

$$f(x^*) = 24.3062 \quad .$$

Ponto inicial viável:

$$x^0 = [2, 3, 5, 5, 1, 2, 7, 3, 6, 10]^T$$

$$f(x^0) = 753$$

## APÊNDICE 4

### SÍMBOLOS, RELAÇÕES BÁSICAS E CIRCUITOS EQUIVALENTES RELEVANTES PARA A DESCRIÇÃO DE UM SISTEMA DE POTÊNCIA E PARA A DEFINIÇÃO DE UM PROBLEMA DE FLUXO DE CARGA ÓTIMO CORRESPONDENTE

#### 1 Símbolos e Relações Básicas

$n_B$	- número de barras;
$n_L$	- número de linhas;
$n_G$	- número de geradores;
$n_C$	- número de compensadores;
$n_T$	- número de transformadores;
$\hat{V}_i = V_{R_i} + jV_{I_i} = V_i e^{j\theta_i}$	- fasor da tensão de barra na barra $i$ ;
$V_{R_i}, V_{I_i}$	- componentes real e imaginário da tensão de barra na barra $i$ ;
$V_i, \theta$	- magnitude e ângulo de fase da tensão de barra na barra $i$ ;

$$V_i^{\min}, V_i^{\max}$$

- limites inferior e superior da magnitude da tensão de barra na barra  $i$ ;

$$\delta_k \equiv \delta_{ij} = \theta_i - \theta_j$$

- diferença do ângulo de fase das tensões de barra associado com a linha  $k$  que conecta a barra  $i$  à barra  $j$ ;

$$\delta_k^{\max} \equiv \delta_{i,j}^{\max}$$

- limite superior da diferença do ângulo de fase associado com a linha  $k$  que conecta a barra  $i$  à barra  $j$ ;

$$P_{G_i}, Q_{G_i}$$

- geração de potência ativa e reativa do gerador  $i$ ;

$$P_{G_i}^{\min}, P_{G_i}^{\max}, Q_{G_i}^{\min}, Q_{G_i}^{\max}$$

- limites inferior e superior das gerações de potência ativa e reativa do gerador  $i$ ;

$$S_{G_i} = \sqrt{P_{G_i}^2 + Q_{G_i}^2}$$

- geração de potência aparente do gerador  $i$ ;

$$S_{G_i}^{\text{nom}}$$

- valor nominal de potência aparente do gerador  $i$ ;

$$P_{L_i}, Q_{L_i}$$

- Potências ativa e reativas consumidas na barra  $i$ ;

$$P_{ij} = G_{ij}(V_{R_i} V_{R_j} + V_{I_i} V_{I_j}) +$$

$$+ B_{ij}(V_{R_i} V_{I_j} - V_{R_j} V_{I_i})$$

- Potência ativa que flui da barra  $i$  para a barra  $j$ ;



$$Q_{ij} = B_{ij}(V_{R_i} V_{R_j} + V_{I_i} V_{I_j}) - G_{ij}(V_{R_i} V_{I_j} - V_{R_j} V_{I_i})$$

- potência reativa que flui da barra i para a barra j;

$$P_i = \sum_{j \in L(i)} P_{ij}$$

- potência ativa injetada na barra i;

$$Q_i = \sum_{j \in L(i)} Q_{ij}$$

- potência reativa injetada na barra i;

$$P_{p_k} = P_{ij} + P_{ji}$$

- perda na transmissão de potência ativa na linha k que conecta a barra i à barra j;

$$Q_{p_k} = Q_{ij} + Q_{ji}$$

- perda na transmissão de potência reativa na linha k que conecta a barra i à barra j;

$$P_{p_k}^{\max}, Q_{p_k}^{\max}$$

- limite superior das perdas na transmissão de potências ativa e reativa;

$$P_p = \sum_{i=1}^{n_B} P_i = \sum_{i=1}^{n_L} P_{p_k}$$

- perda total da transmissão de potência ativa;

$$Q_p = \sum_{i=1}^{n_B} Q_i = \sum_{i=1}^{n_L} Q_{p_k}$$

- perda total da transmissão de potência reativa;

$$C(P_G) = \sum_{i=1}^{n_G} C_i(P_{G_i})$$

- custo total da geração de potência ativa;

$$C_i(P_{G_i}) = a_{1_i} P_{G_i} + a_{2_i} P_{G_i}^2 + a_{3_i} P_{G_i}^3$$

- custo da geração de potência ativa do gerador i;

$$a_{1i}, a_{2i}, a_{3i}$$

- coeficientes da função de custo de geração de potência ativa do gerador  $i$ ;

$$\hat{t}_i = t_{Ri} + jt_{Ii} = t_i e^{j\phi_i}$$

- razão de transformação complexa do transformador complexo  $i$ ;

$$t_{Ri}, t_{Ii}$$

- componentes real e imaginário da razão de transformação complexa do transformador complexo  $i$ ;

$$t_i, \phi_i$$

- magnitude e ângulo de fase da razão de transformação complexa do transformador complexo  $i$ ;

$$t_i^{\min}, t_i^{\max}, \phi_i^{\min}, \phi_i^{\max}$$

- limites inferior e superior da magnitude e do ângulo de fase da razão de transformação complexa do transformador complexo  $i$ ;

$$L(i)$$

- conjunto de índices das barras conectadas a barra  $i$ ;

$$\hat{Z}_{ij}^L = R_{ij}^L + jX_{ij}^L$$

- impedância complexa da linha  $k$  que conecta a barra  $i$  a barra  $j$ ;

$R_{ij}^L, X_{ij}^L$  - resistência e reatância de linha da linha k que conecta a barra i à barra j;

$\hat{Y}_{ij}^L = G_{ij}^L - jB_{ij}^L = \frac{1}{\hat{Z}_{ij}^L}$  - admitância complexa de linha da linha k que conecta a barra i à barra j;

$G_{ij}^L, B_{ij}^L$  - condutância e susceptância de linha da linha k que conecta a barra i à barra j;

$\hat{Y}_{ij}^{LSH} = G_{ij}^{LSH} - jB_{ij}^{LSH}$  - admitância shunt de linha complexa da linha k (que conecta a barra i à barra j) na barra i;

$G_{ij}^{LSH}, B_{ij}^{LSH}$  - condutância e susceptância shunt da linha k (que conecta a barra i à barra j) na barra i;

$G_{bus} = [G_{ij}]$  - matriz da condutância de barra;

$B_{bus} = [B_{ij}]$  - matriz da susceptância de barra;

$\hat{Y}_{bus} = G_{bus} - jB_{bus} = [Y_{ij}]$  - matriz admitância, tal que:

$$\hat{Y}_{ii} = \sum_{j \in L(i)} (\hat{Y}_{ij}^L + \hat{Y}_{ij}^{LSH}) \quad i = 1, \dots, n_B$$

$$\hat{Y}_{ij} = -\hat{Y}_{ij}^L \quad \begin{array}{l} i = 1, \dots, n_B \\ j = 1, \dots, n_B \\ i \neq j \end{array}$$

## 2 Circuito Equivalente para uma Linha de Transmissão com um Transformador Complexo

Uma linha de transmissão e um transformador complexo ideal conectado em série (Figura A4.1) pode ser representado por um circuito  $\pi$  equivalente (Figura A4.2).

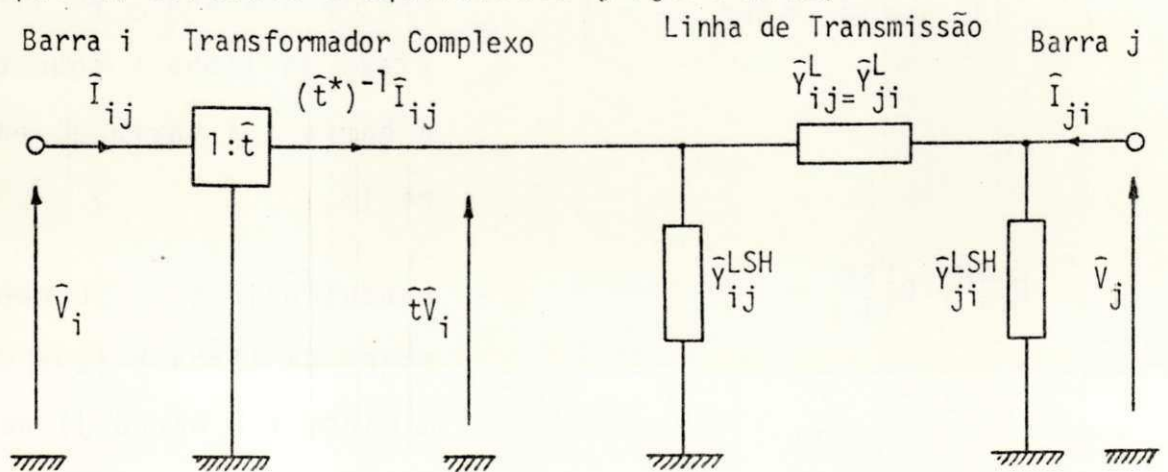


Fig. A4.1 - Uma linha de transmissão e um transformador complexo conectado em série.

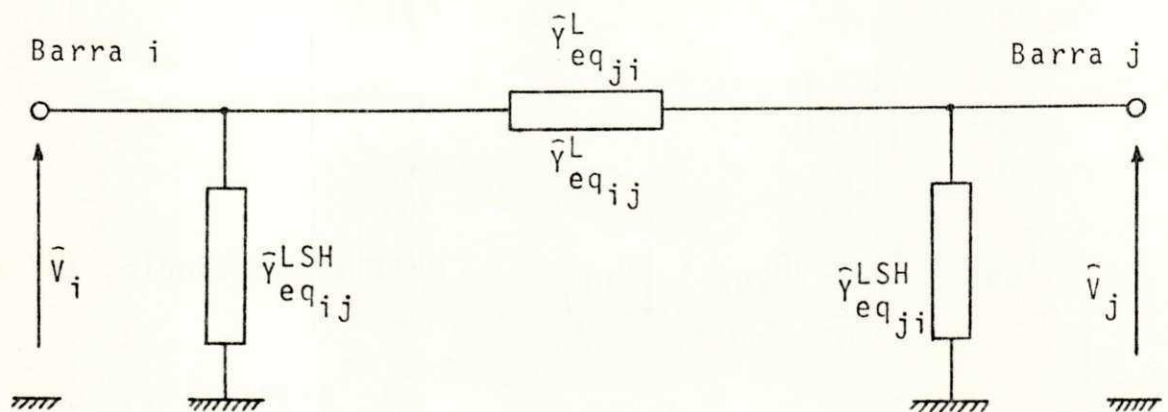


Fig. A4.2 - Um circuito  $\pi$  equivalente.



As admitâncias de um circuito  $\pi$  equivalente podem ser expressas como uma função dos parâmetros conhecidos da linha de transmissão e do transformador complexo:

$$\hat{Y}_{eq\,ij}^L = \hat{t}^* \hat{Y}_{ij}^L$$

$$\hat{Y}_{eq\,ji}^L = \hat{t}^* \hat{Y}_{ij}^L$$

(A4-1)

$$\hat{Y}_{eq\,ij}^{LSH} = \hat{t} \hat{t}^* (\hat{Y}_{ij}^L + \hat{Y}_{ij}^{LSH}) - \hat{t}^* \hat{Y}_{ij}^L$$

$$\hat{Y}_{eq\,ji}^{LSH} = \hat{Y}_{ij}^L + \hat{Y}_{ij}^{LSH} - \hat{t} \hat{Y}_{ij}^L$$

Das relações (A4-1)

$$G_{eq\,ij}^L = t_R G_{ij}^L - t_I B_{ij}^L$$

$$B_{eq\,ij}^L = t_R B_{ij}^L + t_I G_{ij}^L$$

$$G_{eq\,ji}^L = t_R G_{ij}^L + t_I B_{ij}^L$$

$$B_{eq\,ji}^L = t_R B_{ij}^L - t_I G_{ij}^L$$

$$G_{eq\,ij}^{LSH} = (t_R^2 + t_I^2)(G_{ij}^L + G_{ij}^{LSH}) - G_{eq\,ij}^L$$

$$B_{eq\,ij}^{LSH} = (t_R^2 + t_I^2)(B_{ij}^L + B_{ij}^{LSH}) - B_{eq\,ij}^L$$

$$G_{eq\,ji}^{LSH} = (G_{ij}^L + G_{ij}^{LSH}) - G_{eq\,ji}^L$$



$$B_{eqji}^{LSH} = (B_{ij}^L + B_{ij}^{LSH}) - B_{eqji}^L$$

## APÊNDICE 5

### SISTEMA DE POTÊNCIA COM 23 BARRAS, 30 LINHAS E 15 GERADORES E O RESPECTIVO PROBLEMA DE FLUXO DE CARGA ÓTIMO

Os dados para o sistema foram obtidos de Lipowski (1980a). Todos os valores são dados na forma 'por unidade'.

#### 1 Configuração do Sistema e os Dados de Linha

**TABELA A5.1 - CONFIGURAÇÃO DAS LINHAS E OS DADOS DE LINHA**

L I N H A	DESIGNAÇÃO DAS LINHAS		DADOS DE LINHA			
	BARRA -	BARRA	RESISTÊNCIA SÉRIE	REATÂNCIA SÉRIE	SUSCEPTÂNCIA SHUNT	POT. ATIVA LIM. CARGAS
1	1	- 3	0.02420	0.05400	0.00590	0.02420
2	1	- 4	0.03090	0.06930	0.00755	0.03090
3	2	- 5	0.04040	0.08880	0.00985	0.04040
4	5	- 8	0.03250	0.07090	0.00785	0.03250
5	2	- 7	0.06150	0.16200	0.01710	0.06150
6	3	- 6	0.05760	0.15200	0.01600	0.05760
7	4	- 9	0.02660	0.07000	0.00740	0.02660
8	7	- 9	0.02290	0.05040	0.00560	0.02290
9	6	- 8	0.04460	0.10030	0.01090	0.04460
10	10	- 11	0.02330	0.05140	0.02280	0.09320
11	8	- 10	0.05970	0.13150	0.01455	0.05970
12	9	- 10	0.05970	0.13150	0.01455	0.05970

L I N H A	DESIGNAÇÃO DAS LINHAS		DADOS DE LINHA			
	BARRA - BARRA		RESISTÊNCIA SÉRIE	REATÂNCIA SÉRIE	SUSCEPTÂNCIA SHUNT	POT. ATIVA LIM. CARGAS
13	13	- 14	0.00430	0.03510	0.11865	0.16530
14	12	- 14	0.00430	0.03510	0.11865	0.16530
15	12	- 15	0.00380	0.03070	0.10360	0.14610
16	15	- 18	0.00350	0.02880	0.09755	0.13450
17	13	- 23	0.00890	0.07200	0.24355	0.34210
18	16	- 17	0.00100	0.00800	0.02715	0.03840
19	17	- 18	0.00210	0.01670	0.05665	0.08070
20	18	- 19	0.00160	0.01270	0.04310	0.06150
21	19	- 20	0.00450	0.03620	0.12255	0.17300
22	18	- 22	0.00240	0.01920	0.06490	0.09230
23	20	- 21	0.00190	0.01560	0.05280	0.07300
24	21	- 22	0.00140	0.01140	0.03850	0.05380
25	16	- 23	0.00200	0.01640	0.05545	0.07690
26	8	- 12	0.00230	0.08390	0.00000	0.00550
27	8	- 13	0.00230	0.08390	0.00000	0.00550
28	9	- 12	0.00185	0.13000	0.00000	0.00600
29	9	- 13	0.00230	0.08390	0.00000	0.00550
30	1	- 2	0.00250	0.20000	0.00000	0.00200

Os valores atribuídos as susceptâncias shunt são iguais a metade dos valores totais de susceptâncias das linhas.

TABELA A5.2 - LOCALIZAÇÃO DOS GERADORES E DOS TRANSFORMADORES

GERADOR	BARRA	TRANSFORMADOR	LINHA	BARRA
1	1	1	26	12
2	2	2	27	13
3	2	3	28	12
4	11	4	29	13
5	11			
6	11			
7	11			
8	14			
9	20			
10	20			
11	20			
12	23			
13	23			
14	23			
15	23			

## 2 Potência Consumida nas Barras

TABELA A5.3 - POTÊNCIA CONSUMIDA NAS BARRAS

BARRA	CARGA	
	POTÊNCIA ATIVA	POTÊNCIA REATIVA
1	0.6400	0.1600
2	1.0100	0.2500
3	0.	0.
4	0.4700	0.1200
5	0.5100	0.1300
6	0.4100	0.1000
7	0.4800	0.1200
8	0.0100	0.
9	1.5000	0.3800
10	1.7700	0.4400
11	1.3000	0.3200
12	0.0600	0.
13	-0.0400	0.
14	4.8000	1.2000
15	2.0100	0.5000
16	1.3200	0.3300
17	3.4400	0.8600
18	1.0400	0.2600
19	3.7600	0.9400
20	-1.0000	-0.2500
21	3.7500	0.9400
22	-2.1000	-0.5200
23	1.2900	0.3200

## 3 Problema do Fluxo de Carga Ótimo

O problema associado do fluxo de carga ótimo é definido como um problema geral de programação matemática com:

- 75 variáveis;
- 46 restrições de igualdade;



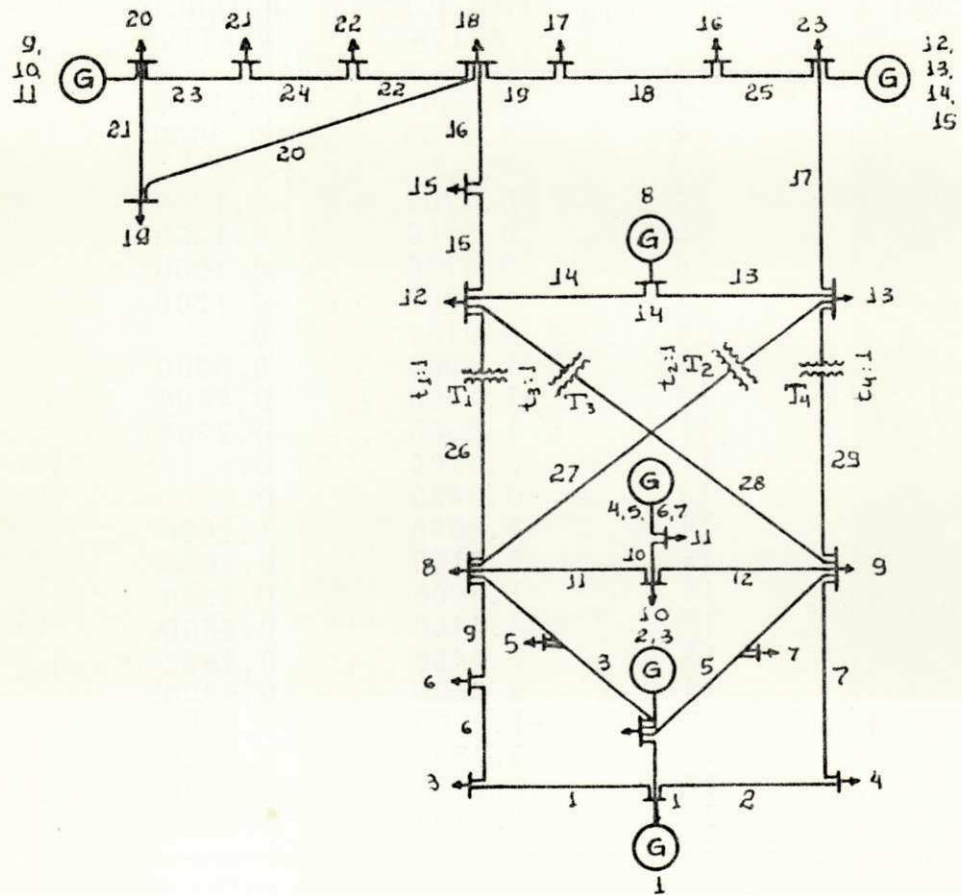


Fig. A5.1 - Configuração do Sistema.



- 136 restrições de desigualdade.

sob a condição imposta que as razões de transformação de todos os transformadores (reais) são fixadas:  $t_i = 1, i = 1, \dots, 4$

### 3.1 Função Objetivo - Custo Total da Geração de Potência Ativa

$$f(x) = C(P_G) = \sum_{i=1}^{15} C_i(P_{G_i})$$

onde

$$C_i(P_{G_i}) = a_{1_i} P_{G_i} + a_{2_i} P_{G_i}^2 + a_{3_i} P_{G_i}^3 - \text{função custo de geração de potência ativa associado ao } i\text{-ésimo gerador.}$$

**TABELA A5.4 - COEFICIENTES DA FUNÇÃO CUSTO DE GERAÇÃO DE POTÊNCIA ATIVA E OS LIMITES DAS GERAÇÕES DE POTÊNCIA ATIVA E POTÊNCIA REATIVA**

GERADOR	COEFICIENTES			LIMITES			
	$a_1$	$a_2$	$a_3$	$p_G^{\min}$	$p_G^{\max}$	$Q_G^{\min}$	$S_G^{\text{nom}}$
1	322.00	0.0	0.0	0.450	1.820	-0.300	2.400
2	322.00	0.0	0.0	0.150	0.610	-0.100	0.800
3	219.40	0.5	0.0	0.600	1.220	-0.200	1.600
4	221.37	-13.5	11.25	0.430	0.575	-0.100	0.750
5	218.14	1.0	0.0	0.430	0.585	-0.100	0.750
6	228.90	-29.0	23.50	0.430	0.585	-0.100	0.750
7	225.90	-29.0	23.50	0.430	0.585	-0.100	0.750
8	85.00	0.0	0.0	5.003	5.004	-1.200	6.600
9	170.56	1.0	0.0	0.220	1.120	-0.250	1.500
10	139.30	1.0	0.0	1.350	3.340	-0.650	4.000
11	118.14	1.0	0.0	1.430	3.570	-0.650	4.200
12	166.56	0.5	0.0	0.440	2.240	-0.500	3.000
13	170.56	1.0	0.0	0.220	1.120	-0.250	1.500
14	132.30	1.0	0.0	1.350	3.340	-0.650	4.000
15	112.16	1.0	0.0	1.430	3.570	-0.650	4.200

### 3.2 Variáveis (75)

$$x_i = V_{R_i} \quad i = 1, \dots, 23$$

$$x_{i+23} = V_{L_i} \quad i = 1, \dots, 22$$

$$x_{i+45} = P_{G_i} \quad i = 1, \dots, 15$$

$$x_{i+60} = Q_{G_i} \quad i = 1, \dots, 15$$

( $V_{I_{23}} = 0$  - componente de referência da tensão na barra.)

### 3.3 Restrições de Igualdade (46)

Equações nodais de fluxo de carga para barras não geradoras (34):

$$c_i(x) = P_k + P_{L_k} = 0 \quad k = f(i) \quad i = 1, \dots, 17$$

$$c_{i+17}(x) = Q_k + Q_{L_k} = 0 \quad k = f(i) \quad i = 1, \dots, 17$$

Equações que relatam a geração total e as gerações individuais nas barras geradoras (12):

$$c_{i+34}(x) = P_{G_i}^{tot} - \sum_{j \in J(i)} P_{G_j} = 0 \quad i = 1, \dots, 6$$

$$c_{i+40}(x) = Q_{G_i}^{tot} - \sum_{j \in J(i)} Q_{G_j} = 0 \quad i = 1, \dots, 6$$

onde

$$P_{G_i}^{tot} = P_{\ell} + P_{L_{\ell}} \quad \ell = f(i) \quad i = 1, \dots, 6$$

$$Q_{G_i}^{tot} = Q_{\ell} + Q_{L_{\ell}} \quad \ell = f(i) \quad i = 1, \dots, 6$$

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
k	3	4	5	6	7	8	9	10	12	13	15	16	17	18	19	21	22	⌘
ℓ	1	2	11	14	20	23												⌘⌘

⌘ Conjunto de Índices das barras não geradoras.

⌘⌘ Conjunto de Índices das barras geradoras.

$$J(1) = \{1\}$$

$$J(2) = \{2, 3\}$$

$$J(3) = \{4, 5, 6, 7\}$$

$$J(4) = \{8\}$$

$$J(5) = \{9, 10, 11\}$$

$$J(6) = \{12, 13, 14, 15\}$$

$$P_{L_i}, Q_{L_i} \quad i = 1, \dots, 23 \quad - \quad \text{potência ativa e potência reativa consumida das barras na Tabela A5.3.}$$

### 3.4 Restrições de Desigualdade (136)

Restrições correspondentes aos limites das tensões (46):

$$c_{i+46}(x) = \sqrt{V_{R_i}^2 + V_{L_i}^2} - v_i^{\min} \geq 0 \quad i = 1, \dots, 23$$

$$c_{i+69}(x) = v_i^{\max} - \sqrt{V_{R_i}^2 + V_{I_i}^2} \geq 0 \quad i = 1, \dots, 23$$

Restrições correspondentes aos limites de gerações

$$c_{i+92}(x) = P_{G_i} - p_{G_i}^{\min} \geq 0 \quad i = 1, \dots, 15$$

$$c_{i+107}(x) = p_{G_i}^{\max} - P_{G_i} \geq 0 \quad i = 1, \dots, 15$$

$$c_{i+122}(x) = Q_{G_i} - q_{G_i}^{\min} \geq 0 \quad i = 1, \dots, 15$$

$$c_{i+137}(x) = S_{G_i}^{\text{nom}} - \sqrt{P_{G_i}^2 + Q_{G_i}^2} \geq 0 \quad i = 1, \dots, 15$$

Restrições correspondentes aos limites de potência ativa transmitida (30):

$$c_{i+152}(x) = p_{P_i}^{\max} - P_{P_i} \geq 0 \quad , \quad i = 1, \dots, 30$$

onde

$$v_i^{\min} = 0.95 \quad i = 1, \dots, 23$$

$$v_i^{\max} = 1.05 \quad i = 1, \dots, 23$$

$$p_{G_i}^{\min}, p_{G_i}^{\max}, q_{G_i}^{\min}, S_{G_i}^{\text{nom}} \quad i = 1, \dots, 15 \quad - \text{limites dados na Tabela A5.4.}$$



$p_i^{\max}$

$i = 1, \dots, 30$  - limites dados na Tabela A5.1.

