

UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO DE CIÊNCIAS FÍSICAS E MATEMÁTICAS  
DEPARTAMENTO DE MATEMÁTICA

**INTRODUÇÃO AOS MÉTODOS VARIACIONAIS  
PARA SISTEMAS LINEARES**

**ALEXANDRA DUARTE CANDIDO**

**FLORIANÓPOLIS, DEZEMBRO DE 2004.**

ALEXANDRA DUARTE CANDIDO

# INTRODUÇÃO AOS MÉTODOS VARIACIONAIS PARA SISTEMAS LINEARES

Trabalho de Conclusão de Curso apresentado ao  
Curso de Matemática - Habilitação Licenciatura  
Departamento de Matemática  
Centro de Ciências Físicas e Matemáticas  
Universidade Federal de Santa Catarina

Orientador: Dr. Márcio Rodolfo Fernandes

FLORIANÓPOLIS - SC

Dezembro de 2004

Esta monografia foi julgada adequada como **TRABALHO DE CONCLUSÃO DE CURSO** no Curso de Matemática - Habilitação Licenciatura, e aprovada em sua forma final pela Banca Examinadora designada pela Portaria nº 71/SCG/04.

---

Prof<sup>a</sup> Carmem Suzane Comitre Gimenez  
Professora da disciplina

Banca Examinadora

---

Prof<sup>o</sup> Dr. Márcio Rodolfo Fernandes  
Professor Orientador

---

Prof<sup>o</sup> Dr. Daniel Norberto Kozakevich

---

Prof<sup>o</sup> Dr. Mário César Zambaldi

# *Agradecimentos*

A Deus Pai, por estar sempre guiando o meu caminho, acompanhando-me e confortando-me, em toda minha trajetória acadêmica.

A minha família que não mediu esforços para que eu alcançasse mais este objetivo.

Aos amigos que fiz durante toda minha graduação, e que de alguma maneira colaboraram para minha formação, em especial aos meus amigos Walisson Pereira Lorigiola e Rui Seara da Conceição Júnior.

Ao Felipe Luy Valério pela confecção das figuras deste trabalho.

Ao professor Márcio Rodolfo Fernandes, pela orientação na realização deste trabalho.

A todos os professores, porque somente eles são capazes de lapidarem com sabedoria as arestas que nos afligem.

# *Dedicatória*

Ofereço este trabalho para meus pais

Divaldo e Marcia,

para minhas irmãs

Fernanda e Joana,

pelo esforço e pela dedicação que sempre consagraram à minha formação.

# *Sumário*

<b>Introdução</b>	<b>7</b>
<b>1 O Problema de Programação Não-Linear</b>	<b>9</b>
<b>2 Condições de Otimalidade para Minimização sem Restrições</b>	<b>11</b>
2.1 Condições de Otimalidade . . . . .	11
<b>3 Convexidade</b>	<b>14</b>
3.1 Conceitos Fundamentais . . . . .	14
3.2 Funções Convexas Diferenciáveis . . . . .	15
<b>4 Modelo de Algoritmo com Buscas Direcionais</b>	<b>17</b>
4.1 Direções de Descida . . . . .	17
4.2 Modelo de Algoritmo . . . . .	18
<b>5 Métodos Clássicos de Descida</b>	<b>21</b>
5.1 Método do Gradiente . . . . .	21
5.2 Método de Newton . . . . .	22
<b>6 Gradientes Conjugados</b>	<b>26</b>
6.1 Método dos Gradientes Conjugados (GC) . . . . .	28
6.2 Pré-Condicionamento e Matrizes Não Simétricas . . . . .	29
<b>7 Exemplos Numéricos</b>	<b>31</b>
7.1 Um Problema de Difusão-Convecção . . . . .	31

7.2 Método de Newton para Minimização Irrestrita . . . . .	32
<b>Conclusão</b>	<b>34</b>
<b>Referências</b>	<b>35</b>

# *Introdução*

Os métodos para a resolução de sistemas lineares que requerem a fatoração da matriz dos coeficientes são chamados Métodos Diretos. Estes podem se tornar impraticáveis se a matriz for de grande porte e esparsa porque seus fatores geralmente serão matrizes cheias, isto é, sem estrutura de esparsidade. Uma exceção a este fato ocorre quando a matriz tem estrutura de banda. Ainda assim, os algoritmos de fatoração podem tornar-se de difícil implementação computacional.

Uma das razões para o grande interesse em sistemas lineares esparsos reside na resolução numérica de equações diferenciais. Sabe-se que as pesquisas nesta área tem sido responsáveis pelo desenvolvimento da maioria das estratégias para o tratamento de esparsidades.

Mais detalhadamente, existem duas formas de atacar um problema esparsa  $Ax = b$ . Uma delas é escolher um método direto e adaptar seu algoritmo de forma a explorar a esparsidade de  $A$ . Esta adaptação envolve o uso de estruturas de dados propícias ao seu problema e de estratégias especiais de pivotamento que minimizem o preenchimento da matriz.

Em contraste com os métodos diretos estão os métodos iterativos. Estes geram uma seqüência de soluções aproximadas através de operações que envolvem essencialmente produtos do tipo matriz-vetor. Desta forma, a estrutura de  $A$  não é alterada, isto é, não há o aparecimento de novos elementos na matriz tornando-se dispensável o uso de estrutura dinâmica de dados.

Nesta classe de métodos, o Métodos dos Gradientes Conjugados (GC), desenvolvido por Hestenes e Stifel, ver Golub (3), tem sido amplamente usado para resolver sistemas lineares esparsos e de grande porte, com  $A$  quadrada de ordem  $n$ , simétrica e definida-positiva. Tal método baseia-se na minimização de uma função quadrática e também pode ser visto como um método direto o que, na ausência de erros de arredondamento, obtém-se a solução do sistema em  $n$  iterações, ou como um método iterativo que, sob certas condições, fornece uma boa aproximação da solução em poucos passos.

Entretanto, se o sistema não possui simetria ou positividade, o método GC ainda

pode ser aplicado, através da formação (não explícita) das Equações Normais

$$A^t Ax = A^t b,$$

embora esta técnica possa retardar a convergência do método.

Ao longo deste trabalho, apresentamos uma coletânea de resultados sobre minimização de funções  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , para a construção do Método dos Gradientes Conjugados(GC). Além disso, testamos o algoritmo GC aplicado à resolução numérica de um problema estacionário de difusão-convecção.

# 1 *O Problema de Programação Não-Linear*

O problema de programação não-linear tem por objetivo solucionar problemas da forma

$$\begin{aligned} &\text{minimizar } f(x) \\ &\text{sujeita a } x \in S, \end{aligned} \tag{1.1}$$

onde  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  e  $S \subset \mathbb{R}^n$ .  $S$  é chamado conjunto factível e (1.1) é a forma genérica dos problemas de programação não-linear ou otimização.

Vamos considerar dois tipos de solução para este problema:

**Definição 1** *Um ponto  $x^* \in S$  é um minimizador local de  $f$  em  $S$  se e somente se existe  $\varepsilon > 0$  tal que  $f(x) \geq f(x^*)$  para todo  $x \in S$  tal que  $\|x - x^*\| < \varepsilon$ .*

*Se  $f(x) > f(x^*)$  para todo  $x \in S$  tal que  $x \neq x^*$  e  $\|x - x^*\| < \varepsilon$ , diremos que  $x^*$  é um minimizador local estrito em  $S$ .*

**Definição 2** *Um ponto  $x^* \in S$  é um minimizador global de  $f$  em  $S$  se existe  $f(x) \geq f(x^*)$  para todo  $x \in S$ . Se  $f(x) > f(x^*)$  para todo  $x \in S$  tal que  $x \neq x^*$ , diremos que se trata de um minimizador global estrito em  $S$ .*

Um resultado fundamental relacionado com problema de otimização é enunciado a seguir.

**Teorema 1** *(Bolzano - Weierstrass)*

*Se  $S$  é um conjunto fechado e limitado, e  $f : S \rightarrow \mathbb{R}$  é contínua, então existe  $x^* \in S$  minimizador global do problema (1.1).*

Prova: Consideremos primeiro a possibilidade de que  $f$  não seja limitada inferiormente

em  $S$ . Então, para cada  $k \in \mathbb{N}$ , existe  $x^k \in S$  tal que

$$f(x^k) \leq -k,$$

portanto,

$$\lim_{k \rightarrow \infty} f(x^k) = -\infty. \quad (1.2)$$

Como  $S$  é um conjunto fechado e limitado, existe  $K_1$  um subconjunto infinito de  $\mathbb{N}$  tal que a subsequência  $\{x^k\}$ ,  $k \in K_1$  converge a um ponto de  $S$ , digamos  $x^*$ . Pela continuidade de  $f$ , isto implica que

$$\lim_{k \in K_1} f(x^k) = f(x^*),$$

o que entra em contradição com (1.2).

Podemos aceitar, portanto, que  $f$  é limitada inferiormente em  $S$ . Seja

$$\gamma = \inf_{x \in S} f(x) > -\infty.$$

Pela definição de ínfimo, para todo  $k \in \mathbb{N}$ , existe  $x^k \in S$  tal que

$$\gamma \leq f(x^k) \leq \gamma + \frac{1}{k},$$

portanto

$$\lim_{k \rightarrow \infty} f(x^k) = \gamma.$$

Seja  $\{x^k\}_{k \in K_1}$  uma subsequência convergente de  $\{x^k\}$  e seja  $x^*$  seu limite. Então, pela continuidade de  $f$ ,

$$\gamma = \lim_{k \in K_1} f(x^k) = f(x^*).$$

Ou seja,  $f(x^*)$  assume o valor ínfimo do  $f$  no conjunto  $S$ . Isto implica que  $x^*$  é um minimizador global de (1.1).

## 2 Condições de Otimalidade para Minimização sem Restrições

Vamos analisar inicialmente o caso em que o conjunto factível é  $\mathbb{R}^n$ . Neste caso, (1.1) é chamado problema de minimização irrestrita.

### 2.1 Condições de Otimalidade

Suponhamos que os seguintes resultados para funções de uma variável, sejam conhecidos.

R1 - Seja  $f : \mathbb{R} \rightarrow \mathbb{R}, f \in C^1$ . Se  $x^*$  é um minimizador local de  $f$  em  $\mathbb{R}$ , então  $f'(x^*) = 0$ .

R2 - Seja  $f : \mathbb{R} \rightarrow \mathbb{R}, f \in C^2$ . Se  $x^*$  é um minimizador local de  $f$  em  $\mathbb{R}$ , então

$$(i) f'(x^*) = 0;$$

$$(ii) f''(x^*) \geq 0.$$

**Proposição 1** (*Condições Necessárias de Primeira Ordem*)

Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}, f \in C^1$ . Se  $x^*$  é um minimizador local de  $f$  em  $\mathbb{R}^n$ , então  $\nabla f(x^*) = 0$ .

Prova: Fixamos  $d \in \mathbb{R}^n$  arbitrário e consideremos a função  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  definida por

$$\phi(\lambda) = f(x^* + \lambda d).$$

Como  $x^*$  é um minimizador local de  $f$ , resulta que  $\lambda = 0$  é um minimizador local de  $\phi$ . Neste caso, por R1, concluímos que  $\phi'(\lambda) = 0$ .

Utilizando a regra da cadeia obtemos  $\phi'(\lambda) = \nabla^t f(x^* + \lambda d)d$ .

Substituindo para  $\lambda = 0$ , resulta  $0 = \phi'(\lambda) = \nabla^t f(x^*)d$ .

Como  $d \in \mathbb{R}^n$  é arbitrário, esta igualdade significa que  $\nabla f(x^*)$  é um vetor ortogonal a todos os vetores do espaço, portanto  $\nabla f(x^*) = 0$ .

**Proposição 2** (*Condições Necessárias de Segunda Ordem*)

Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f \in C^2$ . Se  $x^*$  é um minimizador local de  $f$  em  $\mathbb{R}$ , então

(i)  $\nabla f(x^*) = 0$ ;

(ii)  $\nabla^2 f(x^*)$  é semidefinida positiva.

Prova: A primeira parte desta tese vem da Proposição 1.

Para provar a segunda parte, consideremos  $\phi(\lambda)$ , como na Proposição 1.

$\mathbb{R}^2$  implica que  $\phi''(0) \geq 0$ . Usando a regra da cadeia temos  $\phi''(\lambda) = d^t \nabla^2 f(x^* + \lambda d)d$ , logo,  $\phi''(0) = d^t \nabla^2 f(x^*)d \geq 0$ .

Como  $d \in \mathbb{R}^n$  é arbitrário obtemos que  $\nabla^2 f(x^*)$  é semidefinida positiva.

**Proposição 3** (*Condições Suficientes de Segunda Ordem*)

Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f \in C^2$ . Se  $x^* \in \mathbb{R}^n$ ,  $\nabla f(x^*) = 0$ , e é positiva definida, então  $x^*$  é um minimizador local estrito de  $f$  em  $\mathbb{R}^n$ .

Prova: Seja  $\mathbb{B} = \{h \in \mathbb{R}^n / \|h\| = 1\}$ . Consideremos a função  $\Gamma : \mathbb{B} \rightarrow \mathbb{R}$  dada por

$$\Gamma(h) = h^t \nabla^2 f(x^*) h.$$

$\Gamma$  é uma função contínua e  $\mathbb{B}$  é um conjunto fechado e limitado, portanto  $\Gamma$  atinge um valor máximo e um valor mínimo em  $\mathbb{B}$ . Chamamos  $a$  ao valor mínimo, então

$$\Gamma(h) \geq a > 0 \text{ para todo } h \in \mathbb{B}.$$

Agora, consideremos  $d \in \mathbb{R}^n$ , arbitrário não nulo. Como  $\frac{d}{\|d\|} \in \mathbb{B}$ , temos que

$$d^t \nabla^2 f(x^*) d \geq a \|d\|^2. \tag{2.1}$$

Desenvolvendo  $f$  em série de Taylor em torno de  $x^*$ , temos

$$f(x^* + d) - f(x^*) = \nabla^t f(x^*) d + \frac{1}{2} d^t \nabla^2 f(x^*) d + o(\|d\|^2). \tag{2.2}$$

Desde que, por hipótese,  $\nabla f(x^*) = 0$ , (2.2) implica que

$$f(x^* + d) - f(x^*) \geq \frac{a}{2}\|d\|^2 + o(\|d\|^2).$$

Então, para todo  $d$  tal que  $\|d\|$  é suficientemente pequeno, o primeiro termo do membro direito da desigualdade define o sinal deste lado. Mas  $\frac{a}{2}\|d\|^2 > 0$ .

Portanto, para  $\|d\|$  suficientemente pequeno não nulo (digamos  $0 < \|d\| < \epsilon$ )

$$f(x^* + d) - f(x^*) > 0,$$

ou seja,  $f(x^* + d) > f(x^*)$ . Então, para todo  $x \in \mathbb{B}(x^*, \epsilon)$ ,  $x \neq x^*$ , temos que  $f(x) > f(x^*)$ .

Logo,  $x^*$  é um minimizador local estrito em  $f$ .

## 3 Convexidade

Para caracterizar minimizadores locais, serão úteis as proposições enunciadas no capítulo 2. Identificar se um minimizador local é também um minimizador global não é fácil, a menos que a função objetivo tenha características especiais. É o caso das funções convexas.

### 3.1 Conceitos Fundamentais

**Definição 3** Um subconjunto  $S \subset \mathbb{R}^n$  é convexo se e somente se para todo  $x, y \in S$ ,  $\lambda \in [0, 1]$  se verifica que  $\lambda x + (1 - \lambda)y \in S$ . Ver figura 1.

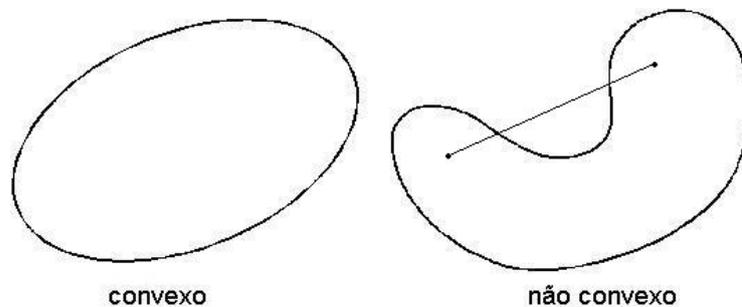


Figura 1: Ilustração da Definição 3

**Definição 4** Uma função  $f$  definida em um convexo  $S$  é convexa se e somente se para todo  $x, y \in S$ ,  $\lambda \in [0, 1]$  se verifica que  $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$ .

Se para todo  $\lambda \in (0, 1)$  e  $x \neq y$  vale que  $f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$ , diremos que  $f$  é estritamente convexa. Ver figura 2.

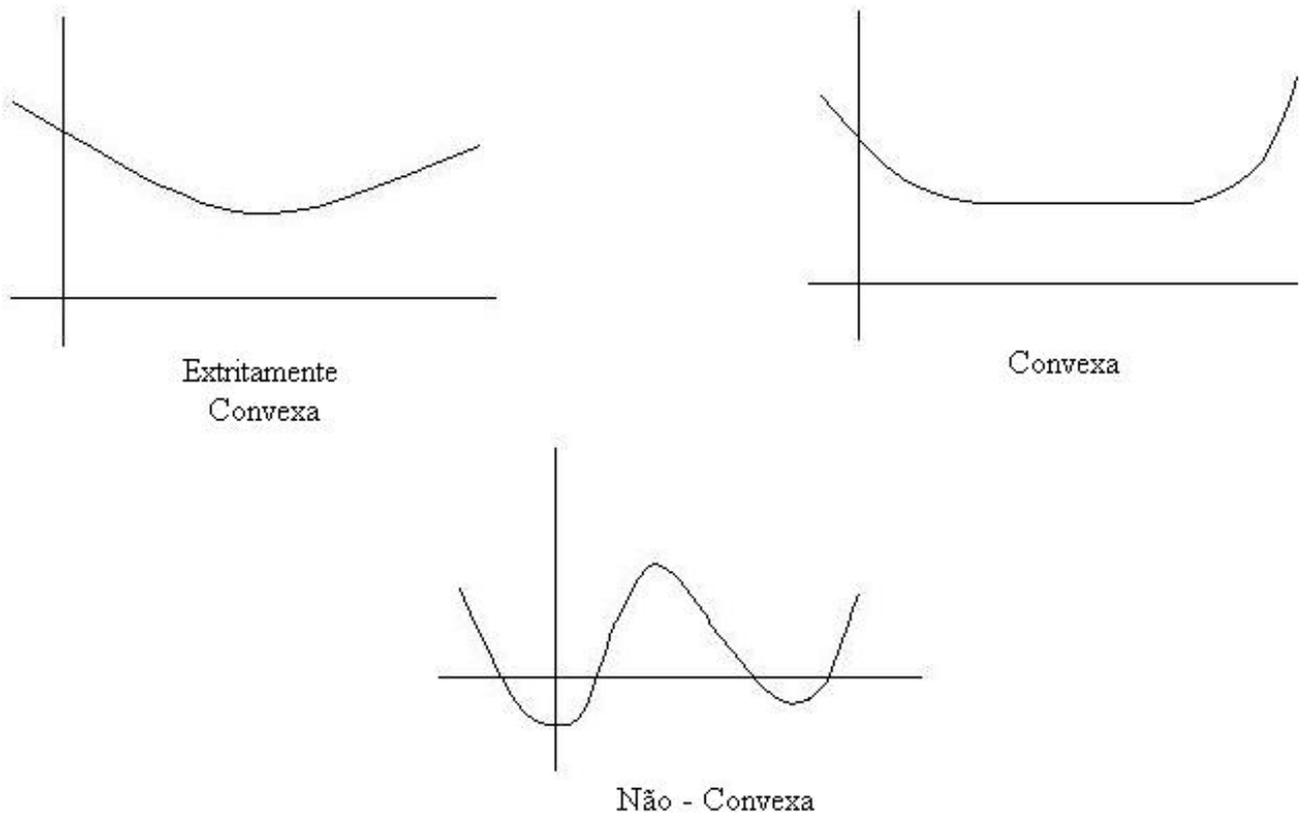


Figura 2: Ilustração da Definição 4

### 3.2 Funções Convexas Diferenciáveis

**Proposição 4** *Seja  $f \in C^1$ . Então,  $f$  é convexa em  $S$  convexo se e somente se para todo  $x, y \in S$  se verifica que*

$$f(y) \geq f(x) + \nabla^t f(x)(y - x).$$

**Proposição 5** *Seja  $f \in C^2$ . Seja  $S \subset \mathbb{R}^n$  convexo com interior não vazio. Então,  $f$  é convexa se e somente se  $\nabla^2 f(x) \geq 0$  para todo  $x \in S$ .*

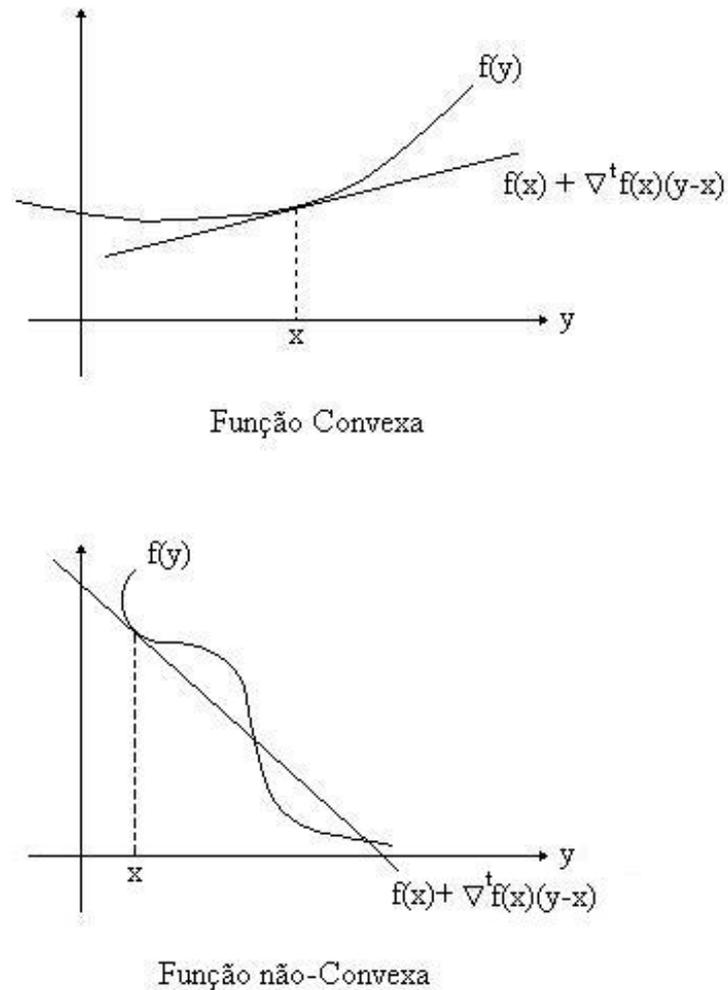


Figura 3: Ilustração da Proposição 4

**Proposição 6** *Seja  $f$  uma função convexa definida em  $S$  convexo. Então:*

- i) O conjunto  $\Gamma \subset S$  onde  $f$  toma seu valor mínimo é convexo;*
- ii) Qualquer minimizador local de  $f$  é um minimizador global de  $f$ .*

**Proposição 7** *Seja  $f \in C^1$  convexa definida em  $S$  convexo. Se existe  $x^* \in S$  tal que para todo  $y \in S$  se verifica que*

$$\nabla^t f(x^*)(y - x^*) \geq 0,$$

*então  $x^*$  é um minimizador global de  $f$  em  $S$ .*

As provas destas proposições podem ser encontradas em Luenberger (4).

## 4 Modelo de Algoritmo com Buscas Direcionais

### 4.1 Direções de Descida

Seja  $x \in \mathbb{R}^n$ , se  $\nabla f(x) \neq 0$ , temos, pela Proposição 1, que  $x$  não é um minimizador local de  $f$  em  $\mathbb{R}^n$ . Portanto, ao seu redor existe  $z \in \mathbb{R}^n$  tal que  $f(z) < f(x)$ .

Vamos caracterizar as direções a partir de  $x$ , nas quais é possível achar um ponto  $z \in \mathbb{R}^n$  que verifique  $f(z) < f(x)$ .

**Proposição 8** *Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f \in C^1$ ,  $x \in \mathbb{R}^n$  tal que  $\nabla f(x) \neq 0$ ,  $d \in \mathbb{R}^n$  tal que  $\nabla^t f(x)d < 0$ . Então existe  $\bar{\alpha} > 0$  tal que  $f(x + \alpha d) < f(x)$  para todo  $\alpha \in (0, \bar{\alpha}]$ .*

Prova: Vamos considerar a função  $\phi(\alpha) = f(x + \alpha d)$ . Então  $\phi(0) = f(x)$  e aplicando a regra da cadeia temos  $\phi'(0) = \nabla^t f(x)d$ .

Como  $\phi'(0) = \lim_{\alpha \rightarrow 0} \frac{\phi(\alpha) - \phi(0)}{\alpha}$ , então para  $0 < \alpha < \bar{\alpha}$ , com  $\bar{\alpha}$  suficientemente pequeno, o sinal de  $\phi'(0)$  e o sinal de  $\phi(\alpha) - \phi(0)$  devem ser o mesmo.

Como  $\nabla^t f(x)d < 0$  temos que  $\phi'(0) < 0$  e  $\phi(\alpha) - \phi(0) < 0$  para  $0 < \alpha < \bar{\alpha}$ , portanto  $f(x + \alpha d) < f(x)$ .

A Proposição 8 nos diz que, sendo  $d \in \mathbb{R}^n$  tal que  $\nabla^t f(x)d < 0$ , com certeza podemos encontrar nessa direção pontos cujo valor da função seja estritamente menor que  $f(x)$ .

As direções  $d \in \mathbb{R}^n$ , tais que  $\nabla^t f(x)d < 0$ , são chamadas direções de descida a partir de  $x$ .

A existência dessas direções indicam um modelo geral de algoritmo para minimizar uma função sem restrições.

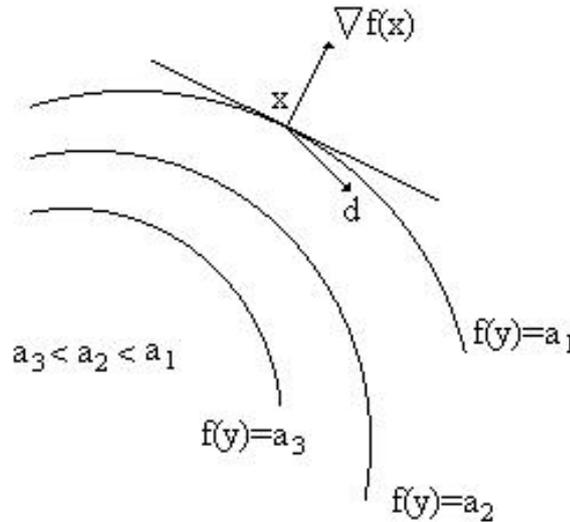


Figura 4: Ilustração da Proposição 8

## 4.2 Modelo de Algoritmo

Se  $x^*$  é uma solução de

$$\text{minimizar } f(x), x \in \mathbb{R}^n$$

e  $x^k$  uma aproximação de  $x^*$ , tal que  $\nabla f(x^k) \neq 0$ ; os passos para definir uma nova estimativa  $x^{k+1}$  são dados pelo algoritmo a seguir.

### Algoritmo 4.1

Passo 1: Escolher  $d_k \in \mathbb{R}^n$  tal que  $\nabla^t f(x^k)d_k < 0$ .

Passo 2: (Determinação do tamanho do passo)

$$\text{Calcular } \lambda_k > 0 \text{ tal que } f(x^k + \lambda_k d_k) < f(x^k).$$

(Este subproblema é chamado de busca linear.)

Passo 3: Fazer  $x^{k+1} = x^k + \lambda_k d_k$ .

Encerra-se o processo para algum valor de  $k$ , digamos  $k_0$ , tendo  $\nabla f(x^{k_0}) = 0$ . Neste caso  $x^{k_0}$  é um ponto estacionário e o Passo 1 não é mais possível. A condição  $\nabla f(x^k) = 0$  é necessária mas não é suficiente para deduzir que  $x^k$  é uma solução do problema. Na realidade, este processo nos indica “candidatos” à solução.

No entanto, é mais provável que o processo continue indefinidamente sem verificar

a condição  $\nabla f(x^k) = 0$  para nenhum valor de  $k$ . Neste caso, levando em consideração o algoritmo, estamos gerando uma seqüência infinita  $\{x^k\}$  de pontos em  $\mathbb{R}^n$ . Então as seguintes perguntas fazem sentido:

1- Existe  $\lim_{k \rightarrow \infty} x^k$ ?

2- Se  $\lim_{k \rightarrow \infty} x^k = x^*$ , é possível garantir alguma das afirmações seguintes?

$x^*$  é uma solução do problema;

$x^*$  é um ponto estacionário.

Vamos dar uns passos no sentido de responder essas perguntas. Claramente, o Algoritmo 4.1 gera uma seqüência de pontos  $\{x^k\}$  tal que a seqüência de números reais associada  $\{f(x^k)\}$  é monótona decrescente.

Para exemplificar, consideremos a função de uma variável  $f(x) = x^2$ . O único minimizador desta função é  $x^* = 0$ . A seqüência definida por

$$x^k = 1 + \frac{1}{k}, \text{ para } k \geq 1$$

pode ser gerada pelo algoritmo porque

$$f(x^{k+1}) = \left(1 + \frac{1}{k+1}\right)^2 < \left(1 + \frac{1}{k}\right)^2 = f(x^k).$$

Porém,

$$\lim_{k \rightarrow \infty} x^k = 1.$$

Este exemplo nos responde a pergunta (2), é negativa.

No exemplo, temos a impressão de que, apesar de haver sempre um decréscimo da função, este decréscimo é pequeno demais devido a distância entre  $x^{k+1}$  e  $x^k$  que se aproxima de zero muito rapidamente.

O decréscimo pode ser muito pequeno mesmo sendo grande a distância entre  $x^{k+1}$  e  $x^k$ , como vemos na figura 5.

Nessa figura,  $f(y) = f(x^k)$  e tomando  $x^{k+1}$  arbitrariamente próximo de  $y$  teremos  $f(x^{k+1}) < f(x^k)$ . Porém a diferença entre estes valores será arbitrariamente pequena.

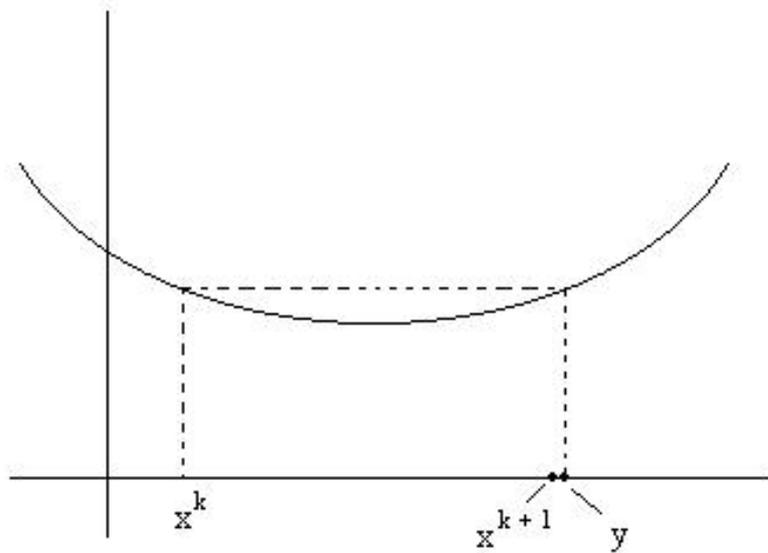


Figura 5: Decréscimo pequeno para distância grande

Há uma terceira situação que pode levar-nos a obter decréscimos excessivamente pequenos do valor da função. Com efeito, consideremos o conjunto de nível que passa por  $x^k$ :

$$\Gamma = \{x / f(x) = f(x^k)\}.$$

Se nos limitássemos a andar sobre  $\Gamma$ , o decréscimo da função seria nulo. Assim, se a direção  $d_k$  é “quase” perpendicular a  $\nabla f(x^k)$ , essa direção é “quase” tangente a  $\Gamma$  em  $x^k$ . Neste caso também podemos ter pouco decréscimo do valor da função na direção  $d_k$ . Na figura 6 vemos tal situação.

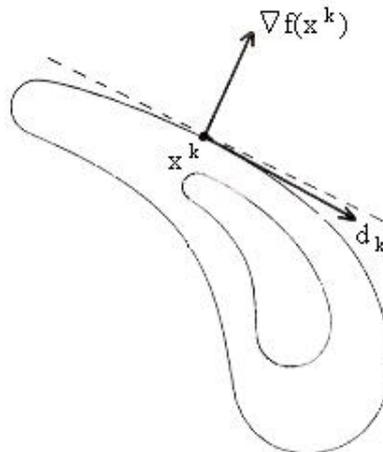


Figura 6: Decréscimo pequeno para direção “quase” tangente

## 5 Métodos Clássicos de Descida

### 5.1 Método do Gradiente

Dentro do contexto do Algoritmo 4.1, este método corresponde a escolher  $d_k$  na direção de  $-\nabla f(x^k)$ .

Consideremos o seguinte algoritmo para minimizar uma função  $f$  definida em  $\mathbb{R}^n$ .

#### Algoritmo 5.1

Se  $x^k \in \mathbb{R}^n$  é tal que  $\nabla f(x^k) \neq 0$  os passos para determinar  $x^{k+1}$  são:

Passo 1: Calcular  $d_k = -\nabla f(x^k)$

Passo 2: (Busca linear exata)

Determinar  $\lambda_k$ , minimizador de  $f(x^k + \lambda d_k)$  sujeita a  $\lambda \geq 0$ .

Passo 3: Fazer  $x^{k+1} = x^k + \lambda_k d_k$ .

Podemos observar que no Passo 2 a busca linear é mais exigente que a do algoritmo 4.1, porque  $\lambda_k$  é o minimizador de  $f$  na direção de  $d_k$ . Chamamos a este processo de busca linear exata. É importante perceber que este subproblema pode não ter solução e portanto o Algoritmo 5.1 nem sempre está bem definido.

Se a função objetivo for quadrática, isto é, se  $f(x) = \frac{1}{2}x^t G x + b^t x + c$ , com  $G$  definida positiva, ou seja,  $G^t = G$  e  $x^t G x > 0$ , para qualquer  $x$  não-nulo, então existe um único  $x^* \in \mathbb{R}^n$  que é minimizador global de  $f$ . Ver figura 7.

Neste caso, a busca linear exata determina

$$\lambda_k = \nabla^t f(x^k) \nabla f(x^k) / \nabla^t f(x^k) G \nabla f(x^k).$$

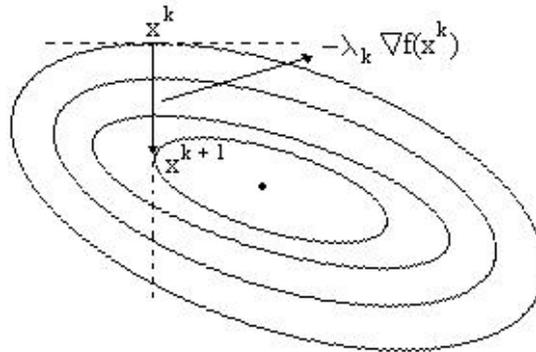


Figura 7: Minimizador global de  $f$

O teorema a seguir garante a convergência da seqüência gerada pelo Algoritmo 5.1, para qualquer aproximação inicial e que a ordem de convergência da seqüência associada  $\{f(x^k)\}$  é linear.

**Teorema 2** *Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função quadrática com matriz hessiana  $G$  definida positiva.*

*Seja  $x^*$  o minimizador global de  $f$ .*

*Dado  $x^0 \in \mathbb{R}^n$ , arbitrário, o Algoritmo 5.1 gera uma seqüência  $\{x^k\}$  tal que*

*(i)  $\lim_{k \rightarrow \infty} x^k = x^*$*

*(ii)  $\lim_{k \rightarrow \infty} f(x^k) = f(x^*)$*

*e*

$$f(x^{k+1}) - f(x^*) \leq ((A - a)/(A + a))^2 (f(x^k) - f(x^*)),$$

*onde  $A$  e  $a$  são o maior e o menor autovalor de  $G$ , respectivamente.*

Prova: Ver Luenberger (4).

## 5.2 Método de Newton

**Proposição 9** *Se  $f$  é uma função quadrática com hessiana  $G$  definida positiva, dado  $x^0 \in \mathbb{R}^n$  arbitrário, a direção  $d \in \mathbb{R}^n$  dada por:*

$$d = -G^{-1}(Gx^0 + b) \tag{5.1}$$

verifica que

$$x^* \equiv x^0 + d \quad (5.2)$$

é o minimizador global de  $f$  em  $\mathbb{R}^n$ . Ver Figura 5.2.

Prova: Seja  $f(x) = \frac{1}{2}x^t Gx + b^t x + c$ . Temos, por (5.2), que  $\nabla f(x^*) = G(x^0 + d) + b$ . Logo, usando (5.2), obtemos que

$$\nabla f(x^*) = G(x^0 - G^{-1}(Gx^0 + b)) + b.$$

Então,  $\nabla f(x^*) = Gx^0 - Gx^0 - b + b = 0$ , o que prova a proposição.

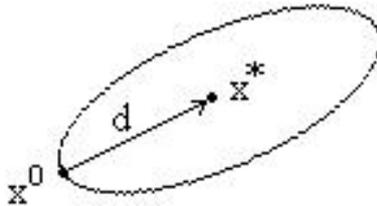


Figura 8: Ilustração da Proposição 9

A direção  $d$  é a solução do sistema linear

$$Gd = -(Gx^0 + b) = -\nabla f(x^0).$$

Assim, temos que, minimizar uma função quadrática com hessiana definida positiva é um problema equivalente a resolver um sistema linear com matriz simétrica e definida positiva.

Se a função não for quadrática e tivermos uma aproximação  $x^k$  da solução de

$$\text{Minimizar } f(x), x \in \mathbb{R}^n,$$

podemos usar o resultado anterior na função quadrática que resulta da consideração dos três primeiros termos do desenvolvimento em série de Taylor de  $f$  em torno de  $x^k$ :

$$q(d) = f(x^k) + \nabla^t f(x^k)d + \frac{1}{2}d^t \nabla^2 f(x^k)d.$$

Chamamos  $c = q(0) = f(x^k)$ ,  $b = \nabla q(0) = \nabla f(x^k)$ ,  $G = \nabla^2 q(0) = \nabla^2 f(x^k)$ .

Se escrevermos  $q(d) = \frac{1}{2}d^t G d + b^t d + c$  e se  $\nabla^2 f(x^k)$  é definida positiva podemos calcular o minimizador global desta quadrática a partir de  $d_0 = 0$ . Assim, temos que

$$d^* = -G^{-1}(Gd_0 + b) = -G^{-1}b = -(\nabla^2 f(x^k))^{-1}\nabla f(x^k).$$

Isto sugere a escolha  $d_k = -(\nabla^2 f(x^k))^{-1}\nabla f(x^k)$  no Passo 1 do Algoritmo 4.1.

A pergunta a seguir é pertinente:

$d_k$  é sempre uma direção de descida?

Se  $\nabla^2 f(x^k)$  não for definida positiva,  $d_k$  não pode ser uma direção de descida.

Exemplo: a função  $f(x, y) = (\frac{1}{2})(x^2 - y^2)$  no ponto  $x^0 = (0, 1)^t$  verifica que:

$$\nabla f(x^0) = (0, -1)^t, \text{ e } \nabla^2 f(x^0) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Neste caso, a direção de Newton é

$$d_0 = (0, -1)^t,$$

e

$$\nabla^t f(x^0)d_0 = 1 > 0.$$

Mesmo tendo  $d_0$  uma direção de subida, pode-se dizer que basta tomar  $\tilde{d} = -d_0$  para obter uma direção de descida. Mas o seguinte exemplo devido a Powell mostra que a situação pode não ter conserto:

A função  $f(x, y) = x^4 + xy + (1 + y)^2$  em  $x^0 = (0, 0)^t$  verifica

$$\nabla f(x^0) = (0, 2)^t \text{ e } \nabla^2 f(x^0) = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}.$$

A solução de  $\nabla^2 f(x^0)d = -(0, 2)^t$  é  $d_0 = -(0, 2)^t$  e  $\nabla^t f(x^0)d_0 = 0$ .

Vamos considerar o algoritmo a seguir:

**Algoritmo 5.2 (Método de Newton)**

Se  $x^k$  é tal que  $\nabla f(x^k) \neq 0$ , os passos para determinar  $x^{k+1}$  são:

Passo 1: Determinar  $d_k$  tal que

$$\nabla^2 f(x^k)d_k = -\nabla f(x^k),$$

(ou seja, resolver este sistema linear. Notemos que este passo pode não estar bem-definido se  $\nabla^2 f(x^k)$  for singular.)

Passo 2: Fazer  $x^{k+1} = x^k + \lambda_k d_k$ , onde  $\lambda_k$  é determinado como no Passo 2 do Algoritmo 4.1.

Para o algoritmo 5.2 temos o seguinte resultado:

**Proposição 10** *Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f \in C^3$ . Seja  $x^*$  um minimizador local de  $f$  em  $\mathbb{R}^n$ , tal que  $\nabla^2 f(x^*)$  é definida positiva. Então, existe  $\epsilon > 0$  tal que se  $x^0 \in \mathbb{B}(x^*, \epsilon)$ , e  $\lambda_k = 1$  para todo  $k \in \mathbb{N}$ , a seqüência  $\{x^k\}$  gerada pelo Algoritmo 5.2 verifica:*

- (i)  $\nabla^2 f(x^k)$  é definida positiva para todo  $k \in \mathbb{N}$ ;
- (ii)  $\lim_{k \rightarrow \infty} x^k = x^*$ ;
- (iii) Existe  $c > 0$  tal que  $\|x^{k+1} - x^*\| \leq c\|x^k - x^*\|^2$  para todo  $k \in \mathbb{N}$ .

Prova: Ver Luenberger (4) .

Este é um resultado de convergência local que diz que se escolhermos  $x^0$  suficientemente perto de  $x^*$ ,

- (i) os sistemas lineares do Passo 1 têm solução única e portanto  $d_k$  está bem-definido para todo  $k \in \mathbb{N}$ ;
- (ii) a seqüência converge a  $x^*$ .

## 6 *Gradientes Conjugados*

Admitiremos que a matriz do sistema linear seja simétrica e positiva definida, ou seja,  $A^t = A$  e  $x^t Ax > 0$ , para qualquer  $x$  não-nulo.

Vimos no capítulo anterior a idéia básica dos métodos do tipo gradiente, para resolver o sistema linear  $Ax = b$ , que é minimizar a função de  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  do tipo

$$f(x) = \frac{1}{2}x^t Ax - b^t x. \quad (6.1)$$

Embora o desenvolvimento seja o mesmo para um número arbitrário de componentes, explicitaremos as equações com  $n = 2$ , para facilitar o entendimento. Utilizando a simetria de  $A$ , (6.1) pode ser reescrita como

$$f(x) = \frac{1}{2}(a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2) - b_1x_1 - b_2x_2.$$

As curvas de nível da função  $f$  são elipses. Assim, esta função tem um único minimizador global, pois o seu gráfico é um parabolóide.

O mínimo da função é atingido no ponto  $x$  que anula o gradiente da função,

$$\nabla f(x_1, x_2) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 - b_1 \\ a_{21}x_1 + a_{22}x_2 - b_2 \end{bmatrix} = Ax - b. \quad (6.2)$$

Se,  $\nabla f(x) = 0$ , então  $Ax = b$ , isto é, a solução do sistema de equações lineares minimiza a função quadrática e vice-versa.

Do Cálculo Diferencial, sabemos que, o vetor gradiente aponta para a direção de crescimento máximo da função. Logo, é natural que, nos passos de busca do mínimo, caminhemos na direção contrária ao gradiente, isto é,

$$x^{k+1} = x^k - s_k \nabla f(x^k). \quad (6.3)$$

A aproximação num passo  $k + 1$  é calculada a partir da aproximação no passo anterior,

caminhando na direção oposta ao gradiente. O parâmetro real  $s_k$  regula o tamanho do passo na  $k$ -ésima iteração.

Usando (6.2), temos que a direção de descida é  $-\nabla f(x^k) = b - Ax^k = r^k$ , que é o resíduo associado à aproximação  $x^k$ .

O tamanho do passo  $s_k$  em (6.3), será usado na minimização do resíduo associado à aproximação que está sendo calculada. Sendo assim, vamos calcular o valor de  $s$  que minimiza

$$f(x + sr) = \frac{1}{2}(x + sr)^t A(x + sr) - b^t(x + sr). \quad (6.5)$$

Fixamos  $x$  e derivamos a função em relação à variável  $s$ , aplicando as regras da cadeia e de derivação de produto:

$$\frac{df}{ds}(x + sr) = \frac{1}{2}r^t A(x + sr) + \frac{1}{2}(x + sr)^t Ar - b^t r.$$

Usando a propriedade distributiva nas operações e as propriedades de transposição de vetores e matrizes para mostrar que  $x^t Ar = (Ar)^t x = r^t Ax$ , uma vez que  $A$  é simétrica, temos

$$\frac{df}{ds}(x + sr) = sr^t Ar + r^t(Ax - b) = sr^t Ar - r^t r.$$

Por fim, igualando a derivada a zero, obtemos o valor de  $s$  que minimiza a função:

$$s = \frac{r^t r}{r^t Ar}.$$

Com este valor de  $s$  em (6.4) e lembrando que  $-\nabla f(x) = r^k$ , o Método do Gradiente pode ser resumido no Algoritmo 6.1:

### Algoritmo 6.1 Método do Gradiente

Dados  $A, b, max, tol$

- 1)  $x^0 = 0$
- 2) Para  $k = 0 : max$ , faça
- 3)  $r = b - Ax$
- 4)  $s = \frac{r^t r}{r^t Ar}$
- 5)  $x^{k+1} = x^k + sr$
- 6) Se  $r^t r < tol$  então
- 7) Saída com  $sol = x^{k+1}$

## 6.1 Método dos Gradientes Conjugados (GC)

A base do método GC é a seguinte:

Propriedade: É possível escolher  $n$  direções linearmente independentes,  $v_1, v_2, \dots, v_n$  e, por meio da minimização da função  $f(x^k + sv_k)$ , em cada uma das direções separadamente, construir uma seqüência de aproximações que forneça o mínimo da função (6.1) após  $n$  passos ( $n$  é o número de equações do sistema).

Neste método, cada novo vetor  $v_j$  é gerado, como descreveremos a seguir. Dada uma aproximação inicial  $x^0$ , tomamos a primeira direção  $v_1 = -\nabla f(x^0)$ . Conhecidas as direções  $v^1, \dots, v^j$ , e as aproximações  $x^1, \dots, x^j$ , tomamos:

- 1)  $v_{j+1}$  tal que  $v_j^t A v_{j+1} = 0$ .

- 2)  $s_{j+1}$  é o número real que minimiza a  $f(x_j + s v_{j+1})$ , ou seja, o mínimo de  $f$  ao longo da reta que passa por  $x_j$  e tem a direção  $v_{j+1}$ .

- 3)  $x_{j+1} = x_j + s_{j+1} v_{j+1}$ .

Existem muitas maneiras de construir vetores que satisfaçam o item 1. No Algoritmo (6.2), que poderá ser usado na implementação do método GC, apresentamos um destes procedimentos. Quanto ao item 2, repetimos os mesmos cálculos realizados ao minimizar a função (6.4) e podemos mostrar que

$$s_{j+1} = \frac{r_j^t v_{j+1}}{v_{j+1}^t A v_{j+1}}$$

onde, conforme a notação anterior,  $r_j = b - Ax_j$ .

### Algoritmo 6.2 Método dos Gradientes Conjugados

Dados  $A, b$

$$x_0 = 0, r_0 = b$$

Enquanto  $r_k \neq 0$  faça

$$k = k + 1$$

se  $k = 1$

$$v_1 = r_0$$

caso contrário

$$v_k = r_{k-1} + \frac{r_{k-1}^t r_{k-1}}{r_{k-2}^t r_{k-2}} v_{k-1}$$

fim

$$s = \frac{r_{k-1}^t r_{k-1}}{v_k^t A v_k}$$

$$x_k = x_{k-1} + s v_k$$

$$r_{k-1} - s A v_k$$

fim

## 6.2 Pré-Condicionamento e Matrizes Não Simétricas

A matriz de Hilbert, serve-nos como exemplo de como o mal condicionamento da matriz do sistema pode prejudicar a aplicação do Método dos Gradientes Conjugados. Isto acontece devido à ampliação dos erros de arredondamento quando a matriz é mal condicionada. O mal condicionamento de uma matriz ocorre quando seus autovalores se distribuem num intervalo muito grande ou quando existem autovalores muito próximos de zero.

Para ilustrar este tipo de comportamento, o número de condição da matriz de Hilbert de dimensão 15 é da ordem de  $10^{17}$ . Tomando  $tol = 10^{-10}$ , o método GC não converge neste caso.

Em alguns casos, é possível acelerar a convergência do Método GC usando um pré-condicionamento do sistema, que consiste em substituir o sistema  $Ax = b$  por outro equivalente,  $M^{-1}Ax = M^{-1}b$ , onde  $M$  é uma matriz “próxima” de  $A$  que satisfaz às seguintes propriedades:

- 1)  $M$  é simétrica positiva-definida;
- 2)  $M^{-1}A$  é bem condicionada;
- 3) a equação  $Mx = b$  é fácil de ser resolvida.

Não há uma regra geral para encontrar matrizes pré-condicionadoras.

Existem recomendações para alguns casos particulares, como quando os elementos da diagonal variam num intervalo grande, o pré-condicionamento pela matriz diagonal,  $M = D$ , pode ser usado. Outras escolhas recaem em matrizes bloco-diagonais, fatoração de Cholesky e fatoração LU incompleta.

Tendo escolhido a matriz pré-condicionadora, o Algoritmo (6.3) pode ser usado na implementação do método, utilizando o Matlab.

### Algoritmo 6.3 Método Gradiente Conjugado com Pré-Condicionamento

Dados  $A$ ,  $b$ ,  $M$  (matriz pré-condicionadora),  $max$ ,  $tol$

$$1) x^0 = 0, r = b, v = M^{-1}b, y = M^{-1}r, aux = y^t r$$

2) Para  $k = 0 : max$ , faça

$$3) z = Av$$

$$4) s = \frac{aux}{v^t z}$$

$$5) x^{k+1} = x^k + sv$$

$$6) r = r - sz$$

$$7) y = M^{-1}r$$

8) Se  $r^t r < tol$  então

9) Saída com  $sol = x^{k+1}$

$$10) aux1 = y^t r$$

$$11) m = aux1/aux; aux = aux1$$

$$12) v = y + mv$$

O Método GC não pode ser usado em sistemas cujas matrizes não são simétricas, uma vez que os argumentos nesta seção não mais se aplicam. Uma saída para esta dificuldade pode ser pré-multiplicar à esquerda, o sistema linear (equações normais):

$$A^t Ax = A^t b.$$

Como  $A^t A$  é uma matriz simétrica positiva definida, o Método GC pode ser usado neste sistema equivalente ao inicial,  $Ax = b$ . Mas se a matriz  $A$  for mal condicionada, isto pode não ser satisfatório, pois os autovalores das matrizes mal condicionadas se distribuem num intervalo grande ou são muito próximos de zero. Como os autovalores de  $A^t A$  são os autovalores de  $A$  elevados ao quadrado, o problema se agrava com a nova forma do sistema.

Nesse caso, existem outras alternativas, dentre eles os métodos GMRES (Generalized Minimum Residual) e CGS (Gradientes Conjugados Quadrático).

## 7 Exemplos Numéricos

### 7.1 Um Problema de Difusão-Convecção

Considere o seguinte problema num domínio  $\Omega \subset \mathbb{R}^2$ :

$$\begin{cases} -\alpha(x, y)\Delta u + \beta^t(x, y)\nabla u = f(x, y) & em \quad \Omega \\ u(x, y) = 0 & em \quad \partial\Omega, \end{cases}$$

com  $\alpha(x, y) > 0$ .

Trata-se de um problema estacionário de difusão-convecção. O termo do Laplaciano corresponde a um fenômeno de difusão cujo coeficiente é  $\alpha(x, y)$  enquanto as derivadas de primeira ordem correspondem ao fenômeno de convecção na direção do campo de velocidades  $\vec{\beta} = (\beta_1(x, y), \beta_2(x, y))^t$ .

Testaremos o método dos Gradiente Conjugados, na resolução do problema acima. Para isso, escolhemos  $\Omega = [0, 1] \times [0, 1]$ ,  $\alpha(x, y) = 2$ ,  $f(x, y) = 1$  e  $\vec{\beta} = (1, 1)^t$  de forma a garantir o bom comportamento do problema.

A discretização foi obtida utilizando-se diferenças finitas centrais e uma malha regular contendo npt pontos em cada direção  $(x, y)$  originando um sistema linear esparsa e sem simetria com dimensão ndim. Para testar a convergência do método, tomamos  $tol = 10^{-8}$  como condição de parada.

Na tabela 1, encontramos o número de iterações it necessárias para a resolução do sistema linear resultante da discretização do problema de difusão-convecção, utilizando o método GC aplicado às equações normais (GCEN), em função da dimensão ndim do sistema linear,

npt	ndim	iterações	it/ndim
17	225	137	0,609
25	529	307	0,580
29	729	413	0,566
37	1225	677	0,552
41	1521	835	0,549

Tabela 1

Encontramos também a relação entre o número de iterações necessárias para a obtenção da convergência e a dimensão do sistema linear (it/ndim).

Podemos observar que, em todas as situações realizadas, o método GC convergiu em menos ndim iterações.

Além disso, é importante notar que, ao aumentarmos a dimensão do sistema linear, a relação entre it e ndim diminui. Isto sugere que o método GC seja apropriado para sistemas lineares de grande porte.

## 7.2 Método de Newton para Minimização Irrestrita

Vamos utilizar o seguinte problema para exemplificar o uso do Método de Newton (Algoritmo 5.2) para obter o ponto de mínimo de

$$f(x_1, x_2) = (x_1 - 2)^4 + (x_1 - 2)^2 x_2^2 + (x_2 + 1)^2$$

partindo da aproximação inicial  $x^0 = (1, 1)^t$ .

Isto equivale à resolução do sistema não-linear  $\nabla f(x) = 0$ , onde

$$\nabla f(x_1, x_2) = \begin{bmatrix} 4(x_1 - 2)^3 + 2(x_1 - 2)x_2^2 \\ 2x_2(x_1 - 2)^2 + 2(x_2 + 1) \end{bmatrix}.$$

	$x^k$	$f(x^k)$
$k = 0$	$(1; 1)^t$	6
$k = 1$	$(1; -0,5)^t$	1,5
$k = 2$	$(1,391; -0,6956)^t$	$4,09 \times 10^{-1}$
$k = 3$	$(1,7459; -0,9487)^t$	$6,49 \times 10^{-2}$
$k = 4$	$(1,9862; -1,0482)^t$	$2,53 \times 10^{-3}$
$k = 5$	$(1,998; -1,0001)^t$	$1,63 \times 10^{-6}$
$k = 6$	$(1,9999; -1,0000)^t$	$2,75 \times 10^{-12}$

Tabela 2

Na tabela 2, encontramos os resultados para tal problema. Nele, podemos reparar a boa aproximação de  $x^* = (2, -1)^t$  obtida e a rapidez com que isso acontece.

## *Conclusão*

Os métodos iterativos usados para resolução de sistemas lineares mostraram-se eficientes, tornando viável a busca de soluções aproximadas para este tipo de problema.

Além disso, este trabalho foi uma boa oportunidade de correlacionar conteúdos estudados tanto nas disciplinas de cálculo quanto nas disciplinas de álgebra linear do curso, bem como tomar contato com Problemas de Programação Não-Linear.

De acordo com as proposições e definições apresentadas nos primeiros capítulos, sobre condições de otimalidade, e fazendo uso do conceito de convexidade para funções, juntamente com as buscas direcionais e os métodos clássicos de descida, pudemos ilustrar ao final deste trabalho que o método Gradiente Conjugado aplicado às equações normais apresentou um bom resultado, nos testes realizados para a resolução de um problema de Difusão-Convecção.

Assim, deixo como sugestão para trabalhos futuros a implementação do algoritmo dos Gradientes Conjugados com Pré-Condicionamento, e além disso, de outros métodos do mesmo tipo: GMRES (Generalized Minimum Residual) e CGS (Gradientes Conjugados Quadráticos).

## *Referências*

- 1 CUNHA, MARIA CRISTINA C.; *“Métodos Numéricos”*, 2ªed., UNICAMP, 2000.
- 2 FRIEDLANDER, ANA; *“Elementos de Programação Não-Linear”*, Editora da UNICAMP, Campinas - SP, 1994.
- 3 GOLUB,G.H.; VAN LOAN C.F.; *“Matrix Computations”*, John Hopkins University Press, 1989.
- 4 LUENBERGER,D.G.; *“Linear and nonlinear programming”*, 2ª ed., Nova York, Addison - Wesley Publishing Company, 1986.
- 5 MARTINEZ, JOSÉ M.; SANTOS, SANDRA A.; *“Métodos Computacionais de Otimização”*, Sociedade Brasileira de Matemática Aplicada e Computacional, Goiânia, 1996.