UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE BLUMENAU
CURSO DE LICENCIATURA EM MATEMÁTICA

João Vitor Pamplona

MÉTODO DE GRADIENTES CONJUGADOS:
um estudo teórico aplicado a quadráticas convexas

Blumenau
2019

João Vitor Pamplona

MÉTODO DE GRADIENTES CONJUGADOS:
um estudo teórico aplicado a quadráticas convexas

Trabalho de Conclusão de Curso submetido ao Curso de Licenciatura em Matemática do Centro de Blumenau da Universidade Federal de Santa Catarina para a obtenção do título de Licenciado em Matemática.
Orientador: Prof. Luiz Rafael dos Santos, Dr.

Blumenau
2019

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Pamplona, João Vitor
Método de Gradientes Conjugados : um estudo teórico aplicado a quadráticas convexas / João Vitor Pamplona ; orientador, Luiz Rafael dos Santos, 2019.
72 p.

Trabalho de Conclusão de Curso (graduação) - Universidade Federal de Santa Catarina, Campus Blumenau, Graduação em Matemática, Blumenau, 2019.

Inclui referências.

1. Matemática. 2. Otimização. 3. Métodos iterativos. 4. Método de Máxima Descida. 5. Método dos Gradientes Conjugados. I. dos Santos, Luiz Rafael. II. Universidade Federal de Santa Catarina. Graduação em Matemática. III. Título.

João Vitor Pamplona

Método de Gradientes Conjugados: um estudo teórico aplicado a quadráticas convexas

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do Título de "Licenciado em Matemática" e aprovado em sua forma final pelo Curso de Licenciatura em Matemática do Centro de Blumenau da Universidade Federal de Santa Catarina.

Blumenau, 28 de novembro de 2019.

Prof. André Vanderlinde da Silva, Dr.
Coordenador Curso de Licenciatura em Matemática

Banca Examinadora:

Prof. Luiz Rafael dos Santos, Dr.
Orientador
Universidade Federal de Santa Catarina

Prof. Felipe Delfini Caetano Fidalgo, Dr.
Avaliador
Universidade Federal de Santa Catarina

Prof. Hugo José Lara Urdaneta, Dr.
Avaliador
Universidade Federal de Santa Catarina

Este trabalho é dedicado aos Antônios e Marias da minha vida.

AGRADECIMENTOS

Agradeço aos meus pais, Antônio e Maria Helena por sempre me apoiarem, não importando as circunstâncias, o apoio de vocês é fundamental na minha vida, amo vocês.

Agradeço as minhas tias, Lu e Tata, por serem praticamente minhas segundas mães e me ajudarem em tudo que eu precisava.

Ao meu orientador, professor Luiz Rafael dos Santos, por acreditar que podíamos estudar este assunto e por me mostrar que todo esforço vale a pena.

Agradeço ao meu cachorro Toco por sempre estar comigo nos momentos difíceis, e que após anos vendo eu estudar já aprendeu alguns teoremas.

Agradeço também aos professores do curso de Licenciatura em Matemática da UFSC Blumenau, em especial ao professor Jorge Cássio, que além de professor foi um grande amigo nesta graduação e aos professores do SOMA por acreditarem no meu potencial.

Agradeço à minha namorada, que esteve presente em todos os momentos da graduação. Maria, dentre todas as coisas que a matemática poderia me oferecer a melhor com certeza foi você, que nosso amor seja contínuo e não-enumerável.

Finalmente, agradeço a todos os amigos que fiz nesta graduação.

RESUMO

Neste trabalho é apresentado o método dos Gradientes Conjugados. Para isso, foi realizado um estudo sobre o problema de otimização, condições de otimalidade e convexidade, buscando entender os principais resultados e definições e apresentando alguns exemplos. Além disso, é estudado o método dos gradientes, para que possamos entender a diferença na velocidade de convergência desse método e o de gradientes conjugados. Finalmente, os resultados matemáticos que embasam o Método de Gradientes Conjugados são apresentados.

Palavras-chave: Otimização, Métodos iterativos, Gradientes conjugados.

ABSTRACT

In this work the Conjugate Gradients method is presented. In order to achieve this objective, we study first mathematical optimization background, optimality conditions and convexity, exposing the main definitions and results and illustrating with some examples. Moreover, the Gradient method is studied, thus we can understand the difference in the convergence speed of this method and the conjugate gradient method. Finally, the mathematical results that lead to the conjugate gradient method are presented.

Keywords: Optimization, Iterative methods, Conjugate Gradient.

LISTA DE FIGURAS

SUMÁRIO

INTRODUÇÃO

Neste trabalho estudaremos o método dos Gradientes Conjugados, exploraremos desde sua fundamentação teórica, que envolvem conceitos de otimização, até a definição do algoritmo propriamente dito para a minimização de uma função quadrática estritamente convexa.

Assim nosso principal objetivo é estudar a fundamentação matemática que suporta o Algoritmo dos Gradientes Conjugados em problemas de otimização quadrática convexa, aplicado ao problema:

$$\min_{x\in\mathbb{R}^n} f(x) = \frac{1}{2}x^\top Ax - b^\top x + c \tag{1}$$

em que $A \in \mathbb{R}^{n\times n}$ é uma matriz simétrica definida positiva e $x, b \in \mathbb{R}^n$ são vetores e $c \in \mathbb{R}$ é um escalar. O vetor x é o vetor das incógnitas.

Também estudaremos outro método: o método de máxima descida ou método do gradiente, como também é conhecido. Tal método trabalha apenas com informações da primeiras derivadas, porém pode demorar a convergir. Já o método dos gradientes conjugados temos convergência em n passos, se a função quadrática for estritamente convexa.

Investigaremos as propriedade de direções conjugadas e em cada iteração do método resolvemos um problema de minimização unidirecional.

O método do gradientes conjugados é utilizado principalmente para a resolução de problemas de grande porte, nos quais não é razoável sequer fazer as n iterações que podem ser demandadas no caso aqui já citado.

Neste contexto, os Capítulos 1 e 2 deste Trabalho de Conclusão de Curso foram baseados no livro de Ribeiro e Karas [11]. No Capítulo 1 apresentamos os conceitos de otimização, condições de otimalidade e como a convexidade nos auxilia. No Capítulo 2 começamos a estudar algoritmos de descida propriamente ditos, mas ainda de uma forma geral.

Já o Capítulo 3 foi baseado em Luenberger e Ye [8], Ribeiro e Karas [11] e Shewchuk [12], nele trabalhamos o método dos gradientes, uma extensão mais específica do capítulo anterior, conferimos a conver-

gência do método e fizemos uma análise geométrica do funcionamento do algoritmo. Por fim, no Capítulo 4 estudamos o método dos gradientes conjugados, tais como suas propriedades e interpretações geométricas e utilizamos, como principais referências Luenberger e Ye [8], Ribeiro e Karas [11], Shewchuk [12] e Watkins [13],

Algumas das figuras presentes no trabalho foram elaboradas com auxílio dos softwares `Julia`[1] e `Ipe`.

1 OTIMIZAÇÃO

Neste capítulo apresentaremos os conceitos básicos de otimização. Para iniciar, vamos apresentar alguns resultados que garantem a existência de um minimizador de determinada função e, em seguida, discutiremos as condições de otimalidade para problemas de minimização irrestrita.

1.1 O PROBLEMA DE OTIMIZAÇÃO

Vamos considerar o problema

$$\begin{array}{c} \min f(x) \\ \text{sujeito à } x \in \Omega \end{array}, \tag{PG}$$

em que $f : \mathbb{R}^n \to \mathbb{R}$ é uma função arbitrária e $\Omega \subset \mathbb{R}^n$ é um conjunto qualquer. Como tal problema sugere minimizar a função, primeiro precisamos saber o que é um minimizador.

Definição 1.1. Considere uma função $f : \mathbb{R}^n \to \mathbb{R}$ e $x^* \in \Omega \subset \mathbb{R}^n$. Dizemos que x^* é um *minimizador local* de f em Ω quando existe $\delta > 0$, tal que $f(x^*) \leq f(x)$ para todo $x \in B(x^*, \delta) \cap \Omega$. Caso $f(x^*) \leq f(x)$ para todo $x \in \Omega$, x^* é dito *minimizador global* de f em Ω.

Quando as desigualdade na Definição 1.1 forem estritas para $x \neq x^*$, chamaremos x^* de minimizador estrito. Se não for mencionado, o conjunto Ω significa que $\Omega = \mathbb{R}^n$ e portanto estamos trabalhando com um problema irrestrito, ou seja, sem que haja restrições aos valores das variáveis.

Vamos agora ver algumas condições que garantem a existência de minimizadores.

Teorema 1.1 (Weierstrass)**.** *Sejam $f : \mathbb{R}^n \to \mathbb{R}$ contínua e $\Omega \subset \mathbb{R}^n$ compacto não vazio. Então existe minimizador global de f em Ω.*

Demonstração. Veja em Lima [7, pp. 187]. ■

Corolário 1.1. *Seja $f : \mathbb{R}^n \to \mathbb{R}$ contínua e suponha que existe $c \in \mathbb{R}$ tal que o conjunto $L = \{x \in \mathbb{R}^n \mid f(x) \leq c\}$ seja compacto e não vazio. Então f tem um minimizador global.*

Demonstração. Pelo Teorema 1.1, existe $x^* \in L$ tal que $f(x^*) \leq f(x)$, para todo $x \in L$. Por outro lado, se $x \notin L$, temos $f(x) > c \geq f(x^*)$. Assim, $f(x^*) \leq f(x)$, para todo $x \in \mathbb{R}^n$. ■

A seguir veremos uma definição que será muito importante ao longo deste trabalho e também um exemplo de função, que utilizando o corolário acima, encontraremos o minimizador.

Definição 1.2. Seja $A \in \mathbb{R}^{n\times n}$ uma matriz simétrica. Dizemos que A é *definida positiva* quando $x^\top Ax > 0$, para todo $x \in \mathbb{R}^n\backslash\{0\}$. Se $x^\top Ax \geq 0$, para todo $x \in \mathbb{R}^n$, A é dita *semidefinida positiva*.

Exemplo 1.1. Seja $f : \mathbb{R}^n \to \mathbb{R}$ dado por $f(x) = x^\top Ax$, em que $A \in \mathbb{R}^{n\times n}$ é uma matriz simétrica. Mostraremos que f tem um minimizador global x^* em $S = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$. Com efeito, como f é contínua e S é compacto, segue do Corolário 1.1 que f possui um minimizador global.

O próximo teorema nos auxiliará no caso em que a função é quadrática.

Teorema 1.2. *Sejam $A \in \mathbb{R}^n$ uma matriz simétrica e $\delta > 0$. Se $x^\top Ax \geq 0$, para todo $x \in \mathbb{R}^n$ tal que $\|x\| = \delta$, então $x^\top Ax \geq 0$, para todo $x \in \mathbb{R}^n$.*

Demonstração. Considere $x \in \mathbb{R}^n\backslash\{0\}$. Tomando $y = \dfrac{\delta x}{\|x\|}$, temos que $\|y\| = \delta$. Portanto, usando a hipótese, temos que $\dfrac{\delta^2}{\|x\|^2} x^\top Ax = y^\top Ay \geq 0$. Assim $x^\top Ax \geq 0$. ■

Agora discutiremos os critérios de otimalidade.

1.2 CONDIÇÕES DE OTIMALIDADE

Veremos agora as condições necessárias e suficientes para caracterizar um minimizador de um problema (PG) com $\Omega = \mathbb{R}^n$ e alguns exemplos de cada uma.

Definição 1.3. Um ponto $x^* \in \mathbb{R}^n$ que cumpre $\nabla f(x^*) = 0$ é dito *ponto crítico* ou *ponto estacionário* da função f.

Teorema 1.3 (Condição necessária de 1ª ordem)**.** *Seja $f : \mathbb{R}^n \to \mathbb{R}$ diferenciável no ponto $x^* \in \mathbb{R}^n$. Se x^* é um minimizador local de f, então*

$$\nabla f(x^*) = 0 \tag{2}$$

Demonstração. Considere $d \in \mathbb{R}^n \backslash \{0\}$ arbitrário. Como x^* é minimizador local, pela Definição 1.1, existe $\delta > 0$ tal que

$$f(x^*) \leq f(x^* + td)$$

para todo $t \in (0, \delta)$.

Pela expansão de Taylor, veja Guidorizzi [3, pp. 465],

$$f(x^* + td) = f(x^*) + t\nabla f(x^*)^\top d + r(t)$$

com

$$\lim_{t \to 0} \frac{r(t)}{t} = 0.$$

Usando a igualdade acima e a desigualdade da Definição 1.1 obtemos,

$$\begin{aligned} f(x^*) &\leq f(x^*) + t\nabla f(x^*)^\top d + r(t) \\ f(x^*) - f(x^*) &\leq t\nabla f(x^*)^\top d + r(t) \\ 0 &\leq t\nabla f(x^*)^\top d + r(t). \end{aligned}$$

Tomando d arbitrário e dividindo-se tudo por t, obtemos

$$0 \leq \nabla f(x^*)^\top d + \frac{r(t)}{t}.$$

Passando o limite quando $t \to 0$, obtemos

$$\nabla f(x^*)^\top d \geq 0. \tag{3}$$

Suponha, por contradição, que $\nabla f(x^*)$ não fosse nulo. Então poderíamos escolher $d = -\nabla f(x^*)$, resultando em

$$\begin{aligned}\|\nabla f(x^*)\|^2 &= \nabla f(x^*)^\top \nabla f(x^*) \\ &= (-\nabla f(x^*))^\top (-\nabla f(x^*)) \\ &= d^\top (-\nabla f(x^*)) \\ &= -(d^\top \nabla f(x^*)) \\ &\leq 0,\end{aligned}$$

por conta de (3), mas que é uma contradição. Logo $\nabla f(x^*) = 0$. ■

Exemplo 1.2. Seja $f : \mathbb{R}^3 \to \mathbb{R}$ dada por $f(x) = \sin(3x_1^2 + x_2^2) + \cos(x_1^2 - x_2^2) + 4x_3$. Verificaremos se f tem minimizadores em $\mathbb{R}^3$. Note que $\nabla f(x) \neq 0$, para todo $x \in \mathbb{R}^3$, pois $\dfrac{\partial f}{\partial x_3}(x) = 4$. Portanto, pelo Teorema 1.3, não existe minimizador de f em $\mathbb{R}^3$.

Teorema 1.4 (Condição necessária de 2ª ordem)**.** *Seja $f : \mathbb{R}^n \to \mathbb{R}$ duas vezes diferenciável no ponto $x^* \in \mathbb{R}^n$. Se x^* é um minimizador local de f, então a matriz Hessiana de f no ponto x^* é positiva semi-definida, isto é,*

$$d^\top \nabla^2 f(x^*) d \geq 0.$$

para todo $d \in \mathbb{R}^n$.

Demonstração. Considere $d \in \mathbb{R}^n \backslash \{0\}$ arbitrário. Usando a expansão de Taylor de segunda ordem em torno de $f(x^* + td)$, (veja Lima [6, pp. 262]) temos,

$$f(x^* + td) = f(x^*) + t\nabla f(x^*)^\top d + \frac{t^2}{2} d^\top \nabla^2 f(x^*) d + r(t),$$

com $\lim\limits_{t \to 0} \dfrac{r(t)}{t^2} = 0$. Como x^* é minimizador local, o Teorema 1.3 garante que $\nabla f(x^*) = 0$. Portanto, para t suficientemente pequeno,

$$0 \leq f(x^* + td) - f(x^*) = \frac{t^2}{2} d^\top \nabla^2 f(x^*) d + r(t).$$

Dividindo por t^2 e passando o limite quando $t \to 0$, obtemos $d^\top \nabla^2 f(x^*) d \geq 0$. ■

Antes de vermos e provarmos a condição suficiente de segunda ordem, vamos demostrar um lema que nos auxiliará.

Lema 1.1. *Se $A \in \mathbb{R}^{n\times n}$ é uma matriz simétrica com λ_1 e λ_n sendo o menor e o maior autovalor, respectivamente, então*

$$\lambda_1 x^\top x \leq x^\top A x \leq \lambda_n x^\top x,$$

para todo $x \in \mathbb{R}^n$

Demonstração. Como A é simétrica, então seus autovalores são reais. Logo, seja $\sigma(A) = \{\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n\} \subseteq \mathbb{R}$ e tome $\lambda_1 = \min \sigma(A)$ e $\lambda_n = \max \sigma(A)$. Como A é diagonalizável, então existem $Q \in \mathbb{R}^{n\times n}$ ortogonal e D diagonal, tal que

$$A = QDQ^\top \text{ ou } D = Q^\top A Q,$$

sendo $D \coloneqq \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Se $x = 0$ vale trivialmente, então suponha $x \in \mathbb{R}^n \backslash \{0\}$, temos que $y = Q^\top x$ é tal que

$$\begin{aligned} \|y\|^2 &= y^\top y \\ &= x^\top Q Q^\top x \\ &= x^\top x \\ &= \|x\|^2. \end{aligned}$$

Assim,

$$\begin{aligned} x^\top A x &= x^\top (QDQ^\top) x \\ &= (Q^\top x)^\top D (Q^\top x) \\ &= y^\top D y \\ &= \sum_{i=1}^{n} y_i \lambda_i y_i \\ &\leq \lambda_n \sum_{i=1}^{n} |y_i|^2 \\ &= \lambda_n y^\top y \\ &= \lambda_n \|x\|^2 \\ &= \lambda_n x^\top x. \end{aligned}$$

Logo $x^\top Ax \leq \lambda_n x^\top x$.

Analogamente temos que

$$\begin{aligned} x^\top Ax &= x^\top (QDQ^\top)x \\ &= (Q^\top x)^\top D(Q^\top x) \\ &= y^\top Dy \\ &= \sum_{i=1}^{n} y_i \lambda_i y_i \\ &\geq \lambda_1 \sum_{i=1}^{n} |y_i|^2 \\ &= \lambda_1 y^\top y \\ &= \lambda_1 \|x\|^2 \\ &= \lambda_1 x^\top x. \end{aligned}$$

Logo $x^\top Ax \geq \lambda_1 x^\top x$. ■

Teorema 1.5 (Condição suficiente de 2ª ordem)**.** *Seja $f : \mathbb{R}^n \to \mathbb{R}$ duas vezes diferenciável no ponto $x^* \in \mathbb{R}^n$. Se x^* é um ponto estacionário da função f e $\nabla^2 f(x^*)$ é definida positiva, então x^* é minimizador local estrito de f.*

Demonstração. Seja λ o menor autovalor de $\nabla^2 f(x^*)$. Como esta matriz é definida positiva, temos $\lambda > 0$ (veja Meyer [9, pg. 559]).

Com isto, pela expansão de Taylor, veja Guidorizzi [3, pp. 465], usando o fato de x^* ser estacionário e usando o fato de que $d^\top \nabla^2 f(x^*) d \geq \lambda \|d\|^2$, para todo $d \in \mathbb{R}^n$, pelo Lema 1.1, temos

$$\begin{aligned} f(x^* + d) &= f(x^*) + \nabla f(x^*)^\top d + \frac{1}{2} d^\top \nabla^2 f(x^*) d + r(d) \\ &= f(x^*) + \frac{1}{2} d^\top \nabla^2 f(x^*) d + r(d) \\ &\geq f(x^*) + \frac{1}{2} \lambda \|d\|^2 + r(d), \end{aligned}$$

em que $\lim_{d \to 0} \dfrac{r(d)}{\|d\|^2} = 0$. Podemos então escrever

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

Como $\lim_{d\to 0}\left(\frac{\lambda}{2}+\frac{r(d)}{\|d\|^2}\right)>0$, existe $\delta>0$ tal que $\frac{\lambda}{2}+\frac{r(d)}{\|d\|^2}>0$, para todo $d\in B(0,\delta)\backslash\{0\}$, donde segue que $f(x^*+d)-f(x^*)>0$, para todo $d\in B(0,\delta)\backslash\{0\}$, ou equivalentemente,

$$f(x^*)<f(x),$$

para todo $x\in B(x^*,\delta)\backslash\{x^*\}$. Isto prova que x^* é minimizador local. ■

Exemplo 1.3. Seja $f:\mathbb{R}^2\to\mathbb{R}$ dada por $f(x)=(x_1-x_2^2)(x_1-\frac{1}{2}x_2^2)$. Verificaremos que $\bar{x}=0$ é o único ponto estacionário de f e não é minimizador. No entanto, fixada qualquer direção $d\in\mathbb{R}^2\backslash\{0\}$, $\bar{x}$ minimiza localmente f ao longo de d.

Temos $\nabla f(x)=\begin{bmatrix}2x_1-\frac{3}{2}x_2^2\\ -3x_1x_2+2x_2^3\end{bmatrix}$. Assim, se $\nabla f(x)=0$, então $x=0$. Além disso, $f(\frac{2}{3}x_2^2,x_2)=-\frac{x_2^4}{18}<0$, o que significa que $\bar{x}$ não é minimizador local de f. Porém, dado $d\in\mathbb{R}^2\backslash\{0\}$, temos

$$f(\bar{x}+td)=t^2(d_1-td_2^2)(d_1-\frac{1}{2}td_2^2).$$

Se $d_1=0$, então $f(\bar{x}+td)=\frac{1}{2}t^4d_2^4\geq 0$. Caso $d_1\neq 0$, a expressão $(d_1-td_2^2)(d_1-\frac{1}{2}td_2^2)$ é positiva em $t=0$ e, por continuidade, também para t próximo de 0, logo $\bar{x}$ minimiza localmente f ao longo de d.

1.3 CONVEXIDADE

Dentre as várias classes de funções estudadas em matemática, existe uma que se destaca pelas excelentes propriedades que possui: a classe das funções convexas. Em otimização, a convexidade permite por exemplo concluir que, para estas funções, minimizadores locais são globais.

Os conjuntos convexos constituem o domínio natural para as funções convexas, conforme veremos a seguir.

Definição 1.4. Um conjunto $C\subset\mathbb{R}^n$ é dito *convexo* quando dados $x,y\in C$, o segmento $[x,y]=\{(1-t)x+ty\mid t\in[0,1]\}$ estiver inteiramente contido em C.

Temos a seguir, nas Figuras 1 e 2 exemplos de conjuntos convexos e não convexos.

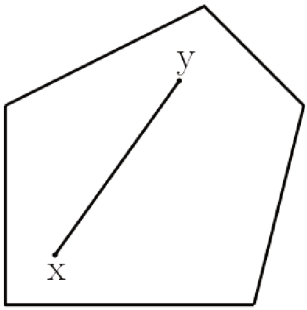

Figura 1 – Conjunto convexo.

Fonte – Elaborado por João Vitor Pamplona (2019).

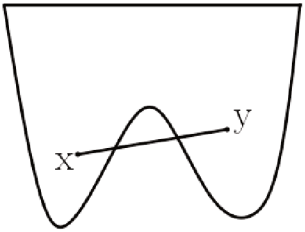

Figura 2 – Conjunto não convexo.

Fonte – Elaborado por João Vitor Pamplona (2019).

Agora definiremos formalmente, o que são funções convexas.

Definição 1.5. Seja $C \subset \mathbb{R}^n$ um conjunto convexo. Dizemos que a função $f : \mathbb{R}^n \to \mathbb{R}$ é *convexa* em C quando

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$$

para todos $x, y \in C$ e $t \in [0, 1]$.

Geometricamente podemos dizer que qualquer arco no gráfico de uma função convexa está sempre abaixo do segmento que liga as

extremidades. Veja nas Figuras 3 e 4 exemplos de função convexa e não convexa, respectivamente.

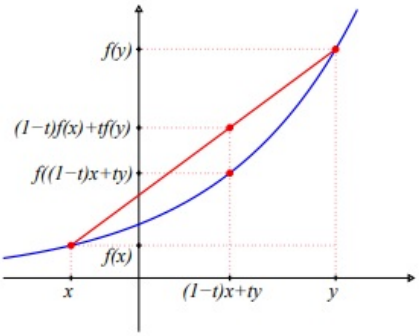

Figura 3 – Função convexa.

Fonte – Elaborado por Ribeiro e Karas [11].

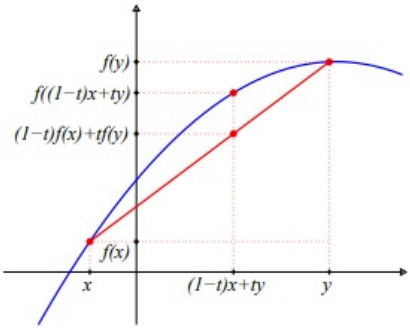

Figura 4 – Função não convexa.

Fonte – Elaborado por Ribeiro e Karas [11].

O próximo teorema justifica o fato de funções convexas serem muito bem vistas em otimização.

Teorema 1.6. *Sejam $C \subset \mathbb{R}^n$ convexo e $f : C \to \mathbb{R}$ uma função convexa. Se $x^* \in C$ é minimizador local de f, então x^* é minimizador global de f.*

Demonstração. Por definição de mínimo local, seja $\delta > 0$ tal que $f(x^*) \leq f(x)$, para todo $x \in B(x^*, \delta) \cap C$. Seja $y \in C$ mas tal que $y \notin B(x^*, \delta)$ e tome $t > 0$ de modo que $t\|y - x^*\| < \delta$. Assim, o ponto $\bar{x} := (1-t)x^* + ty$ tal que $\bar{x} \in C$ satisfaz

$$\begin{aligned}\|\bar{x} - x^*\| &= \|(1-t)x^* + ty - x^*\| \\ &= \|x^* - tx^* + ty - x^*\| \\ &= \|ty - tx^*\| \\ &= |t|\|y - x^*\| \\ &= |t|\|y - x^*\| < \delta,\end{aligned}$$

o que implica que $\bar{x} \in B(x^*, \delta)$ e, portanto, $\bar{x} \in B(x^*, \delta) \cap C$. Deste modo temos,

$$f(x^*) \leq f(\bar{x}) = f((1-t)x^* + ty) \leq (1-t)f(x^*) + tf(y),$$

donde segue que

$$\begin{aligned}f(x^*) - (1-t)f(x^*) &\leq tf(y) \\ f(x^*) - f(x^*) + tf(x^*) &\leq tf(y) \\ tf(x^*) &\leq tf(y) \\ f(x^*) &\leq f(y),\end{aligned}$$

o que finaliza a demonstração. ■

Quando temos a diferenciabilidade da função, podemos caracterizar a convexidade de forma mais simples.

Teorema 1.7. *Sejam $f : \mathbb{R}^n \to \mathbb{R}$ uma função diferenciável e $C \subset \mathbb{R}^n$ convexo. A função f é convexa se, e somente se,*

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

Demonstração. Seja f convexa. Para $x, y \in C$ e $t \in (0, 1]$ quaisquer, definindo $d \coloneqq y - x$ temos $x + td \in C$ e

$$f(x + td) = f(x + t(y - x)) = f((1-t)x + ty) \leq (1-t)f(x) + tf(y).$$

Portanto,

$$f(y) - f(x) \geq \lim_{t \to 0^+} \frac{f(x + td) - f(x)}{t} = \nabla f(x)^\top d,$$

em que a última igualdade é dada pela definição de derivada direcional de f. Usando a definição de d, obtemos $f(y) \geq f(x) + \nabla f(x)^\top (y - x)$.

Para provar a recíproca, considere $z = (1 - t)x + ty \in C$ e observe que, por hipótese

$$f(x) \geq f(z) + \nabla f(z)^\top (x - z) \text{ e } f(y) \geq f(z) + \nabla f(z)^\top (y - z).$$

Multiplicando a primeira desigualdade acima por $(1 - t)$ e a segunda por t, obtemos

$$(1 - t)f(x) \geq (1 - t)f(z) + (1 - t)\nabla f(z)^\top (x - z)$$

e

$$tf(y) \geq tf(z) + t\nabla f(z)^\top (y - z).$$

Somando as duas desigualdades acima,

$$\begin{aligned}
(1 - t)f(x) + tf(y) &\geq f(z) + \nabla f(z)^\top ((1 - t)(x - z) + t(y - z)) \\
&= f(z) + \nabla f(z)^\top (x - z - tx + tz + ty - tz) \\
&= f(z) + \nabla f(z)^\top (x - z - tx + ty) \\
&= f(z) + \nabla f(z)^\top ((1 - t)x + ty - z) \\
&= f(z) + \nabla f(z)^\top (z - z) \\
&= f(z) + \nabla f(z)^\top (0) \\
&= f(z).
\end{aligned}$$

Substituindo novamente $z = (1 - t)x + ty$ derivamos,

$$(1 - t)f(x) + tf(y) \geq f((1 - t)x + ty),$$

que é a definição de função convexa e portanto completa a demostração. ■

Podemos interpretar geometricamente este resultado dizendo que uma função convexa está sempre acima da sua aproximação linear,

pois dados $a, x \in C$, temos $f(x) \geq f(a) + \nabla f(a)^\top (x - a)$. Veja a Figura 5 para um exemplo de uma função real convexa e diferenciável.

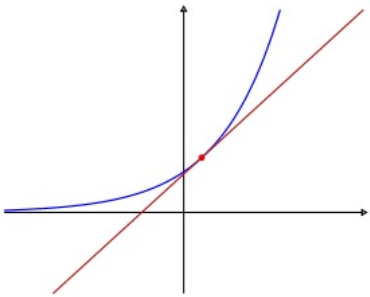

Figura 5 – Aproximação linear de f

Fonte – Elaborado por Ribeiro e Karas [11].

O teorema seguinte nos fornece outro critério para caracterizar convexidade.

Teorema 1.8. *Sejam $f : \mathbb{R}^n \to \mathbb{R}$ uma função de classe C^2 e $C \subset \mathbb{R}^n$ convexo. Se $\nabla^2 f(x) \geq 0$, para todo $x \in C$, então f e convexa em C.*

Demonstração. Dados $x \in C$ e $d \in \mathbb{R}^n$ tal que $x + d \in C$, pelo Teorema de Taylor com resto de Lagrange (veja Guidorizzi [4, pp. 306]), obtemos

$$f(x + d) = f(x) + \nabla f(x)^\top d + \frac{1}{2} d^\top \nabla^2 f(x + td) d,$$

para algum $t \in (0, 1)$. Como $\nabla^2 f(x+td) \geq 0$, logo $\frac{1}{2} d^\top \nabla^2 f(x+td) d \geq 0$ concluindo assim que,

$$f(x + d) \geq f(x) + \nabla f(x)^\top d.$$

Pelo Teorema 1.7, f é convexa. ■

Caracterizaremos agora uma função quadrática em que a matriz que a define é semidefinida positiva como uma função convexa.

Teorema 1.9. *Seja $f : \mathbb{R}^n \to \mathbb{R}$ dada por*

$$f(x) = \frac{1}{2}x^\top Ax - b^\top x + c, \tag{4}$$

em que $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$ e $c \in \mathbb{R}$, $f(x)$ é convexa se, e somente se A é é uma matriz semidefinida-positiva.

Demonstração. Seja $t = \frac{1}{2}$, $w \in C$. Então $x, y \in C$ com $w = x - y \neq 0$. Pela Definição 1.5 temos,

$$f\Big(\frac{x+y}{2}\Big) \leq \frac{1}{2}(f(x) + f(y)).$$

Aplicando à nossa quadrática obtemos,

$$\begin{aligned}
&\frac{1}{2}\Big[\Big(\frac{x+y}{2}\Big)^\top A\Big(\frac{x+y}{2}\Big)\Big] - b^\top\Big(\frac{x+y}{2}\Big) + c \\
&\quad\leq \frac{1}{2}\Big[\Big(\frac{1}{2}x^\top Ax - b^\top x + c\Big)\Big(\frac{1}{2}y^\top Ay - b^\top y + c\Big)\Big],
\end{aligned}$$

e então,

$$\begin{aligned}
\Big(\frac{x+y}{2}\Big)^\top A\Big(\frac{x+y}{2}\Big) &\leq \frac{1}{2}x^\top Ax + \frac{1}{2}y^\top Ay \\
\frac{1}{2}x^\top Ax + \frac{1}{2}y^\top Ay - \Big(\frac{x+y}{2}\Big)^\top A\Big(\frac{x+y}{2}\Big) &\geq 0 \\
x^\top Ax + y^\top Ay - \Big(\frac{x+y}{2}\Big)^\top A(x+y) &\geq 0 \\
x^\top Ax + y^\top Ay - \Big(\frac{x+y}{2}\Big)^\top Ax - \Big(\frac{x+y}{2}\Big)^\top Ay &\geq 0 \\
\Big[x - \frac{x+y}{2}\Big]^\top Ax + \Big[y - \frac{x+y}{2}\Big]^\top Ay &\geq 0 \\
\Big[\frac{x-y}{2}\Big]^\top Ax - \Big[\frac{x-y}{2}\Big]^\top Ay &\geq 0 \\
\Big(\frac{x-y}{2}\Big)^\top A(x-y) &\geq 0 \\
(x-y)^\top A(x-y) &\geq 0 \\
w^\top Aw &\geq 0.
\end{aligned}$$

Assim, como w não-nulo é arbitrário, segue que A é semidefinida positiva, como queríamos. ∎

A seguir veremos um exemplo que, por ser a função convexa, terá solução única por ser estacionário, ou seja, seu minimizador local é global.

Exemplo 1.4. Sejam $A \in \mathbb{R}^{n\times n}$ uma matriz simétrica semidefinida positiva, $b \in \mathbb{R}^n$ e $c \in \mathbb{R}$. Suponha que a função quadrática $f : \mathbb{R}^n \to \mathbb{R}$ dada por

$$f(x) = \frac{1}{2}x^\top Ax - b^\top x + c$$

tem um minimizador local x^*. Pelo Teorema 1.9, temos que f é convexa. Então, usando o Teorema 1.6, x^* é minimizador global.

Uma outra propriedade interessante é que se x^* minimiza a quadrática estrita, isto é, em que A é definida positiva, então x^* é solução única do sistema $Ax = b$. A seguir mostraremos este fato.

Teorema 1.10. *Seja $f : \mathbb{R}^n \to \mathbb{R}$ dada por*

$$f(x) = \frac{1}{2}x^\top Ax - b^\top x + c, \tag{5}$$

em que $A \in \mathbb{R}^{n\times n}$ é uma matriz simétrica definida positiva, $b \in \mathbb{R}^n$ e $c \in \mathbb{R}$, encontrar x^* que minimiza f é equivalente a achar a solução de $Ax = b$.

Demonstração. De fato, seja

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

e $c \in \mathbb{R}$.

Substituindo em (5) obtemos

$$f(x) = \frac{1}{2}\begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} -$$

$$- \begin{bmatrix} b_1 & b_2 & \dots & b_n \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + c.$$

Agrupando de maneira adequada chegamos em:

$$f(x) = \frac{1}{2}[a_{11}x_1^2 + a_{12}x_1x_2 + \ldots + a_{1n}x_1x_n + a_{21}x_1x_2 + a_{22}x_2^2 + \ldots$$
$$+ a_{2n}x_2x_n + \ldots + a_{n1}x_1x_n + a_{n2}x_2x_n + \ldots + a_{nn}x_n^2] +$$
$$- b_1x_1 - b_2x_2 - \ldots - b_nx_n + c.$$

O gradiente desta função é dado por,

$$\nabla f(x) = \begin{bmatrix} a_{11}x_1 + \frac{a_{12}x_2}{2} + \ldots + \frac{a_{1n}x_n}{2} + \frac{a_{21}x_2}{2} + \ldots + \frac{a_{n1}x_n}{2} - b_1 \\ \frac{a_{12}x_1}{2} + \frac{a_{21}x_1}{2} + a_{22}x_2 + \ldots + \frac{a_{2n}x_n}{2} + \ldots + \frac{a_{n2}x_n}{2} - b_2 \\ \vdots \\ \frac{a_{1n}x_1}{2} + \ldots + \frac{a_{2n}x_2}{2} + \ldots + \frac{a_{n1}x_1}{2} + \ldots + a_{nn}x_n - b_n \end{bmatrix}$$

que pode ser escrito como

$$\nabla f(x) = \frac{1}{2}\begin{bmatrix} a_{11}x_1 + a_{21}x_2 + \ldots + a_{n1}x_n \\ a_{12}x_1 + a_{22}x_2 + \ldots + a_{n2}x_n \\ \vdots \\ a_{1n}x_1 + a_{2n}x_2 + \ldots + a_{nn}x_n \end{bmatrix} +$$

$$+ \frac{1}{2}\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

ou

$$\nabla f(x) = \frac{1}{2}\begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & & & \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} +$$

$$+\frac{1}{2}\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Consequentemente

$$\nabla f(x) = \frac{1}{2}A^\top x + \frac{1}{2}Ax - b \tag{6}$$

e como A é simétrica, isto é, $A = A^\top$ temos

$$\nabla f(x) = Ax - b. \tag{7}$$

Como visto, x^* é ponto estacionário se $\nabla f(x^*) = 0$. Então, igualando a equação (7) a zero obtemos o sistema linear $Ax = b$. Por ser A definida positiva, tal sistema tem solução única e, pelo mesmo motivo, e em conjunto com o Teorema 1.5 temos que x^* é solução única de $Ax = b$ se, e somente se, é minimizador global de f. ■

2 ALGORITMOS

Uma maneira de resolver o problema de otimização é encontrar um ponto crítico. Para isso, resolveríamos $\nabla f(x) = 0$, porém dificilmente conseguimos resolver, de forma direta o sistema, nem sempre linear, de n equações e n incógnitas dado por $\nabla f(x) = 0$. Normalmente, a solução é obtida por meio de um processo iterativo. Um processo iterativo consiste em considerar um ponto inicial x^0, obtemos um ponto mais próximo da solução x^1 e repetimos o processo gerando uma sequência $(x^k) \subset \mathbb{R}^n$ na qual a função decresce.

Basicamente temos três aspectos no que diz respeito aos métodos de otimização. O primeiro consiste na criação do algoritmo propriamente dito, que deve levar em conta a estrutura do problema e as propriedades satisfeitas pelas soluções, entre outras coisas.

O segundo aspecto se refere às sequências geradas pelo algoritmo, em que a principal questão é se tais sequências realmente convergem para uma solução do problema.

O terceiro ponto a ser considerado é a velocidade com que a sequência converge para uma solução. Para fins práticos, não basta que uma sequência seja convergente, é preciso que uma aproximação do limite possa ser obtida em um tempo razoável. Deste modo, bons algoritmos são os que geram sequências que convergem rapidamente para uma solução.

Vamos agora descrever um modelo geral de algoritmo para minimizar uma função em $\mathbb{R}^n$. No próximo capítulo estudaremos dois algoritmos específicos, analisando os aspectos mencionados acima.

2.1 ALGORITMOS DE DESCIDA

Uma forma geral de construir um algoritmo consiste em escolher, a partir de cada ponto obtido, uma direção para dar o próximo passo. Uma possibilidade é determinar uma direção em que f decresce.

Definição 2.1. Considere uma função $f : \mathbb{R}^n \to \mathbb{R}$, um ponto $\bar{x} \in \mathbb{R}^n$ e um vetor $d \in \mathbb{R}^n \backslash \{0\}$. Dizemos que d é uma *direção de descida* para

f, a partir de $\bar{x}$, quando existe $\delta \geq 0$ tal que $f(\bar{x} + td) < f(\bar{x})$, para todo $t \in (0, \delta)$.

Abaixo veremos uma definição e um teorema que serão importantes para este capítulo.

Definição 2.2. O limite

$$\frac{\partial f}{\partial d}(\bar{x}) = \lim_{t \to 0} \frac{f(\bar{x} + td) - f(\bar{x})}{t}$$

quando existe e é finito, denomina-se *derivada direcional* de f no ponto $(\bar{x})$ e na direção d.

Teorema 2.1. *Sejam a um ponto de acumulação de $X \subset \mathbb{R}^n$ e $f : X \to \mathbb{R}$ uma função real. Se $b = \lim_{x \to a} f(x)$ é um número positivo então existe $\delta > 0$ tal que $x \in X$ e $0 < |x - a| < \delta$ implicam $f(x) > 0$. O resultado é análogo para $b < 0$.*

Demonstração. Veja Lima [5, pp. 30]. ■

Agora estudaremos uma condição suficiente para uma direção ser de descida.

Teorema 2.2. *Se $\nabla f(\bar{x})^\top d < 0$, então d é uma direção de descida para f, a partir de $\bar{x}$.*

Demonstração. Sabemos que

$$\nabla f(\bar{x})^\top d = \frac{\partial f}{\partial d}(\bar{x}) = \lim_{t \to 0} \frac{f(\bar{x} + td) - f(\bar{x})}{t}$$

Pela hipótese e pela preservação do sinal vista no Teorema 2.1, existe $\delta > 0$ tal que

$$\frac{f(\bar{x} + td) - f(\bar{x})}{t} < 0$$

para todo $t \in (-\delta, \delta)$, $t \neq 0$. Portanto, $f(\bar{x} + td) < f(\bar{x})$, para todo $t \in (0, \delta)$, o que completa a demonstração. ■

Neste teorema há uma interpretação geométrica quando $n = 2$ ou $n = 3$, dizendo que as direções que formam um ângulo obtuso com $\nabla f(\bar{x})$ são de descida. Veja a Figura 6.

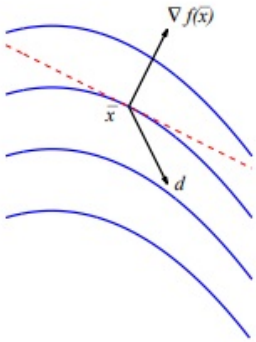

Figura 6 – Ilustração do Teorema 2.2.

Fonte – Elaborado por Ribeiro e Karas [11].

Usaremos o exemplo a seguir para ilustrar quando uma direção d específica é de descida.

Exemplo 2.1. Sejam $f : \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x) = \frac{1}{2}(x_1^2 - x_2^2)$ e $\bar{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Se $d = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$ é tal que $d_1 \leq 0$, então d é uma direção de descida para f, a partir de $\bar{x}$.

Com efeito temos $\nabla f(\bar{x})^\top d = d_1$. Caso $d_1 \leq 0$, podemos aplicar o Teorema 2.2 para concluir o que se pede. Entretanto, se $d_1 = 0$, não podemos usar o teorema, mas basta notar que $f(\bar{x}+td) = f([1, td_2]^\top) = f(\bar{x}) - \frac{(td_2)^2}{2} \leq f(\bar{x})$, para concluir que neste caso, d também é de descida.

O exemplo anterior nos mostra que nada se pode afirmar sobre uma direção d ser ou não de descida, a princípio, quando $\nabla f(\bar{x})^\top d = 0$.

Vamos apresentar agora um algoritmo básico para minimizar f.

Algoritmo 1 Algoritmo básico

Dado: $x^0 \in \mathbb{R}^n$
1: $k = 0$
2: **Enquanto** $\nabla f(x^k) \neq 0$ **Faça**
3: Calcule d^k tal que $\nabla f(x^k)^\top d^k < 0$
4: Escolha $t_k > 0$ tal que $f(x^k + t_k d^k) < f(x^k)$
5: Faça $x^{k+1} = x^k + t_k d^k$
6: $k = k + 1$
7: **Fim Enquanto**

Uma vez que usa direções de descida (Veja etapa 3), o algoritmo acima ou encontra um ponto estacionário em um número finito de iterações ou gera uma sequência ao longo da qual f decresce. Apenas precisamos saber se esta sequência tem algum ponto de acumulação e, em caso afirmativo, se este ponto é estacionário.

Deste modo se quisermos garantir sua convergência, a escolha da direção d^k e do tamanho de passo t_k, não poderá ser arbitrária. Por isso estudaremos métodos para encontrar tais direções e tamanhos de passo.

2.2 MÉTODO DA BUSCA LINEAR EXATA

Dada uma função $f : \mathbb{R}^n \to \mathbb{R}$, um ponto $\bar{x} \in \mathbb{R}^n$ e uma direção de descida $d \in \mathbb{R}^n$, queremos encontrar $\bar{t} > 0$ tal que

$$f(\bar{x} + \bar{t}d) < f(\bar{x}).$$

Precisamos balancear o tamanho do passo t com o decréscimo promovido em f, para que possamos tirar o máximo proveito a cada passo que tomamos, ou seja, achar o ponto que mais minimizou f nesta direção d. Veremos uma abordagem que consiste em fazer uma busca exata a partir do ponto $\bar{x}$ segundo a direção d.

O método consiste em minimizar unidimensionalmente f a par-

tir do ponto $\bar{x}$ na direção d. Basicamente devemos resolver o problema

$$\begin{aligned} \min_t \varphi(t) &:= f(\bar{x} + td) \\ &\text{sujeito à } t > 0 \end{aligned} \tag{Pt}$$

Em geral este problema pode ser difícil de se resolver. Entretanto para certas funções podem existir fórmulas fechadas, como no caso das funções quadráticas. Veremos adiante tais funções, bem como um algoritmo. Porém antes vamos fazer alguns exemplos que podem ser resolvidos de forma direta.

Exemplo 2.2. Considere $f : \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x) = \frac{1}{2}(x_1 - 2)^2 + (x_2 - 1)^2$, $\bar{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $d = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ e $\nabla f(\bar{x}) = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$. Faremos a busca linear exata a partir de $\bar{x}$, na direção d.

Primeiramente note que d é de fato uma direção de descida, pois

$$\nabla f(\bar{x})^\top d = \begin{bmatrix} -1 & -2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -2 < 0.$$

Para fazer a busca, considere

$$\varphi(t) := f(\bar{x} + td) = f(1, t) = \frac{1}{2} + t^2 - 2t + 1 = t^2 - 2t + \frac{3}{2},$$

cujo minimizador $\bar{t}$ satisfaz $\varphi'(\bar{t}) = 2t - 2 = 0$. Assim, como $\varphi(t) = 2t - 2$, temos

$$\bar{t} = 1 \text{ e } \bar{x} + \bar{t}d = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

A Figura 7 abaixo ilustra este exemplo achando $\bar{t}$ tal que $\bar{x} + \bar{t}d$ minimize unidimensionalmente f a partir do ponto $\bar{x}$ na direção d.

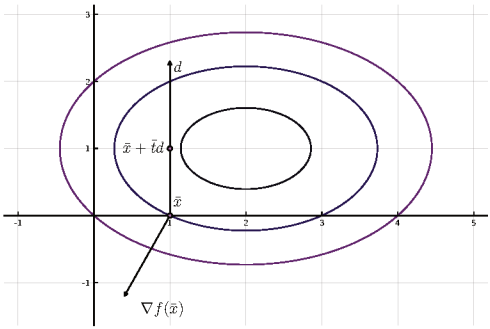

Figura 7 – Ilustração do Exemplo 2.2.

Fonte – Elaborado pelo autor (2019).

Exemplo 2.3. Considere $f : \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x) = \frac{1}{2}(x_1 - 2)^2 + (x_2 - 1)^2$, $\bar{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ e $d = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$. Faça a busca linear exata a partir de $\bar{x}$, na direção d.

Primeiramente note que d é de fato uma direção de descida, pois

$$\nabla f(\bar{x})^\top d = \begin{bmatrix} -1 & -2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = -5 < 0.$$

Para fazer a busca, considere

$$\begin{aligned} \varphi(t) &= f(\bar{x} + td) = f(1 + 3t, t) \\ &= \frac{1}{2}((1 + 3t) - 2)^2 + (t - 1)^2 \\ &= \frac{1}{2}(1 + 6t + 9t^2 - 4 - 12t + 4) + t^2 - 2t + 1 \\ &= \frac{9t^2}{2} - 3t + \frac{1}{2} + \frac{2t^2}{2} - 2t + 1 \\ &= \frac{11t^2}{2} - 5t + \frac{3}{2} \end{aligned}$$

cujo minimizador satisfaz $\varphi'(t) = 11t - 5 = 0$. Assim,

$$\bar{t} = \frac{5}{11} \text{ e } \bar{x} + \bar{t}d = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{5}{11} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1,36 \\ 0,45 \end{bmatrix} = \begin{bmatrix} 2,36 \\ 0,45 \end{bmatrix}$$

A Figura 8 abaixo ilustra este exemplo achando $\bar{t}$ tal que $\bar{x} + \bar{t}d$ minimize unidimensionalmente f a partir do ponto $\bar{x}$ na direção d.

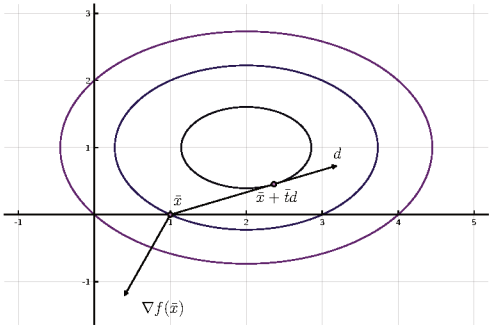

Figura 8 – Ilustração do Exemplo 2.3.

Fonte – Elaborado pelo autor (2019).

2.3 CONVERGÊNCIA GLOBAL DE ALGORITMOS DE DESCIDA

Nesta seção discutiremos a convergência global de algoritmos de descida. Primeiramente consideraremos o Algoritmo 1 com a direção definida por uma transformação do gradiente via matrizes definidas positivas.

Seja $H : \mathbb{R}^n \to \mathbb{R}^{n \times n}$ uma função contínua que associa a cada $x \in \mathbb{R}^n$ uma matriz definida positiva $H(x) \in \mathbb{R}^{n \times n}$. Assim, se $\nabla f(x) \neq 0$, temos que $d = -H(x)\nabla f(x)$ é uma direção de descida. De fato, $\nabla f(x)^\top d = -\nabla f(x)^\top H(x) \nabla f(x) < 0$.

Temos assim uma maneira de obter direções de descida para o Algoritmo 1. Para facilitar, vamos reescrever o algoritmo com esta escolha de direção de busca e com tamanho de passo feita pela busca linear exata.

Algoritmo 2 Algoritmo básico

Dado: $x^0 \in \mathbb{R}^n$
1: $k = 0$
2: **Enquanto** $\nabla f(x^k) \neq 0$ **Faça**
3: Defina $d^k = -H(x)\nabla f(x^k)$
4: Obtenha $t_k > 0$ tal que $f(x^k + t_k d^k) < f(x^k)$
5: Faça $x^{k+1} = x^k + t_k d^k$
6: $k = k + 1$
7: **Fim Enquanto**

Vamos analisar a convergência global do Algoritmo 2 utilizando a seguinte definição e os seguintes teoremas.

Definição 2.3. Um algoritmo é dito *globalmente convergente* quando para qualquer sequência (x^k) gerada pelo algoritmo e qualquer ponto de acumulação $\bar{x}$ de (x^k), temos que $\bar{x}$ é estacionário.

Teorema 2.3. *Seja $(x^k) \subset \mathbb{R}$ uma sequência monótona que possui uma subsequência convergente, então a sequência é convergente.*

Demonstração. Vamos supor (x^n) monótona não-crescente, se provarmos que é limitada então provamos que (x^n) converge.

Temos que (x^{nk}) é convergente, logo existe c tal que $|(x^{nk})| < c$. Como (x^n) é monótona não crescente, para todo $n \in \mathbb{N}$, existe $k \in \mathbb{N}$ tal que $n < nk$, de modo que $-c < (x^{nk}) \leq (x^n)$, logo (x^n) é limitada inferiormente. Para (x^n) monótona não-decrescente se prova de forma análoga. ■

Teorema 2.4. *Seja $f : \mathbb{R}^n \to \mathbb{R}$ diferenciável. Então o Algoritmo 2, com o tamanho do passo calculado pela busca linear exata, é globalmente convergente para um ponto estacionário de f.*

Demonstração. Seja (x^k) uma sequência gerada pelo algoritmo de descida e seja $\bar{x}$ um ponto de acumulação de (x^k), que existe por Bolzano-Weierstrass (veja Lima [7, pp. 96]), com $x^k \xrightarrow{\mathbb{N}'} \bar{x}$.

Suponha por absurdo que $\bar{x}$ não é estacionário, ou seja, $\nabla f(\bar{x}) \neq 0$. Dada a direção $d = -(H(x)\nabla f(x))$, sendo d uma direção de descida,

de fato, $\nabla f(x)^\top d = -(\nabla f(x)^\top H(x)\nabla f(x)) \leq 0$, logo o problema no passo 4 tem solução com $\bar{t} > 0$. Assim $f(\bar{x}) > f(\bar{x} + \bar{t}d)$ e consideremos $\beta \coloneqq f(\bar{x}) - f(\bar{x} + \bar{t}d)$.

Vamos considerar a função auxiliar $h : \mathbb{R}^n \to \mathbb{R}$ definida por

$$h(x) \coloneqq f(x) - f(x - \bar{t}H(x)\nabla f(x))$$

contínua, pois soma de funções contínuas é contínua. Então $h(x^k) \to h(\bar{x}) = \beta$, para a subsequência $(x^k)_{k\in\mathbb{N}'}$.

Por isso, dado $\epsilon > 0$ e para $k \in \mathbb{N}'$ suficientemente grande $|h(x^k) - h(\bar{x})| < \epsilon$ implica em

$$h(\bar{x}) - \epsilon < h(x^k) < h(\bar{x}) + \epsilon$$

e então tomando $\epsilon = \dfrac{\beta}{2}$ temos, $f(x^k) - f(x^k + \bar{t}d^k) \geq \dfrac{\beta}{2}$.

Logo para x^{k+1}, dado pelo Algoritmo 2 obtemos $f(x^{k+1}) = f(x^k + \bar{t}d^k) \leq f(x^k + t^k d^k)$ e como $f(x^k) - f(x^k + td^k) \geq \dfrac{\beta}{2}$, então $-f(x^k + td^k) \geq \dfrac{\beta}{2} - f(x^k)$, donde segue que $f(x^k + \bar{t}d^k) \leq f(x^k) - \dfrac{\beta}{2}$, ou seja, $f(x^{k+1}) \leq f(x^k) - \dfrac{\beta}{2}$. Portanto,

$$f(x^k) - f(x^{k+1}) \geq \frac{\beta}{2} > 0. \tag{8}$$

Por outro lado, temos que, pela continuidade de f, $f(x^k) \xrightarrow{k\in\mathbb{N}'} f(\bar{x})$. Temos também, pelo etapa 5 do Algoritmo 2, que $f(x^k), \forall k \in \mathbb{N}$ é decrescente. Logo, pelo Teorema 2.3, $f(x^k) \xrightarrow{k\in\mathbb{N}'} f(\bar{x})$, contradizendo (8).

Finalmente concluímos que $\bar{x}$ é estacionário. ■

Assim estudamos um algoritmo geral de descida. A seguir vamos explorar este algoritmo de forma mais específica.

3 MÉTODO DOS GRADIENTES

Ainda tentando resolver o problema (PG), estudaremos agora um dos mais famosos métodos de otimização: O método do gradiente, ou método de máxima descida.

Os matemáticos costumam atribuir o Método do Gradiente ao físico Peter Debye, que em 1909 o elaborou em um estudo assintótico das funções de Bessel. O próprio Debye observou que havia emprestado a idéia do método em um artigo de 1863 de Bernhard Riemann. O método também remonta a Cauchy e ao matemático russo Pavel Alexeevich Nekrasov (veja Petrova e Solov'ev [10]).

Tal método iterativo, em cada etapa, faz uma busca na direção d^k oposta ao vetor gradiente da função f em um dado ponto. Fazemos esta escolha pois, de todas as direções que f decresce, a direção oposta ao gradiente é a que f decresce mais rápido. De fato, se $d = -\nabla f(x)$ e $v \in \mathbb{R}^n$ é tal que $\|v\| = \|d\|$, então

$$\begin{aligned}\frac{\partial f}{\partial d}(x) = \nabla f(x)^\top d = -\|\nabla f(x)\|^2 = -\|\nabla f(x)\|\|v\| \\ \leq \nabla f(x)^\top v = \frac{\partial f}{\partial v}(x).\end{aligned}$$

O método do gradiente é definido pelo algoritmo iterativo

$$x^{k+1} = x^k + t_k d^k,$$

em que t_k é um escalar não-negativo que minimiza $\varphi(t) \coloneqq f(x^k + td^k)$, isto é, t_k é calculado usando busca linear exata.

Neste capítulo apenas abordaremos o caso quadrático, em que $f : \mathbb{R}^n \to \mathbb{R}$ é dada por

$$f(x) = \frac{1}{2}x^\top Ax - b^\top x + c \tag{9}$$

com $A \in \mathbb{R}^{n\times n}$ é simétrica definida positiva e $x \in \mathbb{R}^n$. O minimizador de f pode ser encontrado quando acha-se o ponto x^* tal que $\nabla f(x^*) = 0$, o que é equivalente a achar x^* que satisfaça, $Ax = b$.

Tal fato demonstramos no Teorema 1.10. E também pelo Teorema 1.10 temos que a direção oposta ao gradiente $d^k = -\nabla f(x^k) =$

$b - Ax^k$. Vamos construir o algoritmo utilizando o exemplo abaixo para nos auxiliar.

Exemplo 3.1. (Exemplo retirado de Shewchuk [12, pp. 02]) Suponha a função quadrática dada em (9) em que $A = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}$, $b = \begin{bmatrix} 2 \\ -8 \end{bmatrix}$ e $c = 0$, começando em $x^0 = [-2, -2]^\top$. Nosso primeiro passo na direção de máxima descida, ira cair em algum lugar na linha comprida cinza da Figura 9. Ou seja, nós cairemos no ponto

$$x^1 = x^0 + td^0$$

A questão é, quão grande o passo deve ser tomado?

Para isso usaremos a já comentada busca linear exata, que consiste em achar um t para minimizar φ. A Figura 10 ilustra este caso, em que estamos restritos a escolher um ponto na interseção do plano vertical com o paraboloide. A Figura 11 é a parábola definida por esta interseção. Percebemos que o t que estávamos buscando é o vértice dessa parábola. Pois do resultado conhecido do Cálculo de várias variáveis, veja em Guidorizzi [4, pp. 258], t minimiza φ quando a derivada direcional $\frac{d}{dt}f(x_t^1)$ é igual a zero. Pela regra da cadeia, temos que $\frac{d}{dt}f(x^1) = \nabla f(x^1)\frac{d}{dt}x^1 = \nabla f(x^1)^\top d^0$. Igualando a expressão a zero, descobrimos que t precisa ser escolhido tal que d^0 e $\nabla f(x^1)$ sejam ortogonais, veja a Figura 12.

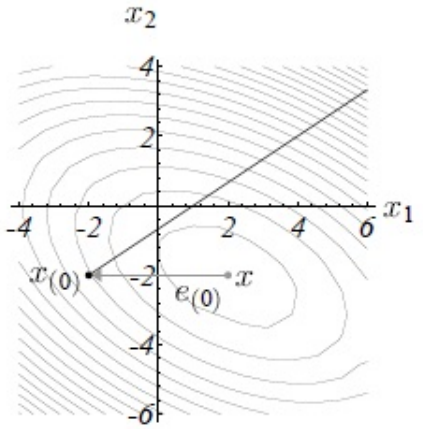

Figura 9 – Tomando o passo na direção de máxima descida de f.

Fonte – Elaborado por Shewchuk [12].

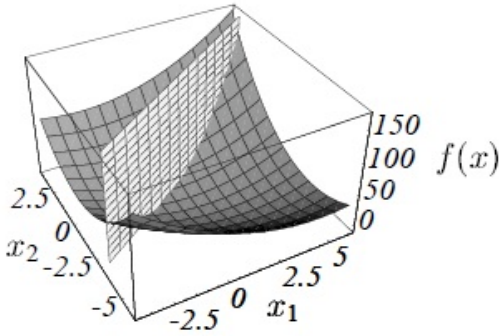

Figura 10 – Achando o ponto que minimiza φ.

Fonte – Elaborado por Shewchuk [12].

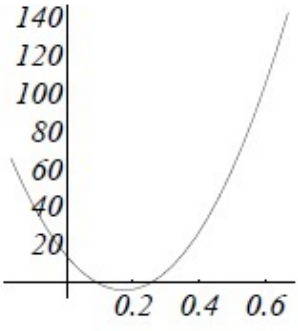

Figura 11 – Intersecção do plano vertical com o paraboloide.

Fonte – Elaborado por Shewchuk [12].

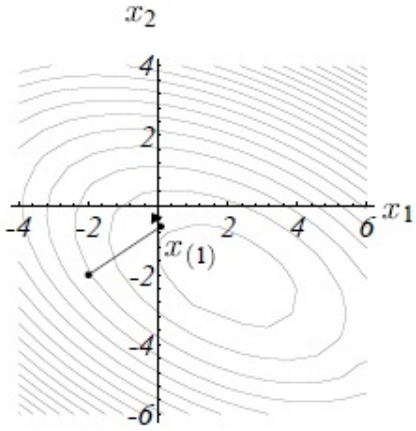

Figura 12 – Propriedade da ortogonalidade dos gradientes.

Fonte – Elaborado por Shewchuk [12].

Para finalizarmos nosso exemplo, precisamos determinar t_k, note que vamos usar d^k como sendo $-\nabla f(x^k)$.

Teorema 3.1. *Se t_k é obtido via busca linear exata, então* $(d_{k+1})^\top d_k = 0$.

Demonstração. Vamos definir uma função auxiliar, $\varphi : \mathbb{R} \to \mathbb{R}$ com $\varphi := f(x^k + td^k)$, temos então que, pela regra da cadeia, sua derivada é

$$\varphi'(t_k) = \nabla f(x^k + t_k d^k)^\top d^k = \nabla f(x^{k+1})^\top d^k.$$

Note que t^k é minimizador de φ, logo $\varphi'(t^k) = 0$.

Por outro lado,

$$\begin{aligned}\varphi'(t^k) &= \nabla f(x^{k+1})^\top d^k \\ &= (Ax^{k+1} - b)^\top (b - Ax^k) \\ &= (d^{k+1})^\top d^k.\end{aligned}$$

Logo, $(d^{k+1})^\top d^k = 0$, provando a ortogonalidade das direções. ■

Teorema 3.2. *O tamanho de passo t_k é determinado, usando busca linear exata, por:*

$$t_k = \frac{(d^k)^\top d^k}{(d^k)^\top A d^k}.$$

Demonstração. Já vimos que $\nabla f(x^k) = -d^k$, então teremos:

$$\begin{aligned}\nabla (d^{k+1})^\top d^k &= 0 \\ (b - Ax^{k+1})^\top d^k &= 0 \\ (b - A(x^k + t_k d^k)^\top d^k) &= 0 \\ (b - Ax^k)^\top d^k - t_k (Ad^k)^\top d^k &= 0 \\ (b - Ax^k)^\top d^k &= t_k (Ad^k)^\top d^k \\ (d^k)^\top d^k &= t_k (d^k)^\top A d^k \\ t_k &= \frac{(d^k)^\top d^k}{(d^k)^\top A d^k}.\end{aligned}$$

■

Unindo tudo que vimos até agora sobre o método do gradiente, temos:

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

$$t_k = \frac{(d^k)^\top d^k}{(d^k)^\top A d^k},$$

$$x^{k+1} = x^k + t_k d^k.$$

Resumindo tudo, temos o seguinte algoritmo.

Algoritmo 3 Método do gradiente

Dado: $x^0 \in \mathbb{R}^n$
1: $k = 0$
2: **Enquanto** $\nabla f(x^k) \neq 0$ **Faça**
3: Defina $d^k = -\nabla f(x^k)$
4: $t_k = \dfrac{(d^k)^\top d^k}{(d^k)^\top A d^k}$
5: Faça $x^{k+1} = x^k + t_k d^k$
6: $k = k + 1$
7: **Fim Enquanto**

Cabe salientar que este algoritmo é exatamente o Algoritmo 1 mas consideramos $H(x^k) = I \in \mathbb{R}^{n\times n}$, para todo $k \in \mathbb{N}$. Isto nos permite aplicar aqui a análise de convergência feita no Capítulo 1.

O Algoritmo 3, com o tamanho do passo t_k calculado pela busca linear exata, é globalmente convergente. Isto é uma consequência imediata do que foi estabelecido no Capítulo 2 e segue diretamente do Teorema 2.4, considerando $H(x) = I \in \mathbb{R}^{n\times n}$

3.1 INTERPRETAÇÃO GEOMÉTRICA

A Figura 13 mostra as 6 iterações do algoritmo para este caso. Esta figura deixa evidente a propriedade de que duas direções consecutivas são ortogonais que provamos no Teorema 3.1.

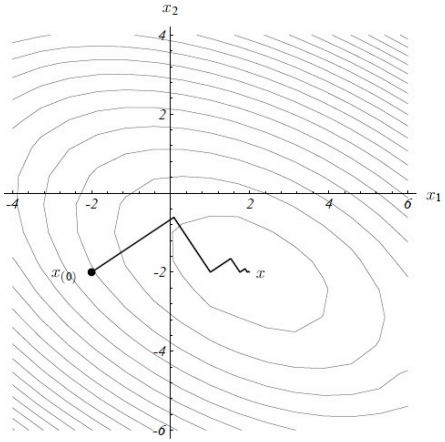

Figura 13 – Método do gradiente convergindo.

Fonte – Elaborado por Shewchuk [12].

Além da propriedade de que duas direções consecutivas são ortogonais já citada, este método possui outras propriedades interessantes. Exemplificaremos no caso em que A é 2×2. As curvas de nível de f são elipses cuja excentricidade depende do número de condição da matriz A, que é dado por $\sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}}$, em que $\lambda_{\max}$ é o maior autovalor da matriz A e $\lambda_{\min}$ o menor. Quanto maior o número de condição, mais excêntricas serão as elipses das curvas de nível.

O ponto de mínimo é encontrado mais rapidamente quando o número de condição da matriz for próximo de 1, logo o ideal é quando $\lambda_{\max} = \lambda_{\min}$, que é a situação da Figura 14, deixando as curvas de nível perfeitamente circuncêntricas. Neste caso achando o minimizador x^* em apenas um passo.

Porém, quando trabalhamos no caso oposto, ou seja, a matriz é mal condicionada, as curvas de nível serão bastante excêntricas, formando algo parecido com um cânion, e isto é um problema, pois, o algoritmo pulará de um lado para o outro, resultando numa demora agonizante para obtermos a solução (Veja a Figura 15). Este comportamento é denominado *zig-zag*.

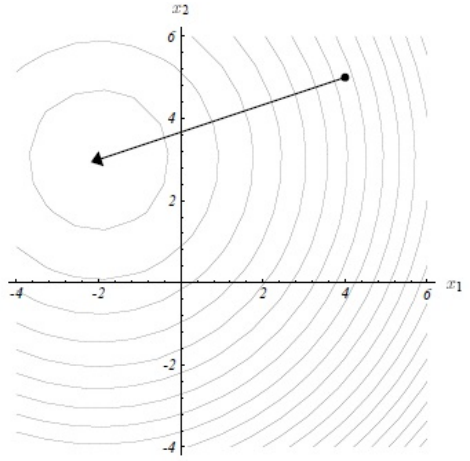

Figura 14 – Curvas de nível de uma matriz bem condicionada.

Fonte – Elaborado por Shewchuk [12].

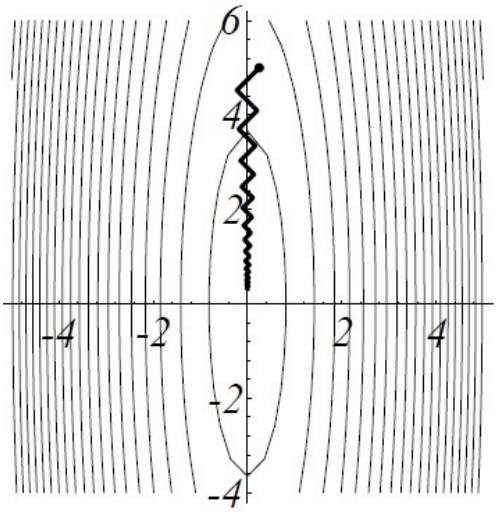

Figura 15 – Curvas de nível de uma matriz mal condicionada.

Fonte – Elaborado por Shewchuk [12].

Veremos a seguir um método que tenta obter convergência mais

rápida e mais barata computacionalmente falando, em comparação com o método do gradiente.

4 MÉTODO DE DIREÇÕES CONJUGADAS

Historicamente o desenvolvimento do trabalho original deste algoritmo foi realizado por diversos pesquisadores, incluindo Cornelius Lanczos e Magnus Hestenes no Instituto de Análise Numérica no Laboratóro Nacional de Matemática Aplicada, em Los Angeles nos Estados Unidos e Eduard Stiefel do Instituto Federal Suíço de Tecnologia em Zurich (veja Golub e O'Leary [2]).

Métodos de direções conjugadas são métodos de primeira ordem (usam apenas informações da função e do gradiente) e possuem, para resolver o problema (1), convergência mais rápida que o método do gradiente. Enquanto este pode gastar uma infinidade de passos para resolver tal problema, o método das direções conjugadas minimizam esta mesma quadrática usando no máximo n passos.

Para começarmos a entender o método, devemos antes nos atentar a alguns detalhes. Usualmente usamos o produto interno como

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i = y^\top x = x^\top y. \tag{10}$$

Porém agora teremos que ampliar este conceito. Dada qualquer matriz A, definida positiva, poderemos definir o produto interno induzido por A como,

$$\langle x, y\rangle_A = y^\top A x = x^\top A y. \tag{11}$$

Chamaremos o produto interno definido em (10) de produto interno usual, que é o produto interno induzido por I.

O produto interno induzido por A, simétrica definida positiva, possui as mesmas propriedades algébricas do produto interno usual. Em particular, $\langle x, x\rangle_A > 0$ se $x \neq 0$. Claramente também teremos mudanças na norma Euclidiana, ao invés de termos $\|x\|^2 = \sqrt{\langle x, x\rangle}$, usaremos o produto interno induzido por A para gerar uma norma diferente,

$$\|x\|_A^2 = \sqrt{\langle x, x\rangle_A}.$$

Relembrando que um método de descida resolve o sistema $Ax = b$ minimizando $f(x) = \frac{1}{2}x^\top Ax - b^\top x + c$. Usando nossa nova notação,

podemos reescrever f como,

$$f(y) = \frac{1}{2}\langle x, x\rangle_A - \langle b, x\rangle + c = \frac{1}{2}\|x\|_A^2 - \langle b, x\rangle + c.$$

Veremos a seguir algumas definições e teoremas que nos auxiliarão na construção do algoritmo dos gradientes conjugados.

Definição 4.1. Seja $A \in \mathbb{R}^{n\times n}$ uma matriz definida positiva. Dizemos que os vetores $d^0, d^1, \ldots, d^k \in \mathbb{R}^n\backslash\{0\}$ são *A-ortogonais* ou *A-conjugados* se

$$\left\langle d^i, d^j \right\rangle_A = 0$$

ou ainda

$$(d^i)^\top A d^j = 0,$$

para todo $i, j = 0, 1, \ldots, k$, com $i \neq j$.

Geometricamente podemos entender os vetores A-ortogonais, como vetores usualmente ortogonais caso as curvas de nível fossem circuncêntricas, como na Figura 16, porém após olhar para as curvas de nível formadas pela matriz A estes vetores ficam como na Figura 17.

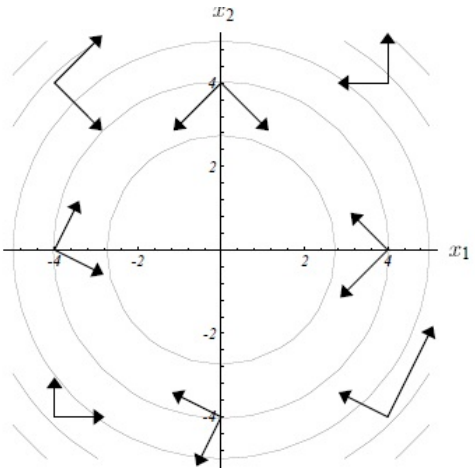

Figura 16 – Vetores usualmente ortogonais em curvas de nível circuncêntricas.

Fonte – Elaborado por Shewchuk [12].

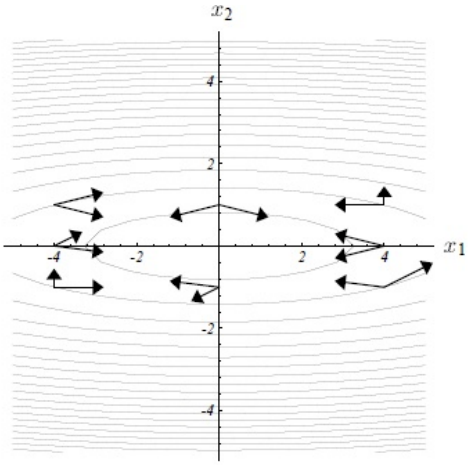

Figura 17 – Vetores A-ortogonais em curvas de nível formadas por A.

Fonte – Elaborado por Shewchuk [12].

Note que, no caso particular em que A é uma matriz identidade, vetores A-ortogonais são ortogonais no sentido usual. No caso geral, podemos provar a independência linear de vetores A-ortogonais.

Lema 4.1. *Seja $A \in \mathbb{R}^{n\times n}$ uma matriz definida positiva. Um conjunto qualquer de vetores A-ortogonais é linearmente independente.*

Demonstração. Sejam $\{d^0, d^1, \ldots, d^k\} \in \mathbb{R}^n$ vetores A-ortogonais e seja a combinação linear nula,

$$\alpha_0 d^0 + \alpha_1 d^1 + \cdots + \alpha_k d^k = 0, \text{ com } \alpha_j \in \mathbb{R}, j = 0, \ldots \ k.$$

Multiplicando ambos os membros da igualdade por $A(d^j)$, obtemos

$$(\alpha_0 d^0 + \alpha_1 d^1 + \cdots + \alpha_k d^k)^\top A d^j = 0^\top A d^j$$

abrindo as contas, temos

$$(\alpha_0 d^0)^\top A d^j + (\alpha_1 d^1)^\top A d^j + \cdots + (\alpha_j d^j)^\top A d^j + \cdots + (\alpha_k d^k)^\top A d^j = 0.$$

Como os vetores $\{d^0, d^1, \ldots, d^k\} \in \mathbb{R}^n$ são A-ortogonais, ficamos com apenas uma parcela

$$\alpha_j (d^j)^\top A d^j = 0.$$

Já que $(d^j)^\top A d^j > 0$, pois A é simétrica positiva definida, segue que

$$\alpha_j = 0.$$

Repetindo o argumento para todo $j = 1, \dots\ k$, obtemos que os vetores são linearmente independentes. ■

Podemos dizer que *a filosofia do método dos gradientes conjugados* consiste em que, dado um ponto inicial x^0, para a solução do Problema (PG) para $f(x)$ quadrática, toma-se um conjunto de direções $\{d_0, d_1, \dots, d_{k-1}\}$, em que a primeira direção é a mesma do Algoritmo 3 e todas as outras são A-conjugadas entre si.

Explicando melhor, é dado um ponto inicial x^0 e uma direção de descida, na qual será feita uma busca linear exata, para alcançarmos um novo ponto x^1. Após repetimos o processo até que seja atingida a solução desejada. Com efeito, no Teorema 4.3 ao final do capítulo, demonstraremos que se f é quadrática, então o algoritmo dá no máximo n iterações.

O Algoritmo 4 descreve o método de maneira geral.

Algoritmo 4 Método dos Gradientes Conjugados

Dado: $x^0 \in \mathbb{R}^n$, f quadrática.
1: $d^0 = -\nabla f(x^0)$
2: $k = 0$
3: **Enquanto** $\nabla f(x^k) \neq 0$ **Faça**
4: Encontre t_k usando busca linear exata
5: $x^{k+1} = x^k + t_k d^k$
6: Encontre d^{k+1} A-conjugado com $\{d^0, d^1, \dots, d^k\}$
7: $k = k + 1$
8: **Fim Enquanto**

Exploraremos agora cada etapa do Algoritmo 4. Começaremos por descrever as direções utilizadas na etapa 6 do Algoritmo 4. Nossa primeira direção será dada pela mesma direção do método do gradiente, ou seja, $d_0 = -\nabla f(x^0)$. A partir desta direção, acharemos uma próxima direção da seguinte maneira:

dado $x^0 \in \mathbb{R}^n$, defina $d^0 = -\nabla f(x^0)$ e, para $k = 0, 1, \ldots, n-2$,

$$d^{k+1} \coloneqq -\nabla f(x^{k+1}) + \beta_k d^k, \tag{12}$$

em que, $x^{k+1} = x^k + t_k d^k$ e β^k deve ser calculado de modo que d^k e d^{k+1} sejam A-ortogonais, ou seja,

$$(d^k)^\top A d^{k+1} = (d^k)^\top A(-\nabla f(x^{k+1}) + \beta_k d^k) = 0.$$

Daí,

$$\begin{aligned}(d^k)^\top A(-\nabla f(x^{k+1})) + (d^k)^\top A\beta_k d^k &= 0\\ -(d^k)^\top A\nabla f(x^{k+1}) &= -\beta_k (d^k)^\top A d^k\\ \beta_k &= \frac{d^k)^\top A\nabla f(x^{k+1}}{d^k)^\top A d^k}.\end{aligned}$$

Os próximos resultados estabelecerão que as direções geradas pelo algoritmo são, de fato, A-ortogonais.

Lema 4.2. *Dado $x^0 \in \mathbb{R}^n$, considere a sequência finita $x^{k+1} = x^k + t_k d^k$. Então com d^k e t_k dado pelo Algoritmo 5, temos*

$$\nabla f(x^k)^\top d^j = 0,$$

para todo $j = 0, 1, \ldots, k-1$.

Demonstração. Provaremos usando indução em k. Para $k = 1$, temos $\nabla f(x^1)^\top d^0 = \nabla f(x^0 + t_0 d^0)^\top d^0 = \varphi'(t_0) = 0$. Por hipótese de indução temos $\nabla f(x^{k-1})^\top d^j = 0$, para todo $j = 0, 1, \ldots, k-2$. Então para $j \in 0, 1, \ldots, k-2$ temos que,

$$\begin{aligned}\nabla f(x^k)^\top d^j &= (A(x^{k-1} + t_{k-1} d^{k-1}) - b)^\top d^j\\ &= (Ax^{k-1} - b + t_{k-1} d^{k-1})^\top d^j\\ &= \nabla f(x^{k-1})^\top d^j + t_{k-1}(d^{k-1})^\top A d^j\\ &= 0 + 0\\ &= 0\end{aligned}$$

Logo, $\nabla f(x^k)^\top d^j = 0$ ■

Definiremos agora, um subespaço vetorial que está intimamente relacionado às direções geradas pelo método de Gradientes Conjugados.

Definição 4.2. Dados $A \in \mathbb{R}^{n\times n}$, $y \in \mathbb{R}^n$ e $k \in \mathbb{N}$, definimos o *r-ésimo espaço de Krylov* gerado por A e y como sendo o subespaço vetorial dado por

$$\mathcal{K}_r(A, y) \coloneqq \text{span}\{y, Ay, A^2y, \ldots, A^{r-1}y\}.$$

O próximo teorema relaciona o espaço gerado pelos gradientes $\nabla f(x^k)$ e o espaço gerado pelas direções d^k, obtidos pelo algoritmo de gradientes conjugados, com um espaço de Krylov específico, gerado à partir de $\nabla f(x^0)$.

Teorema 4.1. *Após j passos do algoritmo dos gradientes conjugados (com $\nabla f(x^k) \neq 0$ em todos os passos), temos que*

$$\text{span}\{d^0, d^1, \ldots, d^j\} = \text{span}\{\nabla f(x^0), \nabla f(x^1), \ldots, \nabla f(x^j)\} = \\ = \mathcal{K}_{j+1}(A, \nabla f(x^0)).$$

Demonstração. Provaremos por indução em j. Para $k = 0$, é trivial, pois $d^0 = -\nabla f(x^0)$. Assumiremos que os espaços são iguais para $j = k$, e mostraremos que a igualdade permanece verdade para $j = k + 1$. Primeiramente mostraremos que

$$\text{span}\{\nabla f(x^0), \nabla f(x^1), \ldots, \nabla f(x^{k+1})\} \subseteq \mathcal{K}_{k+2}(A, \nabla f(x^0)). \qquad (13)$$

Tendo em vista a hipótese de indução, é suficiente mostrar que $\nabla f(x^{k+1}) \in \mathcal{K}_{k+2}(A, \nabla f(x^0))$. Relembrando que $\nabla f(x^{k+1}) = \nabla f(x^k) - t_k A d^k$, basta checarmos Ad^k. Por hipótese de indução, $d^k \in \mathcal{K}_{k+1}(A, \nabla f(x^0)) = \text{span}\{\nabla f(x^0), A\nabla f(x^0), \ldots, A^k \nabla f(x^0)\}$. Então,

$$Ad^k \in \text{span}\{A\nabla f(x^0), A^2\nabla f(x^0), \ldots, A^{k+1}\nabla f(x^0)\} \\ \subseteq \mathcal{K}_{k+2}(A, \nabla f(x^0)).$$

Ademais, $\nabla f(x^k) \in \mathcal{K}_{k+1}(A, \nabla f(x^0)) \subseteq \mathcal{K}_{k+2}(A, \nabla f(x^0))$, então

$$\nabla f(x^{k+1}) = \nabla f(x^k) - t_k A d^k \in \mathcal{K}_{k+2}(A, \nabla f(x^0)).$$

Isto conclui (13).

Em nosso próximo passo vamos provar

$$\text{span}\{d^0, d^1, \ldots, d^{k+1}\} \subseteq \text{span}\{\nabla f(x^0), \nabla f(x^1), \ldots, \nabla f(x^{k+1})\}. \quad (14)$$

Por hipótese de indução,

$$d^i \in \text{span}\{\nabla f(x^0), \nabla f(x^1), \ldots, \nabla f(x^k)\}, \text{ para } i = 0, 1, \ldots, k, \quad (15)$$

então é suficiente mostrar que,

$$d^{k+1} \in \text{span}\{\nabla f(x^0), \nabla f(x^1), \ldots, \nabla f(x^{k+1})\}.$$

Mas isto segue imediatamente de (12), usando (15) novamente.

Juntando (13) e (14), podemos ver que os três espaços estão entrelaçados. Nós podemos mostrar isto, demonstrando que todos possuem a mesma dimensão. Como $\mathcal{K}_{k+2}(A, \nabla f(x^0))$ é formado por $k+2$ vetores, sua dimensão é $k+2$. Se pudermos demonstrar que a dimensão de $\text{span}\{d^0, d^1, \ldots, d^{k+1}\}$ é exatamente $k+2$, estaremos convencidos que os três espaços possuem dimensão $k+2$ e portanto são iguais. Mas já sabemos,pelo Lema 4.1, que $d^0, d^1, \ldots, d^{k+1}$ são linearmente independentes, logo eles formam uma base para $\text{span}\{d^0, d^1, \ldots, d^{k+1}\}$ que possui dimensão $k+2$, o que finaliza a demonstração. ■

Corolário 4.1. *Se x^k e d^k foram gerados pelo Algoritmo 5, então*

$$(d^k)^\top A d^j = 0, \forall j = 0, 1, 2, \ldots, k-1.$$

Demonstração. Provaremos por indução. É fácil ver que o teorema é verdade para $k = 1$. Supondo que seja verdade para k, provaremos para $k+1$. Temos que

$$d_{k+1}^\top A d_i = -\nabla f(x^{k+1})^\top A d_i + \beta_k d_k^\top A d_i.$$

Para $i = k$ o lado direito zera por definição de β_k. Para $i < k$ ambos os termos zeram. De fato, o primeiro termo se anula pois $Ad_i \in \text{span}\{d_1, d_2, \ldots, d_{i+1}\}$. Com efeito, temos que pela hipótese de indução que o método é um método de direções conjugadas a partir

de x^{k+1} e pelo Lema 4.2, que garante que $\nabla f(x^{k+1})$ é ortogonal a $\text{span}\{d_0, d_1, \ldots, d_{i+1}\}$.

O segundo termo zera pela hipótese de indução, o que completa a demonstração. ■

Agora, podemos mostrar uma fórmula fechada para o tamanho do passo.

Teorema 4.2. *O* tamanho de passo t_k *da etapa* 4 *do Algoritmo 4 é dado por,*

$$t_k = \frac{(b - Ax^k)^\top d^k}{(d^k)^\top A d^k}.$$

Demonstração. Como estamos construindo o método para quadráticas, seja a função $f : \mathbb{R}^n \to \mathbb{R}$ como $f(x) = \frac{1}{2}x^\top Ax - b^\top x + c$, definiremos também a função auxiliar $\varphi(t, y) = (x^0 + P_k y + td^k)$, em que a matriz $P_k = \begin{bmatrix} d^0 & d^1 & \ldots & d^{k-1} \end{bmatrix}$ e o vetor $y = \begin{bmatrix} t_0 \\ t_1 \\ \ldots \\ t_{k-1} \end{bmatrix}$.

Tendo em vista que, pela etapa 5 do Algoritmo 4, $x^{k+1} = x^k + t_k d^k$, temos que

$$\begin{aligned} x^k &= x^0 + t_0 d^0 + t_1 d^1 + \cdots + t_{k-1} d^{k-1} \\ &= x^0 + P_k y. \end{aligned}$$

Com isto, podemos ver que

$$\begin{aligned}
\varphi(t,y) &= f(x^0 + P_k y + td^k) \\
&= \frac{1}{2}(((x^0 + P_k y) + td^k))^\top A((x^0 + P_k y) + td^k) - \\
&\qquad - b^\top((x^0 + P_k y) + td^k) + c \\
&= \frac{1}{2}(x^0 + P_k y)^\top A(x^0 + P_k y) - b^\top(x^0 + P_k y) + c + \\
&\qquad + \frac{2t}{2}(x^0 + P_k y)^\top Ad^k + \frac{1}{2}(t^2(d^k)^\top Ad^k) - \\
&\qquad - t(b^\top d^k) \\
&= f(x^0 + P_k y) + t(x^0 + P_k y)^\top Ad^k + \frac{1}{2}(t^2(d^k)^\top Ad^k) - t(b^\top d^k) \\
&= f(x^0 + P_k y) + t(Ax^0 + AP_k y)^\top d^k - \\
&\qquad - t((b^k)^\top d^k) + \frac{1}{2}(t^2(d^k)^\top Ad^k) \\
&= f(x^0 + P_k y) + t(AP_k y)^\top d^k + t(Ax^0 - b)^\top d^k + \\
&\qquad + \frac{1}{2}(t^2(d^k)^\top Ad^k).
\end{aligned}$$

Provaremos que $(AP_k y)^\top d^k = 0$. Temos que $(P_k y)^\top Ad^k = \langle P_k y, d^k \rangle_A$ e portanto

$$\langle P_k y, d^k \rangle_A = y^\top P_k^\top Ad^k = y^\top \begin{bmatrix} (d^0)^\top \\ (d^1)^\top \\ \vdots \\ (d^{k-1})^\top \end{bmatrix} Ad^k = y^\top \begin{bmatrix} \langle d^0, d^k \rangle_A \\ \langle d^1, d^k \rangle_A \\ \vdots \\ \langle d^{k-1}, d^k \rangle_A \end{bmatrix}.$$

Pelo Corolário 4.1, as direções geradas são A-conjugadas. Daí, o vetor

$$\begin{bmatrix} \langle d^0, d^k \rangle_A \\ \langle d^1, d^k \rangle_A \\ \vdots \\ \langle d^{k-1}, d^k \rangle_A \end{bmatrix}$$

se anula. Logo, $(AP_k y)^\top d^k = 0$. Com isso,

$$\varphi(t,y) = f(x^0 + P_k y) + \frac{1}{2}(t^2 d^k Ad^k) + t(Ax^0 - b)^\top d^k$$

portanto,

$$\min_{y,t} \varphi(t,y) = \min_{y} f(x^0 + P_k y) + \min_{t} \frac{1}{2}(t^2 d^k A d^k) + t(Ax^0 - b)^\top d^k.$$

Assim,

$$t_k \in \arg\min \frac{1}{2}(t^2 d^k A d^k) + t(Ax^0 - b)^\top d^k,$$

e então,

$$t_k = \frac{(b - Ax^0)^\top d^k}{(d^k)^\top A d^k},$$

pois trata-se de minimizar uma quadrática unidimensional em t. ■

Agora, juntando tudo que provamos, apresentamos o Algoritmo 5, que descreve o algoritmo de gradientes conjugados propriamente dito, com fórmulas explícitas para t_k e β_k.

Algoritmo 5 Método de gradientes conjugados II

Dado: $x^0 \in \mathbb{R}^n$, faça $d^0 = -\nabla f(x^0)$
1: $k = 0$
2: **Enquanto** $\nabla f(x^k) \neq 0$ **Faça**
3: $\quad t_k = -\dfrac{\nabla f(x^k)^\top d^k}{(d^k)^\top A d^k}$
4: $\quad x^{k+1} = x^k + t_k d^k$
5: $\quad \beta_k = \dfrac{(d^k)^\top A \nabla f(x^{k+1})}{(d^k)^\top A d^k}$
6: $\quad d^{k+1} = -\nabla f(x^{k+1}) + \beta_k d^k$
7: $\quad k = k + 1$
8: **Fim Enquanto**

Finalmente, o teorema a seguir mostra que o Algoritmo 5 minimiza uma quadrática com no máximo n passos, isto é, o método de gradientes conjugados para uma quadrática tem convergência finita.

Teorema 4.3. *Considere a função quadrática* $f(x) = \frac{1}{2}x^\top Ax - b^\top x + c$ *e seu minimizador* x^*. *Dado* $x^0 \in \mathbb{R}^n$, *então a sequência gerada pelo Algoritmo 5 cumpre* $x^n = x^*$.

Demonstração. Pelo Lema 4.1, o conjunto $\{d^0, d^1, \dots\ d^{n-1}\}$ é um conjunto de n vetores linearmente independentes, logo é uma base de $\mathbb{R}^n$. Portanto, existem escalares $\alpha_j \in \mathbb{R}$, $j = 0, 1, \dots, n-1$, tais que

$$x^* - x^0 = \sum_{j=0}^{n-1} \alpha_j d^j. \tag{16}$$

Consideremos $k \in \{0, 1, \dots, n-1\}$ arbitrário. Multiplicando a Relação (16) por $(d^k)^\top A$ e levando em conta que as direções são A-ortogonais, temos que

$$(d^k)^\top A(x^* - x^0) = \alpha_k (d^k)^\top A d^k.$$

Assim,

$$\alpha_k = \frac{(d^k)^\top A(x^* - x^0)}{(d^k) A d^k} \tag{17}$$

Por outro lado, novamente, pelo processo iterativo, temos que

$$x^k = x^0 + t_0 d^0 + t_1 d^1 + \cdots + t_{k-1} d^{k-1},$$

que multiplicando por $(d^k)^\top A$, implica

$$(d^k)^\top A x^k = (d^k)^\top A x^0, \tag{18}$$

pois as direções são A-ortogonais. Substituindo (18) em (17) e usando o Teorema 1.10, obtemos

$$\begin{aligned}
\alpha_k &= \frac{(d^k)^\top A x^* - (d^k)^\top A x^0}{(d^k)^\top A x^k} \\
&= \frac{(d^k)^\top A x^* - (d^k)^\top A x^k}{(d^k)^\top A x^k} \\
&= \frac{(d^k)^\top b - (d^k)^\top A x^k}{(d^k)^\top A x^k} \\
&= \frac{(d^k)^\top (b - A x^k)}{(d^k)^\top A x^k} \\
&= \frac{(d^k)^\top (-\nabla f(x^k)}{(d^k)^\top A x^k} \\
&= -\frac{(d^k)^\top \nabla f(x^k)}{(d^k)^\top A x^k} \\
&= t_k.
\end{aligned}$$

Portanto, de (16) segue que

$$x^* = x^0 + \sum_{j=0}^{n-1} t_j d^j = x^0 + t_0 d^0 + t_1 d^1 + \cdots + t_{n-1} d^{n-1} = x^n,$$

completando a demonstração. ■

Agora finalizaremos este capítulo retomando o Exemplo 3.1, porém agora utilizaremos o método dos gradientes conjugados para resolvê-lo.

Exemplo 4.1. Suponha a função quadrática dada em (9) em que $A = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}$, $b = \begin{bmatrix} 2 \\ -8 \end{bmatrix}$ e $c = 0$ começando em $x^0 = [-2, -2]^\top$. Mostraremos que o método converge em dois passos.

Primeiramente vamos descobrir nossa função.

$$\frac{1}{2} x^\top A x - b^\top x = \frac{x_1}{2}(3x_1 + 2x_2) + \frac{x_2}{2}(2x_1 + 6x_2) - (2x_1 + 8x_2),$$

que possui gradiente igual à

$$\nabla f(x) = \begin{bmatrix} 3x_1 + 2x_2 - 2 \\ 2x_1 + 6x_2 + 8 \end{bmatrix}$$

portanto, no ponto x^0, temos

$$\nabla f(x^0) = \begin{bmatrix} -12 \\ -8 \end{bmatrix}$$

Como $d^0 = -\nabla f(x^0)$ então, $d^0 = \begin{bmatrix} 12 \\ 8 \end{bmatrix}$ Possuindo todas estas informações já podemos encontrar nosso t_0.

$$t_0 = \frac{-\nabla f(x^0)^\top d^0}{(d^0)^\top A d^0} = \frac{\begin{bmatrix} 12 & 8 \end{bmatrix} \begin{bmatrix} 12 \\ 8 \end{bmatrix}}{\begin{bmatrix} 12 & 8 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} 12 \\ 8 \end{bmatrix}} = \frac{13}{75}.$$

E assim podemos achar nosso próximo ponto x^1.

$$x^1 = x^0 + t_0 d^0 = \begin{bmatrix} -2 \\ -2 \end{bmatrix} + \frac{13}{75} \begin{bmatrix} 12 \\ 8 \end{bmatrix} = \begin{bmatrix} 0,08 \\ -0,613 \end{bmatrix},$$

cujo gradiente é,

$$\nabla f(x^1) = \begin{bmatrix} -2,986 \\ 4,48 \end{bmatrix}.$$

Assim obteremos nosso β_0,

$$\beta_0 = \frac{(d^0)^\top A \nabla f(x^1}{(d^0)^\top A d^0} = \frac{\begin{bmatrix} 12 & 8 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} -2,986 \\ 4,48 \end{bmatrix}}{\begin{bmatrix} 12 & 8 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} 12 \\ 8 \end{bmatrix}} = 0,139.$$

Assim nossa próxima direção pode ser obtida através da seguinte conta,

$$d^1 = -\nabla f(x^1) + \beta_0 d^0 = \begin{bmatrix} 2,986 \\ -4,48 \end{bmatrix} + 0,139 \begin{bmatrix} 12 \\ 8 \end{bmatrix} = \begin{bmatrix} 4,654 \\ -3,368 \end{bmatrix}.$$

Com a nova direção em mãos recomeçamos o algoritmo calculando t_1,

$$t_1 = \frac{-\nabla f(x^1)^\top d^1}{(d^1)^\top A d^1} = \frac{\begin{bmatrix} 2,986 & -4,48 \end{bmatrix} \begin{bmatrix} 4,654 \\ -3,368 \end{bmatrix}}{\begin{bmatrix} 4,654 & -3,368 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} 4,654 \\ -3,368 \end{bmatrix}} = 0,4121.$$

Assim, finalmente podemos encontrar nosso x^2, que a princípio deve ser nossa solução.

$$x^2 = x^1 + t_1 d^1 = \begin{bmatrix} 0,08 \\ -0,613 \end{bmatrix} + 0,4121 \begin{bmatrix} 4,654 \\ -3,368 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix},$$

cujo gradiente é,

$$\nabla f(x^2) = \begin{bmatrix} 3 \times 2 + (2 \times -2) - 2 \\ 2 \times 2 + (6 \times -2) + 8 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

finalizando nosso algoritmo em 2 passos, como previsto.

A Figura 18 ilustra este exemplo mostrando a aplicação do algoritmo de gradientes conjugados para minimização de uma função quadrática em $\mathbb{R}^2$ com $x^0 = [-2, -2]^\top$ e convergindo em $[2, -2]^\top$. Observe que obtemos a solução em dois passos como visto no Teorema 4.3

e também conseguimos driblar o problema do *zig-zag*. Lembrando que este exemplo é o mesmo utilizado no método do gradiente, na Figura 13, naquele método demorando 6 iterações.

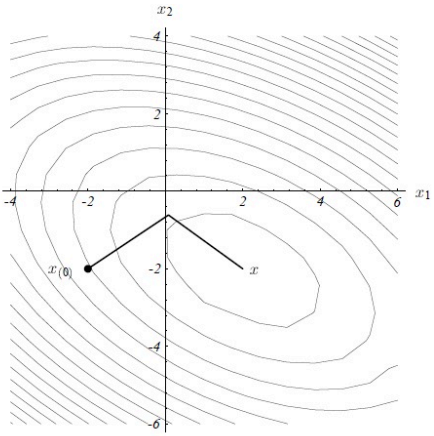

Figura 18 – Método de gradientes conjugados em ação.

Fonte – Elaborado por Shewchuk [12].

CONSIDERAÇÕES FINAIS

Neste trabalho estudamos o Método dos Gradientes Conjugados aplicado a solução de uma função quadrática.

No Capítulo 1, estudamos condições de otimalidade gerais de um problema de otimização, para sabermos como caracterizar um minimizador e como a convexidade nos auxilia em determinados casos.

Já no Capítulo 2 vimos, de maneira geral, o que é um método iterativo e como funciona um algoritmo de descida. Vimos também maneiras de encontrar caminhos para que a solução do Problema (PG) fosse encontrada.

No Capítulo 3 apresentamos o Método do Gradiente, que é um método no qual tomamos a direção oposta ao gradiente para minimizar as funções. Porém conforme vimos este método, em determinados casos, não é muito eficiente, por causa do efeito *zig-zag*.

No Capítulo 4, apresentamos o Método de Gradientes Conjugados, que usa direções conjugadas, e por isso em no máximo n passos encontra a solução do Problema de minimização quadrática.

Finalmente este trabalho possibilitou contato com a Otimização, subárea da matemática que não está presente no currículo da graduação em Licenciatura em Matemática da UFSC- Blumenau, apresentando novas oportunidades de pesquisas futuras. Também é importante enfatizar que os conteúdos de Álgebra Linear aprendidos no decorrer do curso foram aprofundados ao longo do desenvolvimento deste trabalho de conclusão de curso.

REFERÊNCIAS

[1] J. Bezanson, A. Edelman, S. Karpinski e V. B. Shah. "Julia: A Fresh Approach to Numerical Computing". *In: Siam Review* 1 (2017). DOI: 10.1137/141000671.

[2] G. H. Golub e D. P. O'Leary. "Some History of the Conjugate Gradient and Lanczos Algorithms: 1948-1976". *In: SIAM Review* 31.1 (1989), pp. 50–102. ISSN: 00361445.

[3] H. Guidorizzi. *Um curso de cálculo*. Vol. 1. São Paulo: LTC, 2000. ISBN: 9788521622444.

[4] H. Guidorizzi. *Um curso de cálculo*. Vol. 2. São Paulo: LTC, 2001. ISBN: 9788521622451.

[5] E. Lima. *Análise Real*. Vol. 2. Rio de Janeiro: IMPA, 1997.

[6] E. Lima. *Curso de análise*. Vol. 1. Rio de Janeiro: IMPA, 1992.

[7] E. Lima. *Curso de análise*. Vol. 2. Rio de Janeiro: IMPA, 1999.

[8] D. Luenberger e Y. Ye. *Linear and Nonlinear Programming*. International Series in Operations Research & Management Science. Springer US, 2008. ISBN: 9780387745022.

[9] C. Meyer. *Matrix Analysis and Applied Linear Algebra*. Other Titles in Applied Mathematics. Philadelphia: Society for Industrial e Applied Mathematics, 2000. ISBN: 9780898714548.

[10] S. S. Petrova e A. D. Solov'ev. "The Origin of the Method of Steepest Descent". *In: Historia Mathematica* 24.4 (1997), pp. 361–375. ISSN: 0315-0860. DOI: https://doi.org/10.1006/hmat.1996.2146.

[11] A. Ribeiro e E. Karas. *Otimização Contínua: Aspectos Teóricos e Computacionais*. São Paulo: CENGAGE DO BRASIL, 2013. ISBN: 9788522115013.

[12] J. R. Shewchuk. *An Introduction to the Conjugate Gradient Method Without the Agonizing Pain*. Rel. técn. Pittsburgh, PA, USA, 1994.

[13] D. Watkins. *Fundamentals of Matrix Computations*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2010. ISBN: 9780470528334.