UNIVERSIDADE FEDERAL DO PARANÁ

TUANNY ELYZ BRANDELEIRO BRUFATI

MÉTODO DE GRADIENTES CONJUGADOS

CURITIBA
2011

TUANNY ELYZ BRANDELEIRO BRUFATI

MÉTODO DE GRADIENTES CONJUGADOS

Monografia apresentada como requisito parcial à obtenção do grau de Licenciado em Matemática, pelo Departamento de Matemática, Setor de Ciências Exatas, Universidade Federal do Paraná.

Orientadora:

Prof[a]. Dr[a]. Elizabeth Wegner Karas.

Co-orientador:

Prof. Dr. Ademir Alves Ribeiro.

Curitiba

2011

Termo de Aprovação

Monografia apresentada como requisito parcial à obtenção do grau de Licenciado em Matemática, pelo Departamento de Matemática, Setor de Ciências Exatas, Universidade Federal do Paraná, pela seguinte banca examinadora:

Orientadora: Prof[a] Dra Elizabeth Wegner Karas
Universidade Federal do Paraná

Co-orientador: Prof. Dr. Ademir Alves Ribeiro
Universidade Federal do Paraná

Membro da banca: Prof. Dr. Higidio Portillo Oquendo
Universidade Federal do Paraná

Curitiba, 17 de junho de 2011

Agradecimentos

Agradeço a Deus, pelo dom da vida e por me dar forças para suceder nas minhas realizações.

Agradeço à família, Vilmar, Mary e Renan, por estarem sempre me dando amor, carinho e apoio, mesmo distante.

Agradeço à professora Elizabeth, por toda a paciência, dedicação, carinho e afeto, pelos "puxões de orelha", "pitos", nervosismo produtivo e por ter sido um pouco minha mãe também. Agradeço por ter percebido os meus anseios em aprender, por poder me desafiar e dar o suporte suficiente para que eu pudesse aprender coisas novas e criar autonomia. E ainda, por ter sido um espelho na minha vida em muitos aspectos.

Agradeço ao professor Ademir por toda paciência, carisma e por todas as horas que ele dedicou nesse período de sua vida para me ensinar.

Agradeço profundamente aos meus orientadores, Elizabeth, Ademir e Lucas, por terem me proposto situações desafiadoras nas quais, com o apoio deles e de meus amigos, pude desempenhar um papel, diga-se de passagem, brilhante.

Agradeço aos meus amigos e colegas de Iniciação Científica, Beatriz, Flávia, Karla, Leonardo e Wagner, por todo o conhecimento e cultura que dividimos ao longo das nossas reuniões e conversas. Além disso, por algumas risadas, "troca de figurinhas", aprendizados, viagens e experiência.

Agradeço às pessoinhas lindas e amadas que moram e moraram comigo, Aline, Bárbara, Giselle, Marceli, Marta, Talita, Viviam e Yannaê, por terem me oferecido suporte emocional, amor, carinho, sorrisos, abraços e porque foram e são parte de minha família ao longo desses anos de convivência.

Agradeço a todas as pessoas que permitiram a realização deste trabalho, àquelas que continuam acreditando nesse trabalho e àquelas que possibilitaram a sua continuidade.

Agradeço aos colegas e companheiros de trabalho que acreditaram que eu era uma pessoa capaz, por me ensinar, cobrar, questionar, ajudarem a conduzir meus objetivos de forma que eu tivesse ainda mais força para alcançá-los e por abrir novas portas para o meu Sonho, que pude transformar em Objetivo, que é ser uma boa Matemática.

Resumo

Neste trabalho discutimos o Método dos Gradientes Conjugados e algumas variantes do método para minimizar uma função real de várias variáveis. O principal objetivo é obter uma convergência mais rápida que o método de Cauchy e reduzir o custo computacional com relação ao Método de Newton, por não fazer uso de derivadas segundas. A fundamentação teórica dos métodos de gradientes conjugados reside no estudo de direções A-conjugadas, onde A é uma matriz quadrada de dimensão n. É importante ressaltar que as propriedades de convergência do método de Gradientes Conjugados dependem da eficácia da busca unidirecional. Além disso, provamos que os métodos de direções conjugadas minimizam uma função quadrática definida em $\mathbb{R}^n$ em no máximo n iterações com taxa de convergência ótima. Para a análise da complexidade algorítmica estudamos os Polinômios de Chebyshev. Mostramos também que as variantes dos métodos de gradientes conjugados propostas por Polak e Ribière e por Fletcher e Reeves coincidem para funções quadráticas, mas podem ser estendidas para minimização de funções não quadráticas. A partir disso, realizamos testes numéricos e discutimos os gráficos de perfil de desempenho dessas variantes usando buscas direcionais baseadas nos algoritmos de Seção Áurea e de Armijo.

Palavras-chave: otimização irrestrita, gradientes conjugados, complexidade algorítmica.

Abstract

We discuss in this work the Conjugated Gradient Method and some variants of the method to minimizate a real function with several variables. The mainly objective is to obtain a faster convergence than de Gradient method and to reduce the computacional cost. The theoretical basis of the Conjugated Gradient Method is concentrated on the A-conjugated directions study, where A is a n dimensional square matrix. It's important to point out that the method's convergence properties depend of the unidirectional search efficacy. Furthermore, we prove that the conjugated directions method minimize a quadratic function defined on $\mathbb{R}^n$ using a maximum of n steps with great convergence rate. For the algorithmic complexity analysis we study the Chebyshev's Polinomials. We also show that the conjugated gradient method variants proposed by Polak and Ribière and by Fletcher and Reeves coincide for quadratic functions, but they can be extended for non linear minimization. From that, we do numerical tests and we discuss the performance profile graphics of this variants using directional search based on Golden line search and Armijo algorithms.

Key-words: unconstrained optimization, Conjugated Gradient, algorithmic complexity.

Sumário

Introdução

Neste trabalho estudamos o método de gradientes conjugados para otimização irrestrita, desde sua fundamentação teórica, definição do algoritmo propriamente dito para minimização de funções quadráticas, extensão do método para funções não quadráticas e análise de convergência do método.

Nosso objetivo é resolver o problema de minimizar uma função $f : \mathbb{R}^n \to \mathbb{R}$ sujeito a $x \in \mathbb{R}^n$.

Na literatura há vários métodos computacionais. O método de Cauchy faz uma busca na direção oposta ao gradiente da função a cada iteração. Este método, provavelmente o mais antigo, é globalmente convergente, porém possui convergência local linear. Por outro lado, o método de Newton tem convergência quadrática, mas faz uso de derivadas de segunda ordem, o que torna esse método caro. Por isso, estamos particularmente interessados em estudar métodos de gradiente conjugado que possuem convergência mais rápida que o método do gradiente (Cauchy), e além disso, tem custo computacional menor quando comparado ao Método de Newton, por não fazer uso de derivadas segundas.

Estudamos as propriedades de direções conjugadas, a partir das quais se faz uma busca linear. Recaímos, assim, em cada iteração do método em um problema de minimização unidirecional que pode ser resolvido, por exemplo, pelo Método de Seção Áurea ou Armijo. Cabe ressaltar que propriedades de convergência do método de Gradientes Conjugados dependem da eficácia da busca unidirecional.

Discutimos o algoritmo de gradientes conjugados juntamente com as variantes do mesmo, propostas por Polak e Ribière e por Fletcher e Reeves. Essa discussão inicial serviu para investigar o comportamento desses algoritmos e calibrá-los. Os fatos constatados foram que as variantes dos métodos de gradientes conjugados propostas por Polak e Ribière e por Fletcher e Reeves coincidem para funções quadráticas, mas podem ser estendidos para funções não quadráticas. Realizamos testes computacionais para avaliar o desempenho das diferentes variantes do método e assim podermos compará-los através de gráficos de perfil de desempenho.

Provamos que o método dos Gradientes Conjugados minimiza uma função quadrática convexa definida no $\mathbb{R}^n$ em no máximo n iterações com taxa de convergência ótima

com respeito ao valor da função da ordem $O\left(\frac{1}{k^2}\right)$ ao invés de $O\left(\frac{1}{k}\right)$ como o método de Cauchy, onde k representa o número de iterações necessário para minimizar a quadrática. Além disso, fizemos um estudo sobre polinômios de Chebyshev para o estudo da complexidade algorítmica.

O trabalho é organizado da seguinte forma: são apresentadas algumas ferramentas matemáticas, que envolvem resultados de Álgebra Linear e Cálculo no Capítulo 1.

No segundo capítulo é apresentado o problema, são discutidas propriedades das direções A-conjugadas, é apresentado o algoritmo para funções quadráticas, a extensão para não quadráticas e, além disso, são discutidas as formas de busca unidirecionais.

Já no terceiro capítulo, discutimos a minimização em um subespaço, em particular, a minimização em espaços de Krylov. Utilizando do conceito dos Polinômios de Chebyshev, provamos a complexidade algorítmica do método.

No quarto capítulo, são apresentados os testes computacionais utilizando as quatro variantes do método dos gradientes conjugados, ou seja, Polak-Ribière e Fletcher-Reeves, com busca unidirecional exata e inexata. Os algoritmos são comparados através de gráficos de perfil de desempenho.

Em suma, este trabalho envolve uma gama de conceitos e resultados muito importantes para a área de Otimização, e também de boa aplicabilidade.

Capítulo 1

Ferramentas Matemáticas

Este capítulo é dedicado à revisão de alguns conceitos que servirão de ferramenta para o desenvolvimento do nosso trabalho. Alguns conceitos básicos de Álgebra Linear e de Análise são considerados como pré-requisitos. As principais referências são [2, 3, 5, 8, 9, 15].

1.1 Resultados de Álgebra Linear

Definição 1.1 *Considere uma matriz $A \in \mathbb{R}^{n\times n}$. Um vetor $v \neq 0$ é autovetor de A se existe um escalar $\lambda \in \mathbb{R}$ tal que*

$$Av = \lambda v.$$

O escalar λ é o autovalor associado ao autovetor v.

Matriz simétrica

Uma matriz $A \in \mathbb{R}^{n\times n}$ é simétrica quando $A^T = A$. Uma propriedade importante é de que todos os autovalores de uma matriz simétrica são reais. Para provar isto, consideremos um número complexo $z = a + bi$. Seu conjugado é denotado por $\bar{z}$ e é igual a $\bar{z} = a - bi$. Analogamente, a conjugada de uma matriz $B \in \mathbb{C}^{n\times n}$ é a matriz denotada por $\bar{B}$, resultante da conjugação de cada um dos elementos de B.

Teorema 1.2 *Todos os autovalores de uma matriz simétrica são reais.*

Demonstração. Seja A uma matriz simétrica real. Consideremos um autovalor $\lambda \in \mathbb{C}$ de A e seu autovetor correspondente $v \in \mathbb{C}^n$. Assim $Av = \lambda v$, donde segue que $A\bar{v} = \bar{A}v = \bar{\lambda}\bar{v}$. Além disso, temos

$$\lambda \bar{v}^T v = \bar{v}^T(\lambda v) = \bar{v}^T(Av). \tag{1.1}$$

e

$$\bar{\lambda}\bar{v}^T v = (\bar{A}v)^T v = \bar{v}^T(Av). \tag{1.2}$$

De (1.1) e (1.2) decorre

$$(\lambda - \bar{\lambda})(\bar{v}^T v) = 0.$$

Como $v \neq 0$ e $\bar{v}^T v \neq 0$, concluímos que

$$(\lambda - \bar{\lambda}) = 0.$$

Portanto $\lambda = \bar{\lambda}$ e λ é real. □

Exemplo 1.3 *Os autovalores da matriz* $A = \begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix}$ *são* 0 *e* 2.

Lembremos que uma matriz $P \in \mathbb{R}^{n\times n}$ é ortogonal quando $P^T P = I$, ou seja, P é inversível e sua inversa P^{-1} coincide com sua transposta P^T.

O próximo teorema garante que uma matriz simétrica é diagonalizável por uma matriz ortogonal.

Teorema 1.4 *Seja* $A \in \mathbb{R}^{n\times n}$ *uma matriz simétrica, então existe uma matriz ortogonal* $P \in \mathbb{R}^{n\times n}$ *e uma matriz diagonal* D *de ordem* n *tal que* $A = PDP^T$.

Demonstração. Vamos fazer prova por indução em n. Para $n = 1$ temos que $P = [1]$. Como $P^T P = [1] = I_{1\times 1}$, então P é diagonal. Além disso

$$P^T A P = A_{1\times 1}.$$

Como toda matriz 1×1 é diagonal, temos que A é diagonal. Logo, existem P ortogonal e D diagonal tais que $P^T_{1\times 1} A_{1\times 1} P_{1\times 1} = D_{1\times 1}$.

Supondo que vale o resultado para $n-1$, provemos que vale para n. Seja λ um autovalor de A associado ao autovetor v. Como $v \in \mathbb{R}^n$ e o espaço gerado por v tem dimensão um, então seu complemento ortogonal, que denotaremos por $v^\perp$ tem dimensão $n-1$.

Consideremos $\{u_1, ..., u_{n-1}\}$ uma base ortogonal de $v^\perp$ e $M \in \mathbb{R}^{n\times(n-1)}$, sendo

$$M = \begin{bmatrix} u_1 & u_2 & \dots & u_{n-1} \end{bmatrix}.$$

Pelas propriedades de produto de matrizes e da definição de complemento ortogonal, temos que

$$AM = \begin{bmatrix} Au_1 & Au_2 & \dots & Au_{n-1} \end{bmatrix}.$$

Seja agora $B \in \mathbb{R}^{(n-1)\times(n-1)}$ da forma

$$B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1(n-1)} \\ b_{21} & b_{22} & \dots & b_{2(n-1)} \\ \vdots & \vdots & \ddots & \vdots \\ b_{(n-1)1} & b_{(n-1)2} & \dots & b_{(n-1)(n-1)} \end{bmatrix}.$$

Assim

$$MB = \begin{bmatrix} \sum_{i=1}^{n-1} u_i b_{i1} & \sum_{i=1}^{n-1} u_i b_{i2} & \dots & \sum_{i=1}^{n-1} u_i b_{i(n-1)} \end{bmatrix}.$$

Portanto

$$\begin{array}{rcl} AM & = & MB \\ M^T AM & = & M^T MB \\ M^T AM & = & B \end{array}.$$

Assim, concluímos que B é simétrica, pois

$$B^T = (M^T AM)^T = M^T A^T (M^T)^T = M^T AM = B.$$

Pela hipótese de indução existem matrizes quadradas Q e E pertencentes a $\mathbb{R}^{(n-1)\times(n-1)}$, tais que

$$Q^T Q = I \qquad \text{e} \qquad Q^T BQ = E.$$

Sendo

$$P = \begin{bmatrix} v & MQ \end{bmatrix} \qquad \text{e} \qquad P = \begin{bmatrix} \lambda & 0 \\ 0 & E \end{bmatrix},$$

em blocos temos que

$$P^T P = \begin{bmatrix} v^T \\ Q^T M^T \end{bmatrix} . \begin{bmatrix} v & MQ \end{bmatrix} = \begin{bmatrix} v^T v & v^T MQ \\ Q^T M^T v & Q^T M^T MQ \end{bmatrix}.$$

Considerando as q_{ij} entradas de Q e que v é ortogonal, o produto $v^T MQ$ fica

$$\begin{array}{rcl} v^T MQ & = & \begin{bmatrix} v^T \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{n-1} u_i q_{i1} & \sum_{i=1}^{n-1} u_i q_{i2} & \dots & \sum_{i=1}^{n-1} u_i q_{i(n-1)} \end{bmatrix} \\ & = & \begin{bmatrix} \sum_{i=1}^{n-1} v_1^T u_i q_{i1} & \sum_{i=1}^{n-1} v_2^T u_i q_{i2} & \dots & \sum_{i=1}^{n-1} v_{(n-1)}^T u_i q_{i(n-1)} \end{bmatrix}. \end{array}$$

Como cada u_i é ortogonal a v, então $v^T MQ$ é a matriz nula $\mathbb{R}^{1\times(n-1)}$. Finalmente, $Q^T M^T MQ = Q^T IQ = I_{(n-1)\times(n-1)}$ e assim

$$\begin{bmatrix} v^T v & v^T MQ \\ Q^T M^T v & Q^T M^T MQ \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & I_{(n-1)\times(n-1)} \end{bmatrix} = I_{n\times n}.$$

Portanto P é ortogonal e

$$P^TAP = \begin{bmatrix} v^T \\ Q^TM^T \end{bmatrix} A \begin{bmatrix} v & MQ \end{bmatrix} = \begin{bmatrix} v^TAv & v^TAMQ \\ Q^TM^TAv & Q^TM^TAMQ \end{bmatrix}.$$

Como $Av = \lambda v$, temos que

$$v^TAMQ = v^TA^TMQ = (Av)^TMQ = (\lambda v)^TMQ = [0]_{1\times(n-1)}.$$

Analogamente $Q^TM^TAv = [0]_{(n-1)\times 1}$. Logo $Q^TM^TAMQ = Q^TQ = E$. Assim

$$\begin{bmatrix} v^TAv & v^TAMQ \\ Q^TM^TAv & Q^TM^TAMQ \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & E \end{bmatrix} = D.$$

Portanto, existem matrizes P ortogonal e D diagonal, tais que $P^TAP = D$, ou seja

$$A = PDP^T.$$

□

Algumas propriedades de autovalores além das clássicas estudadas nos cursos de Álgebra Linear serão úteis neste trabalho.

Uma das propriedades relaciona os autovalores de A com os autovalores da matriz $p(A)$, onde $p : \mathbb{R} \to \mathbb{R}$ é um polinômio.

Lema 1.5 *Seja $p : \mathbb{R} \to \mathbb{R}$ um polinômio. Se λ é autovalor de uma matriz A com autovetor v, então $p(\lambda)$ é autovalor de $p(A)$ com o mesmo autovetor.*

Demonstração. Consideremos o polinômio $p(t) = \sum_{i=0}^{k} a_it^i$ e λ um autovalor de A com autovetor v. Assim

$$p(A)v = \sum_{i=0}^{k} a_iA^iv = \sum_{i=0}^{k} a_i\lambda^iv = p(\lambda)v,$$

como queríamos demonstrar. □

Podemos ainda generalizar esse resultado através do seguinte lema.

Lema 1.6 *Seja $A \in \mathbb{R}^{n\times n}$ uma matriz simétrica com autovalores $\lambda_1, ..., \lambda_n$ e $p : \mathbb{R} \to \mathbb{R}$ um polinômio. Então $p(\lambda_1), p(\lambda_2), ..., p(\lambda_n)$ são os autovalores de $p(A)$.*

Demonstração. Considere $p(t) = \sum_{i=0}^{k} a_it^i$. Como A é simétrica existe uma matriz P ortogonal e uma matriz D diagonal cujos elementos diagonais são os autovalores de A tais

que $A = PDP^T$. Assim

$$p(A) = \sum_{i=0}^{k} a_i A^i = \sum_{i=0}^{k} a_i P D^i P^T = P \left[\sum_{i=0}^{k} a_i D^i \right] P^T = P \begin{bmatrix} p(\lambda_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & p(\lambda_n) \end{bmatrix} P^T,$$

como queríamos demonstrar. □

1.2 Matriz definida positiva

Nesta seção falaremos um pouco de matriz definida positiva, que tem fundamental importância para o desenvolvimento deste trabalho.

Definição 1.7 *Seja $A \in \mathbb{R}^{n\times n}$ uma matriz simétrica. Dizemos que A é definida positiva quando $x^T Ax > 0$, para todo $x \in \mathbb{R}^n - \{0\}$. Tal propriedade é denotada por $A > 0$. Se $x^T Ax \geq 0$, para todo $x \in \mathbb{R}^n$, A é dita semidefinida positiva, fato este denotado por $A \geq 0$.*

O próximo resultado garante que os autovalores de uma matriz definida positiva são positivos e reciprocamente.

Teorema 1.8 *Uma matriz é definida positiva se, e somente se, seus autovalores são positivos.*

Demonstração. Considere $A \in \mathbb{R}^{n\times n}$ uma matriz simétrica. Então existe uma base ortonormal de autovetores $\{v^1, v^2, \dots, v^n\}$. Sendo o conjunto $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ os autovalores associados, definindo $P = (v^1 v^2 \dots v^n)$como a matriz cujas colunas são os autovetores e $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, temos

$$AP = (Av^1 Av^2 \dots Av^n) = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) = PD.$$

Considerando que cada a_i corresponde a uma linha de A temos que

$$AP = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} . \begin{bmatrix} v^1 & v^2 & \dots & v^n \end{bmatrix} = \begin{bmatrix} Av^1 & Av^2 & \dots & Av^n \end{bmatrix}.$$

Por outro lado

$$PD = \begin{bmatrix} v^1 & v^2 & \dots & v^n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda_1 v^1 & \lambda_2 v^2 & \dots & \lambda_n v^n \end{bmatrix}.$$

Além disso, $P^T P = I$ e, portanto

$$A = PDP^T. \tag{1.3}$$

Isto nos permite caracterizar a positividade de uma matriz em função dos seus autovalores. Basta notar que, dado $x \in \mathbb{R}^n$ e definindo $z = P^T x$, temos

$$x^T Ax = (x^T P)^T D(P^T x) = z^T Dz = \sum_{i=1}^{n} \lambda_i {z_i}^2. \tag{1.4}$$

Se A é definida positiva, então $x^T Ax > 0$ para todo $x \neq 0$ e isto implica que $\sum_{i=1}^{n} \lambda_i {z_i}^2 > 0$ para todo $z \neq 0$. Assim $\lambda_i > 0$, com $i = 0, 1, \dots, n$. Por outro lado, se os autovalores são positivos, então $\sum_{i=1}^{n} \lambda_i {z_i}^2 > 0$ e isto implica que $x^T Ax > 0$. O que conclui nossa demonstração. □

A norma de uma matriz é definida como

$$\|A\| = \sup \{\|Ax\| \mid x \in \mathbb{R}^n, \|x\| = 1\}. \tag{1.5}$$

O próximo teorema garante que a norma de uma matriz definida positiva coincide com seu maior autovalor.

Teorema 1.9 *Seja $A \in \mathbb{R}^{n \times n}$ uma matriz simétrica definida positiva. Então, usando a norma euclidiana em* (1.5)*, temos*

$$\|A\| = \lambda_{\max},$$

onde $\lambda_{\max}$ é o maior autovalor de A.

Demonstração. Consideremos $x \in \mathbb{R}^n$ unitário arbitrário.

$$\|Ax\|^2 = (Ax)^T Ax = (x)^T A^2 x.$$

Como A é simétrica, existem P ortogonal e D diagonal tais que $A = PDP^T$. Substituindo na expressão acima temos

$$\|Ax\|^2 = (x)^T (PDP^T)(PDP^T)x = (x)^T PD^2 P^T x.$$

Consideremos a mudança de variável $y = P^T x$. Assim

$$\|Ax\|^2 = y^T D^2 y = \sum_{i=1}^{n} \lambda_i^2 (y_i)^2.$$

Como A é definida positiva, todos os seus autovalores são positivos. Consideremos $\lambda_{\max}$ o maior deles, temos

$$\|Ax\|^2 \leq \sum_{i=1}^{n} \lambda_{\max}^2 (y_i)^2 = \lambda_{\max}^2 \sum_{i=1}^{n} (y_i)^2.$$

Como x é unitário e P é ortogonal, temos que

$$\|y\|^2 = (P^T x)^T (P^T x) = x^T x = \|x\|^2 = 1.$$

Assim

$$\|Ax\|^2 \leq \lambda_{\max}^2.$$

Logo $\lambda_{\max}$ é uma cota superior para $\|Ax\|$ sempre que $\|x\| = 1$. Como $\lambda_{\max}$ é um autovalor de A, existe um autovetor unitário v tal que $Av = \lambda_{\max} v$ e

$$\|Av\|^2 = (Av)^T Av = (v)^T A^2 v = \lambda_{\max}^2 \|v\|^2 = \lambda_{\max}^2.$$

Como $\|A\| = \max\{\|Ax\| \mid \|x\| = 1, x \in \mathbb{R}^n\}$, concluímos que $\|A\| = \lambda_{\max}$. Como queríamos demonstrar. □

O próximo resultado será útil nos próximos capítulos.

Lema 1.10 *Considere $A \in R^{n \times n}$ uma matriz definida positiva com autovalores $\lambda_1, ..., \lambda_n$, cujo maior autovalor é $\lambda_{\max}$. Seja $p : \mathbb{R} \to \mathbb{R}$ um polinômio. Então*

$$\left\| A(p(A))^2 \right\| \leq \max_{0 \leq z \leq \lambda_{\max}} z(p(z))^2.$$

Demonstração. Considere o polinômio definido por $q(z) = z(p(z))^2$. Pelo Teorema 1.9 e Lema 1.6, temos que

$$\|q(A)\| = \max\{q(\lambda_1), q(\lambda_2), ..., q(\lambda_n)\} \leq \max_{0 \leq z \leq \lambda_{\max}} q(z),$$

completando assim a demonstração. □

1.3 Ferramentas de Cálculo

Definição 1.11 *Consideremos $f : \mathbb{R}^n \to \mathbb{R}$ uma função de classe $\mathcal{C}^2$. O gradiente de f e a Hessiana de f são definidos por*

$$\nabla f = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{pmatrix} \quad e \quad \nabla^2 f = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}.$$

Uma outra definição importante é a da derivada de uma função vetorial, que apresentaremos a seguir.

Definição 1.12 *Consideremos uma função vetorial $f : \mathbb{R}^n \to \mathbb{R}^m$. Sua derivada, chamada de jacobiana, é a matriz*

$$J_f = f' = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

Podemos perceber que a linha i da jacobiana de f é o gradiente transposto da componente f_i. Em particular, para $m = 1$, temos $f' = (\nabla f)^T$. Além disso, $\nabla^2 f = J_{\nabla f}$.

O gradiente de uma função será muito útil em todas as nossas discussões ao longo deste trabalho. Por isso precisamos enfatizar que ele tem propriedades muito importantes, tais como:

1. O gradiente é uma direção de crescimento da função;
2. É a direção de crescimento mais rápido;
3. O gradiente é perpendicular à curva de nível da função.

Definidos os elementos necessários, agora podemos apresentar a aproximação de Taylor de Primeira Ordem.

Teorema 1.13 *Considere $f : \mathbb{R}^n \to \mathbb{R}$ uma função de classe $\mathcal{C}^1$ e $a \in \mathbb{R}^n$. Podemos escrever*

$$f(x) = f(a) + \nabla f(a)^T (x - a) + r(x),$$

com $\lim\limits_{x \to a} \dfrac{r(x)}{\|x - a\|} = 0.$

O polinômio $p_1(x) = f(a) + \nabla f(a)^T(x-a)$ é chamado polinômio de Taylor de ordem 1 da função f. Dentre todos os polinômios de grau menor ou igual a 1, ele é o que melhor aproxima f. E é também o único que satisfaz

$$p(a) = f(a) \quad \text{e} \quad p'(a) = f'(a).$$

O limite nulo no Teorema 1.13 significa que para x próximo de a o resto $r(x)$ é muito pequeno e vai para zero mais rápido que $\|x-a\|$.

É importante e conveniente perceber que podemos reescrever o Teorema 1.13 fazendo uma mudança de variável. Assim, definindo $d = x - a$, temos

$$f(a+d) = f(a) + \nabla f(a)^T d + r(d),$$

com $\lim\limits_{d\to 0} \dfrac{r(d)}{\|d\|} = 0$.

Função Convexa

Definição 1.14 *Seja $C \in \mathbb{R}^n$ 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 $t \in [0,1]$.

Teorema 1.15 *Sejam $f : \mathbb{R}^n \to \mathbb{R}$ uma função diferenciável e $C \subset \mathbb{R}^n$ convexo. A função f é convexa em C se, e somente se,*

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

para todos $x, y \in C$.

Demonstração. Considerando f convexa, para $x, y \in C$ e $t \in (0,1]$ quaisquer, $d = y - x$ e usando a definição de convexidade, temos que $x + td \in C$ e

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

Portanto

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

Precisamos agora provar a recíproca. Para isso, consideremos $z = (1-t)x + ty$ e notemos que

$$f(x) \geq f(z) + \nabla f(z)^T (x - z)$$

e

$$f(y) \geq f(z) + \nabla f(z)^T (y - z).$$

Multiplicando a primeira delas por $(1-t)$ e a segunda por t obtemos que

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

conforme queríamos demonstrar. □

Deste teorema temos uma consequência imediata, que será exibida no corolário a seguir.

Corolário 1.16 *Sejam $f : \mathbb{R}^n \to \mathbb{R}$ uma função convexa e diferenciável e $C \subset \mathbb{R}^n$ convexo. Se $\nabla f(x^*)^T (y - x^*) \geq 0$ para todo $y \in C$, então x^* é um minimizador global de f em C. Em particular, todo ponto estacionário é minimizador global.*

1.4 Polinômios de Chebyshev

Nesta seção estudaremos os polinômios de Chebyshev (ou Tchebychev), que foram desenvolvidos para estudar a aproximação por mínimos quadrados e probabilidade. Seus resultados se aplicam à interpolação polinomial (usando as raízes dos polinômios de primeira ordem).

Lagrange e Legendre haviam estudado conjuntos individuais de polinômios ortogonais, mas Chebyshev foi o primeiro a perceber as importantes consequências dos estudos da teoria em geral.

Estes polinômios receberam esse nome após Pafnuty Chebyshev definí-los como uma sequência de polinômios ortogonais e eles são facilmente obtíveis de forma recursiva.

Os polinômios de Chebyshev de primeira ordem $T_k(x)$ e de segunda ordem $U_k(x)$ são polinômios de grau k e a sequência de polinômios de todos os graus forma uma sequência polinomial.

As referências desta seção são [2, 14].

Definição 1.17 *O polinômio de Chebyshev de grau k, $T_k : [-1, 1] \to \mathbb{R}$ é definido por*

$$T_k(x) = \cos[k \arccos(x)].$$

Discutiremos a seguir algumas propriedades dos polinômios de Chebyshev. A primeira delas corresponde a uma importante relação de recorrência.

Lema 1.18 *Para $k \geq 1$ e $x \in [-1, 1]$, vale a seguinte relação de recorrência:*

$$T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x)$$

onde $T_0(x) = 1$ e $T_1(x) = x$.

Demonstração. Pela definição dos polinômios de Chebyshev temos que:

$$T_0(x) = \cos[0 \arccos(x)] = \cos 0 = 1$$

e

$$T_1(x) = \cos[1 \arccos(x)] = x.$$

Para cada $x \in [-1, 1]$, consideremos a seguinte mudança de variável $\theta = \arccos x$, com $\theta \in [0, \pi]$. Pela definição dos polinômios de Chebyshev, temos para $k \geq 1$

$$T_k(x) = \cos(k\theta).$$

Usando as relações trigonométricas de soma de arcos, obtemos

$$T_{k+1}(x) = cos((k + 1)\theta) = \cos(k\theta)\cos(\theta) - \sin(k\theta)\sin(\theta)$$

e

$$T_{k-1}(x) = cos((k - 1)\theta) = \cos(k\theta)\cos(\theta) + \sin(k\theta)\sin(\theta).$$

Somando os membros das igualdades anteriores, segue que

$$T_{k+1}(x) = 2\cos(k\theta)\cos(\theta) - T_{k-1}(x).$$

Voltando a variável x e usando a definição dos polinômios de Chebyshev

$$\begin{aligned} T_{k+1}(x) &= 2\cos[k \arccos(x)]\cos[\arccos(x)] - T_{k-1}(x) \\ &= 2xT_k(x) - T_{k-1}(x), \end{aligned}$$

conforme queríamos demonstrar. □

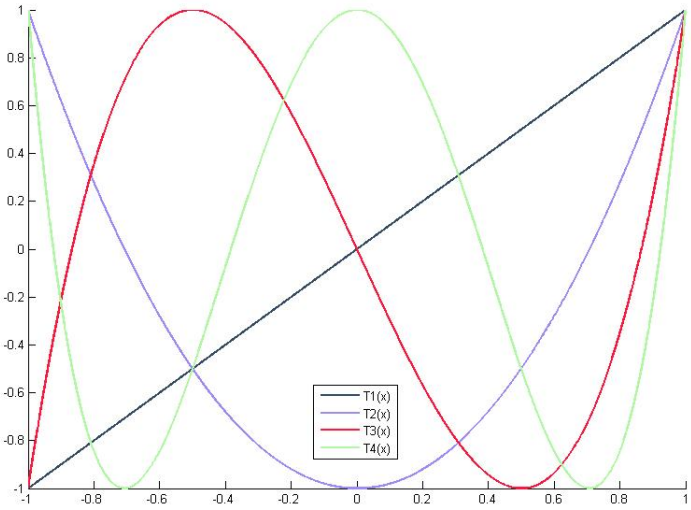

Figura 1.1: Alguns polinômios de Chebyshev.

Segue diretamente do Lema anterior que

$$\begin{aligned}
T_0(x) &= 1 \\
T_1(x) &= x \\
T_2(x) &= 2xT_1(x) - T_0(x) = 2x^2 - 1 \\
T_3(x) &= 2xT_2(x) - T_1(x) = 4x^3 - 3x \\
T_4(x) &= 2xT_3(x) - T_2(x) = 8x^4 - 8x^2 + 1 \\
T_5(x) &= 2xT_4(x) - T_3(x) = 16x^5 - 20x^3 + 5x \\
T_6(x) &= 2xT_5(x) - T_4(x) = 32x^6 - 48x^4 + 18x^2 - 1.
\end{aligned}$$

A Fig. 1.1 ilustra alguns destes polinômios.

Consideremos agora um polinômio de grau k. Ele pode ser escrito da forma

$$P_k(x) = \sum_{i=0}^{k} a_i x^i = a_0 x^0 + a_1 x^1 + ... + a_k x^k.$$

Definição 1.19 *Seja* $P_k(x) = \sum_{i=0}^{k} a_i x^i$ *um polinômio de ordem* k*. Sua norma é definida por* $\|P_k(x)\| = \max\{|P_k(x)|, x \in [-1,1]\}$.

Lema 1.20 *Seja* $T_k(x) = \cos[k \arccos(x)]$*, com* $x \in [-1,1]$*. Temos que* $\|T_k(x)\| = 1$.

Demonstração. Por definição, temos que

$$\|T_k(x)\| = \max\{|\cos[k \arccos(x)]|\} = 1.$$

□

Pode-se provar por indução que para $k > 1$ par temos

$$T_k(x) = 2^{k-1}x^k + ... + (-1)^{\frac{k}{2}}, \tag{1.6}$$

com $a_1 = a_3 = ... = 0$. Para o caso em que k é ímpar, o polinômio fica

$$T_k(x) = 2^{k-1}x^k + ... + (-1)^{\frac{k-1}{2}}kx, \tag{1.7}$$

com $a_0 = a_2 = a_4 = ... = 0$.

Observamos pelas expressões dos polinômios que quando k é par, T_k é uma função par, e quando k é ímpar, T_k é ímpar. Provaremos isto no lema a seguir.

Lema 1.21

$$T_k(x) = \begin{cases} T_k(-x), & \textit{se } k \textit{ é par}, \\ -T_k(-x), & \textit{se } k \textit{ é ímpar}. \end{cases} \tag{1.8}$$

Demonstração. A demonstração será feita por indução. Para $k = 0$

$$T_0(x) = 1 = T_0(-x).$$

Para $k = 1$

$$T_1(x) = x = -T_1(-x).$$

Suponhamos que (1.8) vale para $k \leq m$. Vamos provar que vale para $k = m + 1$. Pela fórmula de recorrência dada pelo Lema (1.18) temos que

$$T_{m+1}(x) = 2xT_m(x) - T_{m-1}(x)$$

e

$$T_{m+1}(-x) = -2xT_m(-x) - T_{m-1}(-x).$$

Suponha inicialmente que $m+1$ é par. Neste caso m é ímpar e $m-1$ é par. Logo, usando as igualdades anteriores e a hipótese de indução

$$\begin{aligned} T_{m+1}(x) &= 2x(-T_m(-x)) - T_{m-1}(-x) \\ &= -2xT_m(-x) - T_{m-1}(-x) \\ &= T_{m+1}(-x), \end{aligned}$$

e portanto T_{m+1} é uma função par. Analogamente, suponha agora que $m+1$ é ímpar. Neste caso m é par e $m-1$ é ímpar. Usando novamente as igualdades e a hipótese de indução

$$\begin{aligned} T_{m+1}(x) &= 2x(T_m(-x)) - (-T_{m-1}(-x)) \\ &= -(-2xT_m(-x) - T_{m-1}(-x)) \\ &= -T_{m+1}(-x), \end{aligned}$$

e portanto T_{m+1} é uma função ímpar, completando a demonstração. □

Considere o espaço dos polinômios de grau menor ou igual a k definidos em $[-1,1]$ munido do produto interno

$$\langle p, q\rangle = \int_{-1}^{1} p(x)\omega q(x)dx$$

onde $\omega : \mathbb{R} \to \mathbb{R}$ definida por $\omega(x) = \frac{1}{\sqrt{(1-x^2)}}$ é uma função peso.

Teorema 1.22 *Os polinômios de Chebyshev $T_k(x)$ são ortogonais para $(-1,1)$ com relação à função peso $\omega(x) = (1-x^2)^{-\frac{1}{2}}$.*

Demonstração. Sejam $T_n, T_m : [-1,1] \to \mathbb{R}$ polinômios de chebyshev de grau menor ou igual a k. Para mostrar a ortogonalidade dos polinômios de Chebyshev consideremos

$$< T_n, T_m >= \int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}}dx = \int_{-1}^{1} \frac{\cos[n\arccos(x)]\cos[m\arccos(x)]}{\sqrt{1-x^2}}dx. \qquad (1.9)$$

Realizando a mudança de variável $\theta = \arccos(x)$ e derivando temos

$$d\theta = -\frac{1}{\sqrt{1-x^2}}dx.$$

Assim, usando as propriedades de integração podemos reescrever (1.9) como

$$< T_n, T_m >= -\int_{\pi}^{0} \cos(n\theta)\cos(m\theta)d\theta = \int_{0}^{\pi} \cos(n\theta)\cos(m\theta)d\theta.$$

Considerando $n \neq m$ e usando que

$$\cos(n\theta)\cos(m\theta) = \frac{1}{2}[\cos(n+m)\theta + \cos(n-m)\theta],$$

temos

$$\begin{aligned} < T_n, T_m > &= \tfrac{1}{2}\int_{\pi}^{0} \cos((n+m)\theta)d\theta + \tfrac{1}{2}\int_{0}^{\pi} \cos((n-m)\theta)d\theta \\ &= \left[\tfrac{1}{2(n+m)}\sin((n+m)\theta) + \tfrac{1}{2(n-m)}\sin((n-m)\theta)\right]_{0}^{\pi} \\ &= 0. \end{aligned}$$

Analogamente para $n = m$, temos

$$\begin{aligned} < T_n, T_n > &= \tfrac{1}{2}\int_{\pi}^{0} \cos((2n)\theta)d\theta + \tfrac{1}{2}\int_{0}^{\pi} d\theta \\ &= \left[\tfrac{1}{2(2n)}\sin((2n)\theta) + \tfrac{\theta}{2}\right]_{0}^{\pi} \\ &= \tfrac{\pi}{2}. \end{aligned}$$

Completando assim a demonstração. □

Capítulo 2

Métodos de gradientes conjugados

O método de direções conjugadas tem por objetivo uma convergência mais rápida que a do método do gradiente (método de Cauchy), além de ter uma vantagem de redução no custo computacional do método de Newton por só usar derivadas de primeira ordem. Para mais referências, consulte [1, 7, 13].

2.1 O problema

Dada uma função $f : \mathbb{R}^n \to \mathbb{R}$ considere o problema de minimização irrestrita da seguinte forma:

$$\text{minimizar } f(x), \text{ com } x \in \mathbb{R}^n. \tag{2.1}$$

Estamos inicialmente interessados no caso em que f é uma função quadrática, ou seja

$$f(x) = \frac{1}{2}x^T Ax + b^T x + c, \tag{2.2}$$

com $A \in \mathbb{R}^{n\times n}$ uma matriz simétrica definida positiva, $b \in \mathbb{R}^n$ e $c \in \mathbb{R}$. A função f tem um único minimizador x^*, que é global e satisfaz

$$\nabla f(x^*) = Ax^* + b = 0. \tag{2.3}$$

Olhando para a vizinhança do minimizador, podemos dizer que a menos de deslo-

camentos e translações, o gráfico dessas funções apresenta um comportamento característico que é mostrado na Fig. 2.1. Além disso, podemos apresentar o gráfico das curvas de nível dessa função. É possível perceber que quanto mais escuras as curvas de nível, mais baixos são seus níveis e que o centro das curvas de nível é justamente o minimizador de f.

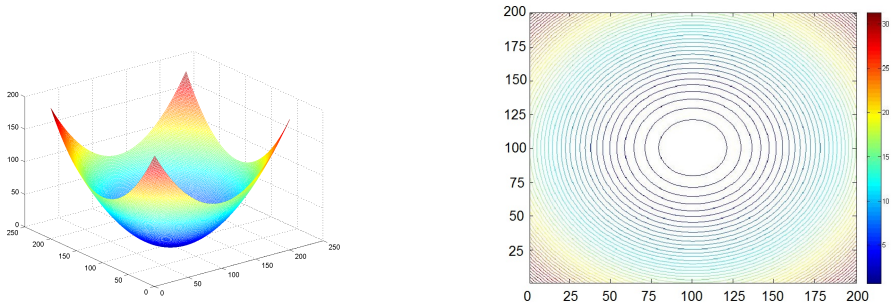

Figura 2.1: Gráfico de f e das curvas de nível.

É importante lembrar que uma função quadrática é fortemente convexa se, e somente se, sua matriz associada é definida positiva.

2.2 Direções Conjugadas

Definição 2.1 *Um conjunto de vetores* $d^0, ..., d^k \in \mathbb{R}^n - \{0\}$ *é dito A-conjugado, se:*

$$(d^i)^T A d^j = 0$$

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

Lema 2.2 *Direções conjugadas são linearmente independentes: para qualquer matriz* $A \in \mathbb{R}^{n \times n}$ *simétrica definida positiva, qualquer conjunto de direções A-conjugadas é linearmente independente.*

Demonstração. Vamos supor que um dos vetores do conjunto possa ser escrito como combinação linear dos outros:

$$d^0 = \sum_{i=1}^{k} \beta_i d^i$$

com $\beta_i \in \mathbb{R}$, $i = 1, ..., k$.

$$0 < (d^0)^T A d^0 = \sum_{i=1}^{k} (\beta_i d^i)^T A d^0 = \sum_{i=1}^{k} \beta_i (d^i)^T A d^0.$$

Pela Definição 2.1 , temos que $\sum_{i=1}^{k} \beta_i (d^i)^T A d^0 = 0$, o que é uma contradição. Portanto direções conjugadas são linearmente independentes. □

Pelo Lema 2.2, um conjunto de vetores A-conjugados no $\mathbb{R}^n$ não pode conter mais de n elementos.

Tendo isso em vista, dado um conjunto de vetores A-conjugados $\{d^0, d^1, ..., d^{n-1}\}$, $x^k \in \mathbb{R}^n$, $\alpha_k \in \operatorname{argmin} f(x^k + \alpha d^k)$, nosso objetivo é considerar um iterando x^{k+1} definido por

$$x^{k+1} = x^k + \alpha_k d^k, \tag{2.4}$$

com $k = 0, ..., n-1$.

Consideremos a função f definida por (2.2). Sabendo que $\nabla f(x) = Ax + b$, temos que

$$\nabla f(x^{k+1}) = Ax^{k+1} + b = A(x^k + \alpha_k d^k) + b = \nabla f(x^k) + \alpha_k A d^k. \tag{2.5}$$

Queremos agora definir como serão calculados os valores do comprimento de passo α_k, sendo que eles satisfazem a seguinte relação:

$$f(x^k + \alpha_k d^k) = \min_{\alpha \in \mathbb{R}} f(x^k + \alpha d^k). \tag{2.6}$$

Dados x^k, $d^k \in \mathbb{R}^n$. Definamos a função $\varphi : \mathbb{R} \to \mathbb{R}$ por

$$\varphi(\alpha) = f(x^k + \alpha d^k)$$

Na Fig. 2.2 temos o gráfico de f, um segmento de reta representando os pontos $x^k + \alpha d^k$ e uma curva sendo o gráfico de φ. Calculemos sua derivada.

$$\begin{aligned} \varphi'(\alpha) &= \nabla f(x^k + \alpha d^k)^T d^k \\ &= (\nabla f(x^k) + \alpha A d^k)^T d^k \\ &= \nabla f(x^k)^T d^k + \alpha (d^k)^T A d^k. \end{aligned}$$

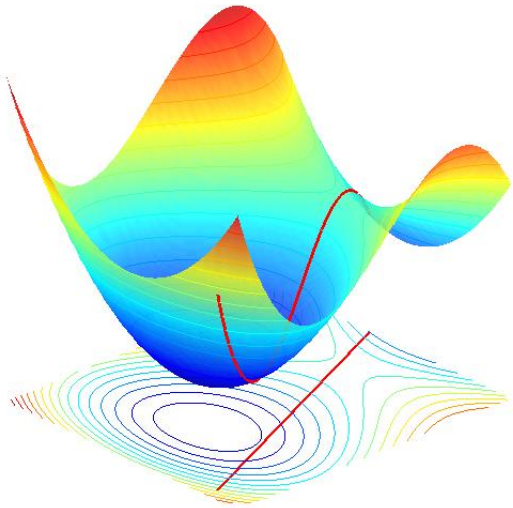

Figura 2.2: Restrição de uma função a um segmento.

Como α_k é o minimizador de $\varphi(\alpha)$, então $\varphi'(\alpha_k) = 0$. Logo:

$$\alpha_k = -\frac{\nabla f(x^k)^T d^k}{(d^k)^T A d^k}. \tag{2.7}$$

Definido α_k, obtemos

$$\nabla f(x^{k+1})^T d^k = \nabla f(x^k + \alpha_k d^k)^T d^k = \varphi'(\alpha_k) = 0, \tag{2.8}$$

o que significa que $\nabla f(x^{k+1})$ é ortogonal a d^k.

Métodos de direções conjugadas não são métodos de descida. A natureza de cada método é definida pela construção do conjunto de vetores A-conjugados associado.

O que nos interessa é que todo método de direções A-conjugadas encontra solução para o problema de minimização irrestrita de uma função quadrática em no máximo n iterações, resultado que será provado no Teorema a seguir.

Teorema 2.3 *Seja f uma função quadrática. Para qualquer ponto inicial $x^0 \in \mathbb{R}^n$, com qualquer escolha de conjunto de vetores A-conjugados, o ponto x^n obtido de acordo com* (2.4) *é a solução global do problema* (2.1).

Demonstração. Seja x^* o minimizador de f e $\{d^0, d^1, \ldots, d^{n-1}\}$ um conjunto de vetores A-conjugado. Pelo Lema (2.2), o conjunto $\{d^0, d^1, \ldots, d^{n-1}\}$ forma uma base do $\mathbb{R}^n$, então

podemos escrever

$$x^* - x^0 = \sum_{k=0}^{n-1} \gamma_k d^k, \tag{2.9}$$

para alguns $\gamma_k \in \mathbb{R}$ com $k = 0, 1, ..., n-1$. Fixando k, com $k \in \{0, \ldots, n-1\}$, vamos multiplicar a igualdade (2.9) por $(d^k)^T A$, obtendo

$$(d^k)^T A(x^* - x^0) = \gamma_k (d^k)^T A d^k, \tag{2.10}$$

porque as direções são A-conjugadas. Logo:

$$\gamma_k = \frac{(d^k)^T A(x^* - x^0)}{(d^k)^T A d^k}. \tag{2.11}$$

Por outro lado, pela relação (2.4) temos

$$x^k = x^0 + \alpha_0 d^0 + \ldots + \alpha_{k-1} d^{k-1},$$

que multiplicada por $(d^k)^T A$ fornece

$$(d^k)^T A x^k = (d^k)^T A x^0 \tag{2.12}$$

pois as direções são A-conjugadas. Substituindo isto em (2.11) temos

$$\gamma_k = \frac{(d^k)^T A(x^* - x^k)}{(d^k)^T A d^k}$$

e sabendo que $Ax^* = -b$ obtemos

$$\begin{aligned} \gamma_k &= \frac{-(d^k)^T b - (d^k)^T A x^k}{(d^k)^T A d^k} \\ &= \frac{-(d^k)^T \nabla f(x^k)}{(d^k)^T A d^k} \\ &= \alpha_k. \end{aligned}$$

Portanto, podemos reescrever a relação (2.9) da forma

$$x^* = x^0 + \sum_{k=0}^{n-1} \alpha_k d^k = x^n,$$

onde a última igualdade segue da relação (2.4). □

Provaremos a seguir uma importante relação que mostra que o gradiente em um ponto dado é ortogonal às direções anteriores.

Lema 2.4 *Dado $x^0 \in \mathbb{R}^n$, seja a sequência definida em* (2.4)*. Então*

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

para todo $j < k$.

Demonstração. Vamos provar por indução em k. Para $k = 1$ temos que $\nabla f(x^1)^T d^0 = 0$ pela relação (2.8). Temos por hipótese de indução que vale para $k-1$ e queremos provar que vale para k. Para $j = k-1$ temos que $\nabla f(x^k)^T d^j = 0$, pela relação (2.8). E para $j < k-1$, pela relação (2.5), temos

$$\nabla f(x^k)^T d^j = (\nabla f(x^{k-1}) + \alpha_{k-1} A d^{k-1})^T d^j.$$

Usando a hipótese de indução e pelo fato das direções serem A-conjugadas, temos que

$$\nabla f(x^k)^T d^j = \nabla f(x^{k-1})^T d^j = 0,$$

que conclui a demonstração. □

Teorema 2.5 *Dados $x^0 \in \mathbb{R}^n$, seja $\{d^0, d^1, ..., d^k\}$ um conjunto de direções A-conjugadas e α_k definido de acordo com a relação* (2.6)*. Então x^{k+1} é solução global de*

$$\min_{x \in D_k} f(x)$$

onde $D_k = x^0 + [d^0, d^1, ..., d^k]$

Demonstração. Se $x^{k+1} \in D_k$, então ele pode ser escrito da forma

$$x^{k+1} = x^0 + \sum_{i=0}^{k} t_i d^i.$$

E todo $x \in D_k$ pode ser assim escrito

$$x = x^0 + \sum_{i=0}^{k} s_i d^i.$$

Então, $x - x^{k+1} = \sum_{i=0}^{k} r_i d^i \in [d^0, \ldots, d^k]$, onde $r_i = s_i - t_i$. Pelo Lema 2.4 temos que

$$\nabla f(x^{k+1})^T (x - x^{k+1}) = 0.$$

Como f é convexa e D_k é convexo, aplicando o Corolário 1.16 concluímos que o ponto x^{k+1} é solução global do problema. □

Observações

Existe uma simples interpretação das propriedades das direções conjugadas. Se a matriz Hessiana da função quadrática é diagonal, então as curvas de nível da função são elipses cujos eixos são alinhados com os eixos coordenados. Assim podemos encontrar o minimizador desta função realizando minimização unidimensional ao longo das direções coordenadas $e_1, e_2, ..., e_n$; a partir de qualquer ponto $x^k \in \mathbb{R}^n$.

A Fig. 2.3 mostra as curvas de nível de uma função em $\mathbb{R}^2$ e que a partir de um ponto x^0 tomando como direção inicial a direção dos eixos coordenados, conseguimos encontrar o minimizador global desta função em somente duas iterações, o ponto x^2.

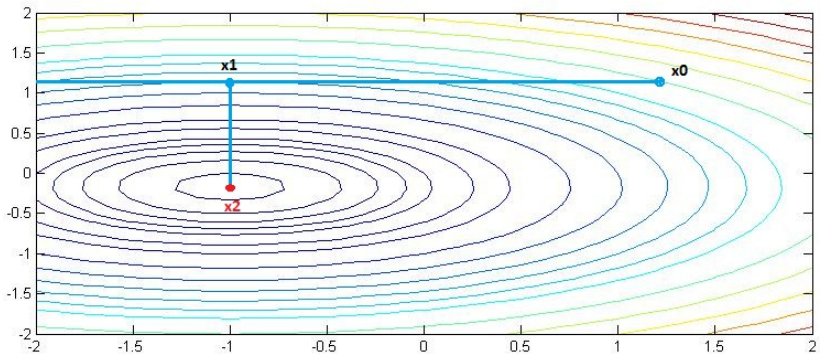

Figura 2.3: Curva de nível de f com matriz Hessiana diagonal.

Já quando A não é diagonal, as curvas de nível da função f ainda são elipses, mas elas geralmente não são alinhadas com as direções coordenadas. A estratégia de minimização sucessiva em torno e ao longo dessas direções não garante que chegaremos na solução em n iterações (ou então num número finito de iterações).

A Fig. 2.4 mostra as curvas de nível de uma função em $\mathbb{R}^2$ e que a partir de um ponto x^0 tomando como direção inicial a direção dos eixos coordenados, não garante que encontraremos solução em n iterações.

2.3 Método dos Gradientes Conjugados

Nesta seção estudaremos um modo de gerar direções conjugadas e apresentaremos o método dos gradientes conjugados para minimizar uma função quadrática.

Dado $x^0 \in \mathbb{R}^n$ definamos

$$d^0 = -\nabla f(x^0) \tag{2.13}$$

e

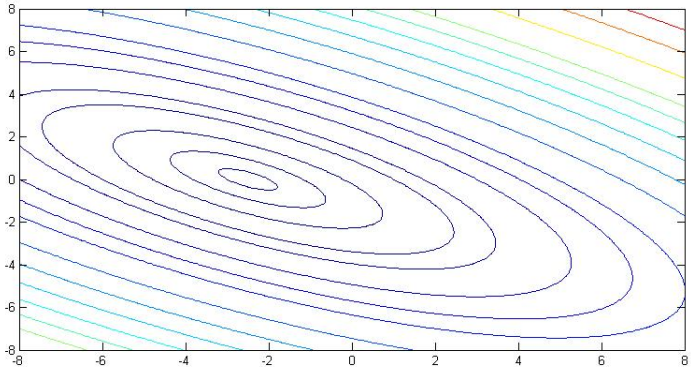

Figura 2.4: Curva de nível de f com matriz Hessiana não diagonal.

$$d^{k+1} = -\nabla f(x^{k+1}) + \beta_k d^k \tag{2.14}$$

com x^{k+1} dado pela relação (2.4) para $k = 0, 1, ..., n-1$. O parâmetro β_k é escolhido de modo que duas direções consecutivas d^k e d^{k+1} sejam A-conjugadas. Como

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

de onde

$$\beta_k = \frac{(d^k)^T A \nabla f(x^{k+1})}{(d^k)^T A d^k}. \tag{2.15}$$

Vamos agora apresentar o Algoritmo dos Gradientes Conjugados:

Algoritmo 2.6 *Algoritmo GC*

Dado: $x^0 \in \mathbb{R}^n$
$k = 0$
$d^0 = -\nabla f(x^0)$
REPITA enquanto $\nabla f(x^k) \neq 0$
    Calcule $\alpha_k = argmin f(x^k + \alpha d^k)$
    $x^{k+1} = x^k + \alpha_k d^k$
    Avalie $\nabla f(x^{k+1})$
    Calcule β_k
    $d^{k+1} = -\nabla f(x^{k+1}) + \beta_k d^k$
    $k = k + 1$

Ressaltamos que o Algoritmo GC está bem definido, ou seja, se $\nabla f(x^k) \neq 0$ então $d^k \neq 0$ e o novo ponto pode ser calculado. De fato, por (2.8) obtemos que

$$\nabla f(x^k)^T d^k = \nabla f(x^k)^T \left(-\nabla f(x^k) + \beta_{k-1} d^{k-1}\right) = -\| \nabla f(x^k) \|^2 \neq 0,$$

portanto $d^k \neq 0$.

O próximo teorema mostra que as direções geradas pelo Algoritmo GC são, de fato, A-conjugadas e que os gradientes são ortogonais.

Teorema 2.7 *Seja x^{k+1} definido por (2.4) e $\{d^0, \ldots, d^k\}$ um conjunto de direções geradas pelo Algoritmo. Então*

$$\nabla f(x^{k+1})^T \nabla f(x^j) = 0$$

e

$$(d^{k+1})^T A d^j = 0,$$

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

Demonstração. Denotemos $g_i = \nabla f(x^i)$. Vamos provar por indução em k. Para $k = 1$, pela relação (2.8), obtemos $g_1^T g_0 = -g_1^T d^0 = 0$. E pela definição de β_0 temos que $(d^1)^T A d^0 = 0$. Suponhamos que vale para k, queremos provar que vale para $k+1$. Usando a definição de d^{k+1} e a hipótese de indução que as direções $d^0, \ldots, d^k$ são A-conjugadas, concluímos que $g_{k+1}^T d^j = 0$ para $j = 0, \ldots, k$. Dessa forma,

$$g_{k+1}^T g_j = g_{k+1}^T(-d^j + \beta_{j-1} d^{k-1}) = 0. \tag{2.16}$$

Pela definição de β_k, temos que $(d^{k+1})^T A d^k = 0$. Além disso, para $j < k$, a hipótese de indução nos fornece

$$(d^{k+1})^T A d^j = (-g_{k+1} + \beta_k d^k)^T A d^j = -g_{k+1} A d^j.$$

Usando a Definição (2.1) e o que foi estabelecido pela relação (2.5), obtemos

$$(d^{k+1})^T A d^j = -g_{k+1}^T \left(\frac{g_{j+1} - g_j}{t_j}\right) = 0.$$

□

Variantes dos métodos de gradientes conjugados

O Lema a seguir mostra outras formas de calcular o parâmetro β_k que apresenta grande importância teórica e prática, pois não requer o cálculo da Hesssiana da função f.

Lema 2.8 *Seja β_k definido em 2.15. Então*

$$\beta_k = \frac{\nabla f(x^{k+1})^T \left(\nabla f(x^{k+1}) - \nabla f(x^k)\right)}{\nabla f(x^k)^T \nabla f(x^k)} \tag{2.17}$$

$$\beta_k = \frac{\nabla f(x^{k+1})^T \nabla f(x^{k+1})}{\nabla f(x^k)^T \nabla f(x^k)} \tag{2.18}$$

Demonstração. Por (2.5)

$$Ad^k = \frac{\nabla f(x^{k+1}) - \nabla f(x^k)}{\alpha_k}.$$

Substituindo na expressão de β_k, temos que

$$\begin{aligned} \beta_k &= \frac{\nabla f(x^{k+1})^T Ad^k}{(d^k)^T Ad^k} \\ &= \frac{\nabla f(x^{k+1})^T \left(\dfrac{\nabla f(x^{k+1}) - \nabla f(x^k)}{\alpha_k}\right)}{(d^k)^T \left(\dfrac{\nabla f(x^{k+1}) - \nabla f(x^k)}{\alpha_k}\right)}. \end{aligned}$$

Usando 2.8 obtemos

$$\begin{aligned} \beta_k &= \frac{\nabla f(x^{k+1})^T \left(\nabla f(x^{k+1}) - \nabla f(x^k)\right)}{(d^k)^T \nabla f(x^{k+1}) - (d^k)^T \nabla f(x^k)} \\ &= \frac{\nabla f(x^{k+1})^T \left(\nabla f(x^{k+1}) - \nabla f(x^k)\right)}{-(d^k)^T \nabla f(x^k)}. \end{aligned}$$

Verifiquemos que $-(d^k)^T \nabla f(x^k) = \nabla f(x^k)^T \nabla f(x^k)$. De fato, usando a definição de d^k dada em (2.14) e (2.5), temos

$$\begin{aligned} -(d^k)^T \nabla f(x^k) &= -(-\nabla f(x^k) + \beta_{k-1} d^{k-1})^T \nabla f(x^k) \\ &= \nabla f(x^k)^T \nabla f(x^k) - \beta_{k-1} (d^{k-1})^T \nabla f(x^k) \\ &= \nabla f(x^k)^T \nabla f(x^k). \end{aligned}$$

Portanto

$$\beta_k = \frac{\nabla f(x^{k+1})^T \left(\nabla f(x^{k+1}) - \nabla f(x^k)\right)}{\nabla f(x^k)^T \nabla f(x^k)}. \tag{2.19}$$

provando (2.17). A relação (2.18) segue de (2.17) e do fato que $\nabla f(x^{k+1})$ é ortogonal a $\nabla f(x^k)$, como provado no Teorema (3.6). □

A expressão (2.17) é devida a Polak e Ribière (GCPR) e a fórmula (2.18) é devida a Fletcher e Reeves (GCFR). As expresõas (2.15), (2.17) e (2.18) coincidem para funções quadráticas como provado no lema anterior. No entanto, serão úteis quando estendermos

o método dos gradientes conjugados para funções não quadráticas, conforme discutiremos na próxima seção.

2.4 Generalização para funções não quadráticas

Nesta seção estenderemos o algoritmo estudado na seção anterior para minimização de funções não quadráticas, os quais são obtidos através de simples mudanças no Algoritmo GC. Observe que as fórmulas (2.17) e (2.18), devidas a Polak e Ribière e a Fletcher e Reeves para o cálculo do parâmetro β_k independem da matriz A. O passo α_k é definido como $argmin f(x^k + \alpha d^k)$.

No caso de funções quadráticas, o passo α_k é calculado de forma explícita pela fórmula (2.7). Para funções não quadráticas faremos uma busca linear ao longo da direção d^k, sendo que esta busca pode ser pelo método da Seção Áurea ou pela busca de Armijo. Além disso, devemos considerar o gradiente da função em questão.

2.4.1 Buscas unidirecionais

Uma da idéias fundamentais para a construção de um algoritmo é escolher, a partir de cada ponto obtido uma direção para dar o próximo passo. Nesse caso, a possibilidade mais razoável para essa escolha é determinar uma direção em que f decresce.

Definição 2.9 *Considere uma função $f : \mathbb{R}^n \to \mathbb{R}$, um ponto $x^k \in \mathbb{R}^n$ e uma direção de descida $d^k \in \mathbb{R}^n$. Dizemos que d^k é uma direção de descida para f, apartir de x^k, quando existe $\delta > 0$ tal que*

$$f(x^k + \alpha d^k) < f(x^k),$$

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

Como nem sempre podemos suceder com esse tipo de avaliação de função, apresentaremos uma condição suficiente para que uma direção seja de descida.

Teorema 2.10 *Se $\nabla f(x^k)^T d^k < 0$, então d^k é uma direção de descida para f, a partir de x^k.*

Demonstração. Temos por hipótese que

$$0 > \nabla f(x^k)^T d^k = \frac{\partial f}{\partial d^k}(x^k) = \lim_{\alpha \to 0} \frac{f(x^k + \alpha d^k) - f(x^k)}{\alpha}$$

e pela preservação do sinal, existe $\delta > 0$ tal que

$$\frac{f(x^k + \alpha d^k) - f(x^k)}{\alpha} < 0,$$

para todo $\alpha \in (-\delta, \delta)$, $\alpha \neq 0$. Portanto

$$f(x^k + \alpha d^k) < f(x^k)$$

para todo $\alpha \in (0, \delta)$, completando assim a demonstração. □

Busca exata

Neste caso, a busca na direção é feita de forma exata. Discutimos o método de Seção Áurea. Para definirmos este método precisamos do conceito de função unimodal.

Definição 2.11 *Uma função contínua $\varphi : [0, \infty) \to \mathbb{R}$ é dita unimodal quando admite um conjunto de minimizadores $[t_1, t_2]$, e é estritamente decrescente em $[0, t_1]$ e estritamente crescente em $[t_2, \infty)$.*

Alguns exemplos de funções unimodais são mostrados nas figuras abaixo. Notemos que o intervalo $[t_1, t_2]$ pode ser degenerado.

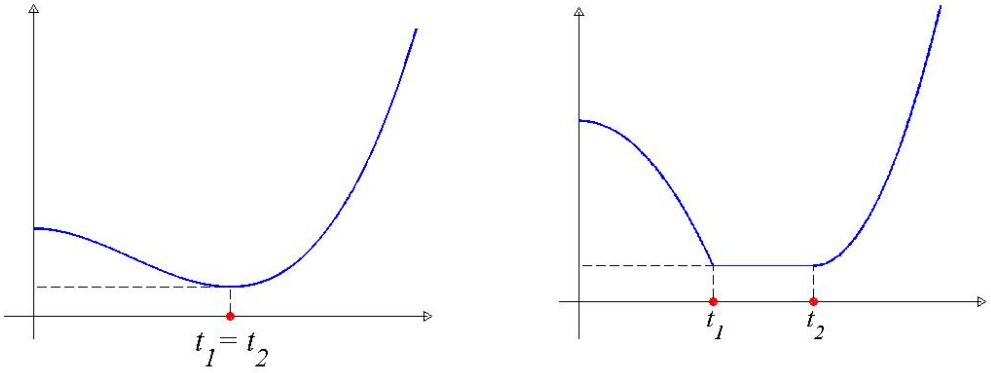

Figura 2.5: Exemplos de função unimodal.

Para sabermos como funciona o algoritmo, observemos a Fig.2.6 e os seguintes itens.

Suponhamos que o minimizador da função φ pertence ao intervalo $[a, b]$.

1. Consideremos $a < u < v < b$ no intervalo $[0, \infty)$.

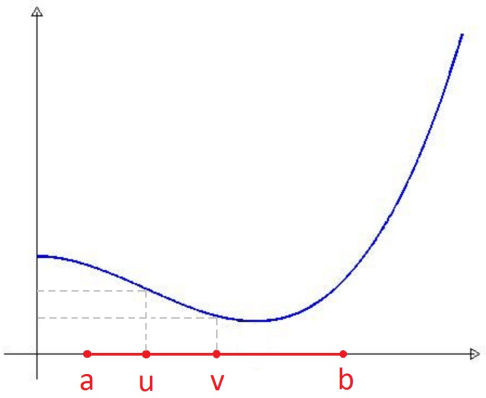

Figura 2.6: Ilustração do método da Seção Áurea.

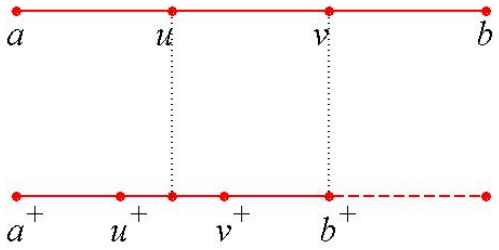

Figura 2.7: Divisão de $[a, b]$.

2. Se $\varphi(u) < \varphi(v)$ então o trecho $(v\ ,\ b]$ não pode conter o minimizador e portanto podemos descartá-lo.
3. Se $\varphi(u) \geq \varphi(v)$ então o trecho $[a, u)$ não pode conter o minimizador e portanto podemos descartá-lo.
4. Particionamos o intervalo que sobrou e repetimos o processo.

Denotaremos o novo intervalo particionado por u^+ e v^+ de $[a^+, b^+]$.

O critério relevante nesse método é a escolha do posicionamento de u e v. À primeira vista, o que é mais óbvio é dividir o intervalo em três partes iguais, conseguindo assim descartar 33,3% do mesmo a cada iteração como ilustrado na Fig. 2.7. No entanto, conseguiremos um resultado mais eficaz se escolhermos u e v dividindo o segmento $[a, b]$ na razão áurea, que será apresentada na próxima definição.

Definição 2.12 *Um ponto c divide o segmento $[a, b]$ na razão áurea quando a razão entre o maior segmento e o segmento todo é igual à razão entre o menor e o maior dos segmentos. Tal razão é conhecida como número de ouro e vale* $\dfrac{\sqrt{5}-1}{2} \approx 0.618$.

Assim, temos que u e v devem satisfazer

$$\frac{b-u}{b-a} = \frac{u-a}{b-u} \qquad \text{e} \qquad \frac{v-a}{b-a} = \frac{b-v}{v-a}. \tag{2.20}$$

Considerando θ_1 e θ_2 tais que

$$u = a + \theta_1(b-a) \qquad \text{e} \qquad v = a + \theta_2(b-a),$$

substituindo u na primeira expressão de (2.13) temos

$$\frac{b-(a+\theta_1(b-a))}{b-a} = \frac{a+\theta_1(b-a)-a}{b-(a+\theta_1(b-a))}$$

que pode ser reescrita como

$$\frac{(b-a)(1-\theta_1)}{b-a} = \frac{\theta_1(b-a)}{(b-a)(1-\theta_1)},$$

assim obtemos

$$(1-\theta_1) = \frac{\theta_1}{(1-\theta_1)}.$$

Analogamente, substituindo v na primeira expressão de (2.13) temos que:

$$\theta_2 = \frac{1-\theta_2}{\theta_2}.$$

Resolvendo esta equação e notando que θ_1 e $\theta_2 \in [0, 1]$, temos que: $\theta_1 = \dfrac{3-\sqrt{5}}{2} \approx 0,382$ e $\theta_2 = \dfrac{\sqrt{5}-1}{2} \approx 0,618$. Cabe notar que θ_1 e θ_2 satisfazem as seguintes relações:

$$(\theta_2)^2 = \theta_1 \qquad \text{e} \qquad \theta_1 + \theta_2 = 1. \tag{2.21}$$

Se dividirmos o intervalo segundo a razão áurea, conseguimos descartar cerca de 38% do intervalo $[a, b]$ a cada iteração e diminuimos o número de avaliações de função, pois além disso reaproveitamos um dos pontos u ou v e a avaliação de $\varphi(u)$ ou $\varphi(v)$, conforme veremos nos resultados seguintes. Se o ponto v é aproveitado na próxima etapa, ele passa a ser u^+. Para u temos $v^+ = u$.

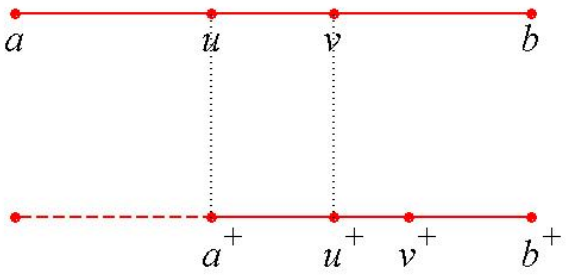

Figura 2.8: $u^+ = v$.

Lema 2.13 *Na seção áurea, se* $[a, u)$ *é descartado então* $u^+ = v$.

Demonstração. Como $[a, u)$ foi descartado então $a^+ = u$ e $b^+ = b$. Logo, temos

$$\begin{aligned} u^+ &= a^+ + \theta_1(b^+ - a^+) \\ &= u + \theta_1(b - u) \\ &= a + \theta_1(b - a) + \theta_1(b - (a + \theta_1(b - a))) \\ &= a + (2\theta_1 - (\theta_1)^2)(b - a) \end{aligned}$$

Usando as relações (2.21), temos $(\theta_1)^2 = 3\theta_1 - 1$. Então

$$\begin{aligned} u^+ &= a + (2\theta_1 - 3\theta_1 + 1)(b - a) \\ &= a + (1 - \theta_1)(b - a) \\ &= a + \theta_2(b - a) = v. \end{aligned}$$

□

A Fig.(2.8) ilustra o Lema (2.13).

Lema 2.14 *Na seção áurea, se* $(v, b]$ *é descartado então* $v^+ = u$.

Demonstração. Como $(v, b]$ é descartado então $a^+ = a$ e $b^+ = v$. Usando as relações (2.21) obtemos:

$$\begin{aligned} v^+ &= a^+ + \theta_2(b^+ - a^+) = a + \theta_2(v - a) \\ &= a + \theta_2(a + \theta_2(b - a) - a) \\ &= a + \theta_2^2(b - a) \\ &= a + \theta_1(b - a) = u. \end{aligned}$$

□

Algoritmo 2.15 *Seção Áurea*

Dados: $\rho > 0$; $\epsilon > 0$
Fase 1: Obtenção do intervalo $[a, b]$
$a_0 = 0$, $s_0 = \rho$ e $b_0 = 2\rho$
$k = 0$
REPITA enquanto $\varphi(b_k) < \varphi(s_k)$
    $a_{k+1} = s_k$, $s_{k+1} = b_k$ e $b_{k+1} = 2b_k$
    $k = k + 1$
$a = a_k$, $b = b_k$
Fase 2: Obtenção de $\alpha \in [a, b]$
$a_0 = a$, $b_0 = b$
$u_0 = a_0 + \theta_1(b_0 - a_0)$, $v_0 = a_0 + \theta_2(b_0 - a_0)$
$k = 0$
REPITA enquanto $b_k - a_k > \epsilon$
    SE $\varphi(u_k) < \varphi(v_k)$
        $a_{k+1} = a_k$, $b_{k+1} = v_k$, $v_{k+1} = u_k$, $u_{k+1} = a_{k+1} + \theta_1(b_{k+1} - a_{k+1})$
    SENÃO
        $a_{k+1} = u_k$, $b_{k+1} = b_k$, $u_{k+1} = v_k$, $v_{k+1} = a_{k+1} + \theta_2(b_{k+1} - a_{k+1})$
    $k = k + 1$
Defina $\alpha = \dfrac{u_k + v_k}{2}$

Busca inexata

Abrindo mão da pretenção de obter um minimizador exato na direção dada, discutamos o critério de Armijo que se satisfaz com um decréscimo suficiente na função.

Consideremos um ponto $x^k \in \mathbb{R}^n$, uma direção de descida $d^k \in \mathbb{R}^n$ e $\eta \in (0, 1)$. A regra de Armijo consiste em encontrar α_k tal que

$$f(x^k + \alpha_k d^k) \leq f(x^k) + \eta \alpha_k \nabla f(x^k)^T d^k.$$

o que significa que teremos uma redução no valor da função e, além disso, que ela será proporcional ao comprimento do passo. O próximo teorema garante esse resultado.

Teorema 2.16 *Considere uma função diferenciável $f : \mathbb{R}^n \to \mathbb{R}$, um ponto $x^k \in \mathbb{R}^n$, uma direção de descida $d^k \in \mathbb{R}^n$ e $\eta \in (0, 1)$. Então existe $\delta > 0$ tal que*

$$f(x^k + \alpha d^k) \leq f(x^k) + \eta \alpha \nabla f(x^k)^T d^k,$$

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

Demonstração. Se temos $\nabla f(x^k)^T d^k = 0$, o resultado segue da definição de direção de descida. Suponhamos agora que $\nabla f(x^k)^T d^k < 0$. Assim, como $\eta < 1$, temos

$$\lim_{\alpha \to 0} \frac{f(x^k + \alpha d^k) - f(x^k)}{\alpha} = \nabla f(x^k)^T d^k < \eta \nabla f(x^k)^T d^k.$$

Portanto, existe $\delta > 0$ tal que

$$\frac{f(x^k + \alpha d^k) - f(x^k)}{\alpha} < \eta \nabla f(x^k)^T d^k,$$

para todo $\alpha \in (0, \delta)$. Logo

$$f(x^k + \alpha d^k) \leq f(x^k) + \eta \alpha \nabla f(x^k)^T d^k,$$

completando assim a demonstração. □

Para entendermos melhor essa condição, consideremos a função $\varphi : [0, \infty) \to \mathbb{R}$ dada por

$$\varphi(\alpha) = f(x^k + \alpha d^k).$$

A aproximação de primeira ordem dessa função em torno de $\alpha = 0$ é

$$p(\alpha) = \varphi(0) + \alpha \varphi'(0) = f(x^k) + \alpha \nabla f(x^k)^T d^k.$$

Assim a definição de busca de Armijo pode ser reescrita da forma

$$\varphi(0) - \varphi(\alpha) = f(x^k) - f(x^k + \alpha d^k) \geq \eta(p(0) - p(\alpha)).$$

Isto significa que queremos um passo cuja redução seja pelo menos uma fração η da redução que conseguimos no modelo linear. Fato este ilustrado na Fig. 2.9.

Notemos que a condição de Armijo é satisfeita para os pontos em que φ está abaixo de q, onde

$$q(\alpha) = f(x^k) + \eta \alpha \nabla f(x^k)^T d^k.$$

Algoritmo 2.17 *Busca de Armijo*

Dados: $x^k \in \mathbb{R}^n$; $d^k \in \mathbb{R}^n$ (direção de descida), $\eta \in (0, 1)$
$t = 1$
REPITA enquanto $f(x^k + \alpha d^k) > f(x^k) + \eta \alpha \nabla f(x^k)^T d^k$
    $t = 0,8t$

Porém, vale ressaltar que além de usar a busca de Armijo [6] precisamos de mais

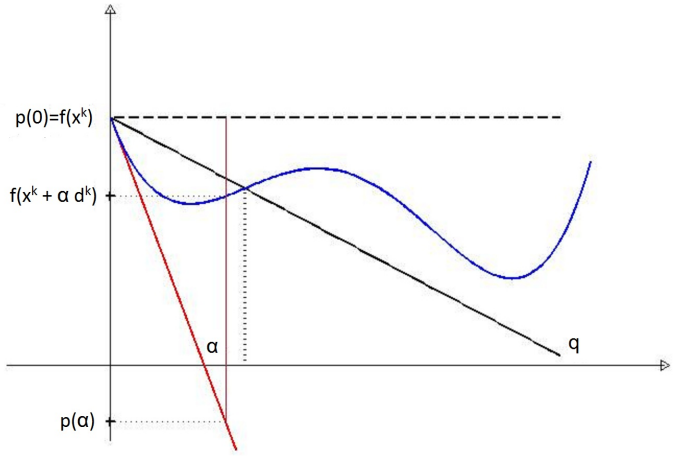

Figura 2.9: Armijo.

uma condição para garantir que d^k é uma direção de decréscimo da função dada pela condição forte de Wolfe, que veremos mais adiante. Para mais detalhes, consulte [13].

Para garantir o bom desempenho do algoritmo, precisamos ser mais precisos sobre a escolha da busca linear do parâmetro α_k, porque de modo geral a direção d^k pode não ser de descida.

Consideremos $d^k = -\nabla f(x^k) + \beta_{k-1} d^{k-1}$ dada no algoritmo e façamos o produto interno com o gradiente da função f em x^k

$$\nabla f(x^k)^T d^k = \nabla f(x^k)^T(-\nabla f(x^k) + \beta_{k-1} d^{k-1})$$
$$\nabla f(x^k)^T d^k = - \parallel \nabla f(x^k) \parallel^2 + \beta_{k-1} \nabla f(x^k)^T d^{k-1}$$

Se a busca linear é exata, ou seja, se α_{k-1} é o minimizador de f ao longo da direção d^{k-1}, então $\nabla f(x^k)^T d^{k-1} = 0$. Neste caso, se $\nabla f(x^k)^T d^k < 0$, então d^k é uma direção de descida. Mas se a busca linear não é exata, podemos ter:

$$\nabla f(x^k)^T d^k = - \parallel \nabla f(x^k) \parallel^2 + \beta_{k-1} \nabla f(x^k)^T d^{k-1}.$$

Mas também podemos ter que $\nabla f(x^k)^T d^k > 0$, implicando que d^k é uma direção de crescimento da função. Podemos evitar isso fazendo com que α_k satisfaça as condições fortes de Wolfe, dadas por

$$f(x^k + \alpha_k d^k) \leq f(x^k) + c_1 \alpha_k \nabla f(x^k)^T d^k \qquad \text{e} \qquad \mid \nabla f(x^k + \alpha_k d^k)^T d^k \mid \leq c_2 \mid \nabla f(x^k)^T d^k \mid \tag{2.22}$$

onde $0 < c_1 < c_2 < \frac{1}{2}$. A primeira desigualdade é o critério de Armijo com $\eta = c_1$.

Lema 2.18 *Considere o Algoritmo GC com o β_k de Fletcher-Reeves e α_k satisfazendo a condição forte de Wolfe* (2.22)*. Então o método gera direções de descida d^k que satisfazem a seguinte desigualdade:*

$$-\frac{1}{1-c_2} \leq \frac{\nabla f(x^k)^T d^k}{\| \nabla f(x^k) \|^2} \leq \frac{2c_2-1}{1-c_2}, \tag{2.23}$$

para todo $k = 0, 1, ...$

Demonstração. Notemos que a função $t : \mathbb{R} \to \mathbb{R}$ definida por $t(\xi) := \frac{(2\xi-1)}{(1-\xi)}$ é monótona crescente no intervalo $[0, \frac{1}{2}]$ e que $t(0) = -1$ e $t(\frac{1}{2}) = 0$. No entanto, pelo fato de $c_2 \in (0, \frac{1}{2})$, nós temos:

$$-1 < \frac{2c_2-1}{1-c_2} < 0. \tag{2.24}$$

A condição de descida $\nabla f(x^k)^T d^k \leq 0$ segue imediatamente uma vez estabelecida a condição (2.23).

A prova é por indução. Para $k = 0$, o termo médio na expressão anterior é -1, então usando (2.24), podemos ver que as duas desigualdades em (2.23) são satisfeitas. Assumamos que (2.23) é assegurada par algum $k \geq 1$. Da definição de β_k e d^{k+1} no Algoritmo GC-FR, temos:

$$\frac{\nabla f(x^{k+1})^T d^{k+1}}{\|\nabla f(x^{k+1})\|^2} = -1 + \beta_k \frac{\nabla f(x^{k+1})^T d^k}{\|\nabla f(x^{k+1})\|^2} = -1 + \frac{\nabla f(x^{k+1})^T d^k}{\|\nabla f(x^k)\|^2}. \tag{2.25}$$

Usando a condição de busca linear $| \nabla f(x^k + \alpha_k d^k)^T d^k | \leq c_2 | \nabla f(x^k)^T d^k |$, temos que $| \nabla f(x^{k+1})^T d^k | \leq -c_2 \nabla f(x^k)^T d^k$, de onde, combinando com (2.25), obtemos

$$-1 + c_2 \frac{\nabla f(x^k)^T d^k}{\|\nabla f(x^k)\|^2} \leq \frac{\nabla f(x^{k+1})^T d^{k+1}}{\|\nabla f(x^{k+1})\|^2} \leq -1 - c_2 \frac{\nabla f(x^k)^T d^k}{\|\nabla f(x^k)\|^2}. \tag{2.26}$$

Que substituindo pelo termo $\frac{\nabla f(x^k)^T d^k}{\|\nabla f(x^k)\|^2}$ à esquerda da hipótese de indução (2.23) nós obtemos:

$$-1 - \frac{c_2}{1-c_2} \leq \frac{\nabla f(x^{k+1})^T d^{k+1}}{\|\nabla f(x^{k+1})\|^2} \leq -1 + \frac{c_2}{1-c_2},$$

que nos mostra que (2.23) é assegurada para $k+1$ como queremos. □

Aplicando o lema a $| \nabla f(x^k + \alpha_k d^k)^T d^k | \leq c_2 | \nabla f(x^k)^T d^k |$ veremos que $\nabla f(x^k)^T d^k = - \| \nabla f(x^k) \|^2 + \beta_{k-1} \nabla f(x^k)^T d^{k-1}$ é negativa, e concluímos que qualquer procedimento de busca linear que forneça α_k satisfazendo:

$$f(x^k + \alpha_k d^k) \leq f(x^k) + c_1 \alpha_k \nabla f(x^k)^T d^k \qquad \text{e} \qquad | \nabla f(x^k + \alpha_k d^k)^T d^k | \leq c_2 | \nabla f(x^k)^T d^k | \tag{2.27}$$

irá assegurar que todas as direções d^k são direções de descida para a função f.

Quando f é uma função não quadrática fortemente convexa e a busca linear é exata, desde que os gradientes sejam ortogonais, os parâmetros β_k de GCPR e GCFR são iguais. Quando aplicados para funções não quadráticas com busca linear inexata, as propriedades dos dois algoritmos diferem fortemente. Experiências numéricas indicam que o Algoritmo GCPR tende a ser o mais eficiente entre os dois. Um fato supreendente sobre o Algoritmo GCPR é que a condição forte de Wolfe não garante que d^k seja sempre uma direção de descida. Se definirmos o parâmetro β_k como

$$\beta_k^+ = \{ \max\{\beta_k, 0\}\}$$

então uma simples adaptação da forte condição de Wolfe assegura que as propriedades de descida são mantidas.

Implementações do método dos gradientes conjugados para uma função qualquer geralmente possuem propriedades semelhantes às do método dos gradientes conjugados para uma função quadrática. Uma interpolação quadrática (ou cúbica) ao longo da direção de busca d^k está incorporada no processo de busca linear. Esta característica garante que quando f é uma função quadrática estritamente convexa, o comprimento do passo α_k é escolhido exatamente como sendo o minimizador unidimensional, então o método do gradiente conjugado para funções não quadráticas se reduz ao método para funções quadráticas.

Outra modificação que é frequentemente usada em procedimentos de gradientes conjugados para funções não quadráticas é reiniciar o algoritmo a cada n etapas definindo $\beta_k = 0$, tomando assim uma direção de maior descida. Este processo serve para refrescar o algoritmo, apagando informações velhas que podem não ser mais úteis. É possível até mesmo provar um forte resultado teórico: a reinicialização leva à convergência quadrática na n-ésima etapa, ou seja:

$$\| x_{k+n} - x \| = O(\| x_k - x^* \|^2).$$

Com um pouco de análise, podemos ver que este resultado não é tão surpreendente. Consideremos uma função quadrática que seja fortemente convexa numa vizinhança da solução e não seja quadrática em todos os outros pontos do domínio. Assumindo que o algoritmo é convergente para a solução em questão, as iterações irão eventualmente entrar na região quadrática. Em algum ponto o algoritmo será reiniciado naquela região, e daquele ponto em diante, suas características serão simplesmente aquelas do método de gradiente conjugado para funções quadráticas. Em particular, terminação finita irá ocorrer dentro de n passos de reinício. O reinício é importante, porque as propriedades de terminação finita do algoritmo se mantém somente quando a direção de busca inicial d^0 é igual ao oposto do gradiente.

Ainda que f não seja exatamente quadrática perto da solução, o Teorema de Taylor implica que ela pode ser aproximada muito de perto por uma quadrática. No entanto, mesmo não esperando terminação em n passos depois do reinício, não é surpreendente que um progresso substancial é feito em direção a solução, como indicado por:

$$\| x_{k+n} - x \| = O(\| x_k - x^* \|^2).$$

Embora este último resultado seja interessante de um ponto de vista teórico, ele pode não ser relevante num contexto prático, porque métodos de gradientes conjugados para funções não quadráticas podem ser recomendados somente para resolver problemas de dimensão grande (n grande). Em tais problemas o reinício pode nunca acontecer, desde que uma solução aproximada seja encontrada em menos que n passos.

Capítulo 3

Complexidade Algorítmica

Neste capítulo discutimos a minimização de uma função quadrática em subespaços com o intuito de provar que os métodos de gradiente conjugados têm complexidade ótima da ordem $O\left(\frac{1}{k^2}\right)$ em termos do valor da função.

3.1 Minimização em um subespaço

Nesta seção veremos como minimizar num subespaço do $\mathbb{R}^n$ uma função quadrática $f : \mathbb{R}^n \to \mathbb{R}$ definida por

$$f(x) = \frac{1}{2}x^T Ax + b^T x + c \tag{3.1}$$

onde $A \in \mathbb{R}^{n\times n}$ é uma matriz definida positiva, $b \in \mathbb{R}^n$ e $c \in \mathbb{R}$. A principal referência desta seção é [3].

Iremos agora fornecer algumas definições essenciais para o desenvolvimento desta seção.

Considere V um subespaço vetorial do $\mathbb{R}^n$. Seja $S \in \mathbb{R}^{n\times k}$ uma matriz cujas colunas formam uma base para V, onde k é a dimensão do subespaço V. Denote as colunas de S por $d^0, d^1, ..., d^{k-1}$.

Definição 3.1 *Dado um ponto x e um subespaço vetorial V, a variedade linear paralela a V e que contém x é o conjunto de todos os pontos $x + Ss$ tais que $Ss \in V$.*

Portanto, se $y \in x + Ss$ ele pode ser escrito da forma

$$y = x + s_0 d^0 + s_1 d^1 + \ldots + s_{k-1} d^{k-1}, \tag{3.2}$$

onde $d^0, d^1, ..., d^{k-1}$ são as k colunas da matriz S.

A figura a seguir ilustra um exemplo desse tipo de minimização em dois subespaços S e T de dimensão um.

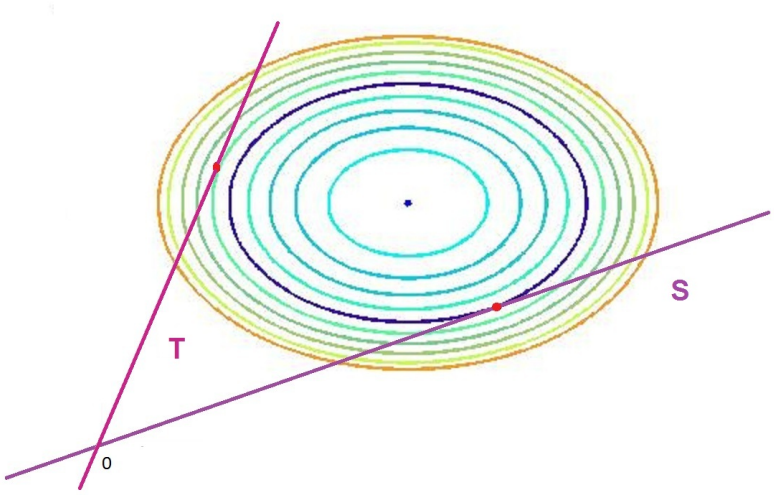

Figura 3.1: Minimização em dois subespaços de dimensão 1.

Vamos agora considerar o problema de minimizar f numa sequência de variedades de dimensão crescente. Lembrando que se a dimensão final for igual a n significa que estamos minimizando a função no $\mathbb{R}^n$. Para isso provemos o seguinte resultado.

Teorema 3.2 *Seja f a função quadrática com Hessiana $A \in \mathbb{R}^{n\times n}$, $\bar{x} \in \mathbb{R}^n$ definida em (2.2), V um subespaço de $\mathbb{R}^n$ e suponha que S^TAS é não-singular. Então o ponto crítico de f na variedade $\bar{x} + Ss$ é dado por*

$$x^+ = \bar{x} - S(S^TAS)^{-1}S^T\nabla f(\bar{x}) \tag{3.3}$$

e satisfaz $S^T\nabla f(x^+) = 0$. Se S^TAS é definida positiva, x^+ é o minimizador de f na variedade $\bar{x} + Ss$.

Demonstração. Considere a função $\tau : \mathbb{R}^k \to \mathbb{R}$ definida por $\tau(s) = f(\bar{x} + Ss)$. Então

$$\begin{aligned} \tau'(s) &= S^T\nabla f(\bar{x} + Ss) = S^T[A(\bar{x} + Ss) + b] \\ &= S^T(A\bar{x} + ASs + b) \\ &= S^T\nabla f(\bar{x}) + S^TASs. \end{aligned} \tag{3.4}$$

Seja s^+ um minimizador de τ. Assim $\tau'(s^+) = 0$. Como S^TAS é não-singular existe $(S^TAS)^{-1}$ e consequentemente

$$s^+ = -(S^TAS)^{-1}S^T\nabla f(\bar{x}).$$

Portanto, o minimizador x^+ de f na variedade $\bar{x} + Ss$ é

$$x^+ = \bar{x} - S(S^TAS)^{-1}S^T\nabla f(\bar{x}). \tag{3.5}$$

que por (3.4) satisfaz $S^T\nabla f(x^+) = S^T\nabla f(\bar{x} + Ss^+) = 0$. Além disso, se S^TAS é definida positiva, $\tau''(s) = S^TAS > 0$, donde concluímos que τ é uma função quadrática fortemente convexa. Como consequência, s^+ é o minimizador da τ e x^+ é o minimizador global de f na variedade $\bar{x} + Ss$, completando a demonstração. □

Como $S^T\nabla f(x^+) = 0$ temos que

$$(d^j)^T\nabla f(x^+) = 0, \tag{3.6}$$

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

Consideremos agora uma sequência de subespaços V_k de dimensão $k+1$; com $k = 0, \ldots, n-1$. Dado $\bar{x} \in \mathbb{R}^n$, considere o minimizador na variedade $\bar{x} + S_k s$ o ponto

$$x^{k+1} = \bar{x} - S_k^T(S_k^TAS_k)^{-1}S_k^T\nabla f(\bar{x}). \tag{3.7}$$

Seja x^k o minimizador na variedade $\bar{x} + S_{k-1}s$, então podemos escrever x^k da forma

$$x^k = \bar{x} + s_0d^0 + s_1d^1 + \ldots + s_{k-1}d^{k-1} = \bar{x} + s_0d^0 + s_1d^1 + \ldots + s_{k-1}d^{k-1} + 0d^k, \tag{3.8}$$

com isso concluímos que $x^k \in \bar{x} + S_{k-1}s \subset \bar{x} + S_k s$ e também que

$$V_0 \subset V_1 \subset \ldots \subset V_k. \tag{3.9}$$

Portanto, se $x^k \in \bar{x} + S_{k-1}s$ e $\bar{x} + S_{k-1}s \subset \bar{x} + S_k s$, então $x^k \in \bar{x} + S_k s$.

Vamos agora provar que se dois pontos pertencem à mesma variedade então podemos escrever a variedade em função de qualquer um desses pontos.

Lema 3.3 *Se $\bar{x}$ e x^k pertencentem à variedade linear paralela ao subespaço V_k, então $\bar{x} + S_k s = x^k + S_k s$.*

Demonstração. Considere $\bar{x}$ e x^k pertencentes à variedade $\bar{x} + S_k s$. Logo $x^k = \bar{x} + S_k\hat{s}$ para algum $\hat{s}$. Considere y arbitrário na variedade $\bar{x} + S_k s$, logo, para algum $\bar{s}$,

$$\begin{aligned} y &= \bar{x} + S_k\bar{s} \\ &= x^k + -S_k\hat{s} + S_k\bar{s} \\ &= x^k + S_k(-\hat{s} + \bar{s}) \\ &= x^k + S_k s^*, \end{aligned}$$

e portanto y pertence à variedade $x^k + S_k s$. Analogamente se y pertence à variedade $x^k + S_k s$,

$$\begin{aligned} y &= x^k + S_k \bar{s} \\ &= \bar{x} + S_k \hat{s} + S_k \bar{s} \\ &= \bar{x} + S_k(\hat{s} + \bar{s}) \\ &= \bar{x} + S_k s^*, \end{aligned}$$

o que prova que x pertence à variedade $\bar{x} + Ss$. Portanto, podemos concluir que $\bar{x} + S_k s = x^k + S_k s$. □

Pelo Teorema (3.3) e Lema (3.3), temos que o minimizador de f na variedade $x^k + S_k s$ é da forma

$$x^{k+1} = x^k - S_k(S_k^T A S_k)^{-1} S_k^T \nabla f(x^k). \tag{3.10}$$

Pelo fato que

$$S_{k-1}^T \nabla f(x^k) = \begin{bmatrix} d^{0^T} \\ d^{1^T} \\ \vdots \\ d^{k-1^T} \end{bmatrix} \cdot \begin{bmatrix} \nabla f(x^k) \end{bmatrix} = 0,$$

temos que

$$S_k^T \nabla f(x^k) = \begin{bmatrix} d^{0^T} \\ d^{1^T} \\ \vdots \\ d^{k-1^T} \\ d^{k^T} \end{bmatrix} \cdot \begin{bmatrix} \nabla f(x^k) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ d^{k^T} \nabla f(x^k) \end{bmatrix} = \left((d^k)^T \nabla f(x^k)\right) e^k$$

onde

$$e_k = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \in \mathbb{R}^{k+1}.$$

Substituindo em (3.10), temos que

$$x^{k+1} = x^k - \left((d^k)^T \nabla f(x^k)\right) S_k(S_k^T A S_k)^{-1} e_k.$$

3.2 Minimização nos espaços de Krylov

Dada uma matriz $A \in \mathbb{R}^{n\times n}$ simétrica e $x^0 \in \mathbb{R}^n$, definimos o k-ésimo espaço de Krylov por

$$\kappa_k = \left[Ax^0, A^2x^0, ..., A^kx^0\right].$$

Nesta seção discutimos a minimização da função quadrática (3.1) na variedade $V_k = x^0 + \kappa_k$. Para simplificar a notação consideremos novamente a função $f : \mathbb{R}^n \to \mathbb{R}$ definida por

$$f(x) = \frac{1}{2}x^TAx + b^Tx + c,$$

com $A \in \mathbb{R}^{n\times n}$ uma matriz simétrica definida positiva, $b \in \mathbb{R}^n$ e $c \in \mathbb{R}$. A função f tem um único minimizador x^*, que é global e satisfaz

$$\nabla f(x^*) = Ax^* + b = 0.$$

Consideremos agora a seguinte mudança de variável $y = x - x^*$, que nos fornece $x = y + x^*$. Subtituindo temos

$$\begin{aligned} f(y + x^*) &= \tfrac{1}{2}(y + x^*)^TA(y + x^*) + b^T(y + x^*) + c \\ &= \tfrac{1}{2}y^TAy + (x^*)^TAy + b^Ty + \tfrac{1}{2}(x^*)^TAx^* + b^Tx^* + c, \end{aligned}$$

onde $\frac{1}{2}(x^*)^TAx^* + b^Tx^* + c$ é uma constante que será denotada por $\bar{c}$. Assim usando isto e o fato que x^* é o minimizador de f,

$$\begin{aligned} f(y + x^*) &= \tfrac{1}{2}y^TAy + (x^*)^TAy + b^Ty + \bar{c} \\ &= \tfrac{1}{2}y^TAy + (Ax^* + b)^Ty + \bar{c} \\ &= \tfrac{1}{2}y^TAy + \bar{c}. \end{aligned}$$

Dessa forma, sem perda de generalidade e a menos de uma translação podemos considerar a função $f : \mathbb{R}^n \to \mathbb{R}$ definida por

$$f(x) = \frac{1}{2}x^TAx.$$

Olhando para a vizinhança do minimizador, podemos dizer que o gráfico dessas funções apresentam um comportamento característico que é mostrado na Fig. 3.2.

Lema 3.4 *Dado $x^0 \in \mathbb{R}^n$, considere a variedade linear $V_k = x^0 + \kappa_k$, a função quadrática $f(x) = \frac{1}{2}x^TAx$. Então para todo $x \in V_k$, existe um polinômio $q_k : \mathbb{R} \to \mathbb{R}$ de grau menor ou igual a k, com $a_0 = 1$, tal que:*

$$x = q_k(A)x^0 \quad e \quad f(x) = \frac{1}{2}(x^0)^TA(q_k(A))^2x^0.$$

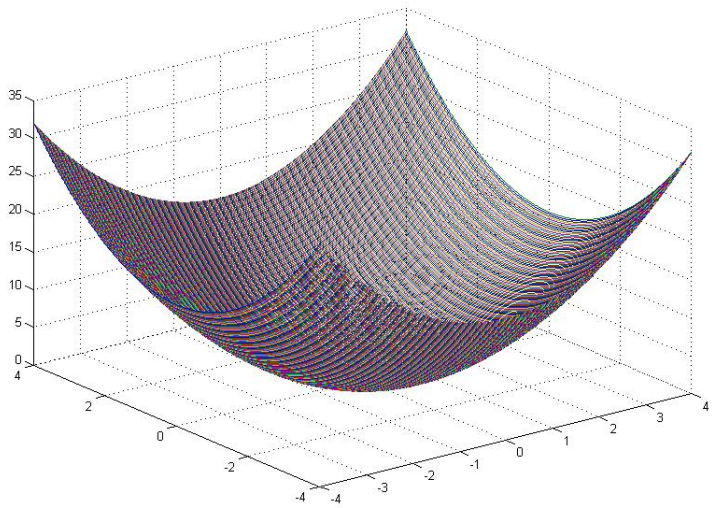

Figura 3.2: Gráfico de uma função quadrática definida em $\mathbb{R}^2$.

Demonstração. Seja x um ponto pertencente a variedade linear $V_k = x^0 + \left[Ax^0, A^2x^0, ..., A^kx^0\right]$. Ele pode ser escrito da forma

$$x = a_0x^0 + a_1Ax^0 + a_2A^2x^0 + ... + a_kA^kx^0$$

com $a_0 = 1$. Colocando x^0 em evidência, temos

$$x = (a_0I + a_1A + a_2A^2 + ... + a_kA^k)x^0,$$

que pode ser reescrito da forma

$$x = q_k(A)x^0,$$

com $q_k : \mathbb{R}^k \to \mathbb{R}$ dado por $q(t) = \sum_{i=0}^{k} a_it^i$, com $a_0 = 1$, provando a primeira igualdade. Além disso, pela definição da função

$$f(x) = \frac{1}{2}(x^0)^T(q_k(A))^TA(q_k(A))x^0.$$

Usando o fato que A é simétrica,

$$\begin{aligned}(q_k(A))^TA &= (a_0I + a_1A + a_2A^2 + ... + a_kA^k)^TA \\ &= A(a_0I + a_1A + a_2A^2 + ... + a_kA^k) \\ &= A(q_k(A))\end{aligned}$$

e consequentemente $f(x) = \frac{1}{2}(x^0)^TA(q_k(A))^2x^0$ completando a demonstração. □

Corolário 3.5 *Se* $x^k = \text{argmin}_{x \in V_k} \{f(x)\}$, *então*

$$f(x^k) \leq \frac{1}{2}(x^0)^T A(q_k(A))^2 x^0$$

para todo polinômio q_k *no espaço* $\mathcal{P}$ *dos polinômios de grau menor ou igual a* k *e* $a_0 = 1$.

Demonstração. Pela definição de x^k,

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

para todo $x \in V_k$. Usando o lema anterior segue o resultado. □

3.3 Complexidade algorítmica

Nesta seção mostramos que os métodos de gradientes conjugados para minimização de funções quadráticas são ótimos, no sentido de Nesterov [12].

O próximo teorema, demonstrado em [10][pág. 270] relaciona o espaço gerado pelos gradientes ($\nabla f(x^k)$) e o espaço gerado pelas direções (d^k) obtidas pelo método de gradientes conjugados com os espaços de Krylov.

Teorema 3.6 *Considere as sequências* (x^k) *geradas, a partir de* $x^0 \in \mathbb{R}^n$ *e das direções* (d^k), *pelo Algoritmo 2.6 de gradientes conjugados para minimizar* $f(x) = \frac{1}{2}x^T Ax$. *Se o método não termina em* x^k, *então:*

(a) $\kappa_k = \left[\nabla f(x^0), \nabla f(x^1), ..., \nabla f(x^{k-1})\right]$

(b) $\kappa_k = \left[d^0, d^1, ..., d^{k-1}\right]$.

Demonstração. Demonstramos simultaneamente (a) e (b) por indução. como $\kappa_1 = [\nabla f(x^0)]$ e $d^0 = -\nabla f(x^0) = -Ax^0$, as afirmações são verdadeiras para $k = 1$. Agora suponha que valem para k. Vamos provar que valem para $k + 1$. Pela definição da função e pelo algoritmo temos (2.5) , ou seja:

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

Pela hipótese de indução (a), temos que $\nabla f(x^{k-1}) \in \kappa_k \subset \kappa_{k+1}$ e por (b),

$$d^{k-1} = \sum_{i=1}^{k} a_i A^i x^0.$$

Logo

$$Ad^{k-1} = \sum_{i=1}^{k} a_i A^{i+1} x^0 \in \kappa_{k+1}.$$

Consequentemente $\nabla f(x^k) \in \kappa_{k+1}$. Além disso, $\nabla f(x^k) \notin \kappa_k = \left[d^0, d^1, ..., d^{k-1}\right]$, pois caso contrário $\nabla f(x^k) = 0$ tendo em vista o Lema 2.4 que diz que $\nabla f(x^k)$ é ortogonal a todas as direções d^i para todo $i < k$. Assim (a) está provado. Por 2.13 e 2.14, temos que

$$d^k = -\nabla f(x^k) + \beta_{k-1} d^{k-1} \in \kappa_{k+1}$$

pelo item (a) e pela hipótese de indução, completando assim a demonstração. □

O próximo teorema, provado em [14] garante que a complexidade algorítmica dos métodos de gradientes conjugados em relação ao valor da função é da ordem $O\left(\frac{1}{k^2}\right)$.

Cabe ressaltar que o método clássico de máxima descida tem complexidade $O\left(\frac{1}{k}\right)$.

Teorema 3.7 *Considere a sequência (x^k) gerada pelo método dos gradientes conjugados a partir de $x^0 \in \mathbb{R}^n$, para a função $f(x) = \frac{1}{2} x^T A x$. Então*

$$f(x^k) \leq \frac{L \left\|x^0\right\|^2}{2(2k+1)^2},$$

onde L é o maior autovalor de A.

Demonstração. Considere (x^k) a sequência gerada pelo algoritmo de gradientes conjugados, isto significa que x^k é minimizador de f na variedade

$$x^0 + \left[d^0, d^1, ..., d^{k-1}\right].$$

Pelo Teorema (3.6) $x^k = \text{argmin}_{x \in V_k} \{f(x)\}$ com $V_k = x^0 + \kappa_k$. Pelo Corolário 3.5

$$f(x^k) \leq \frac{1}{2} (x^0)^T A (q_k(A))^2 x^0,$$

para todo polinômio q_k no espaço $\mathcal{P}$ dos polinômios de grau menor ou igual a k e $a_0 = 1$. Usando o Lema 1.10 temos que

$$\begin{aligned} f(x^k) &\leq \frac{1}{2} \left\|x^0\right\|^2 \left\|A (q_k(A))^2\right\| \\ &\leq \frac{1}{2} \left\|x^0\right\|^2 \max_{0 \leq z \leq L} \left\{z \left(q_k(z)\right)^2\right\} \end{aligned} \tag{3.11}$$

onde L é o maior autovalor de A. Em particular, (3.11) vale para o polinômio

$$q_r(z) = \frac{T_{2k+1}(\frac{\sqrt{z}}{\sqrt{L}})}{(-1)^k(2k+1)(\frac{\sqrt{z}}{\sqrt{L}})}, \tag{3.12}$$

onde $T_{2k+1}(x) = \cos\left[(2k+1)\arccos(x)\right]$ é o polinômio de Chebyshev. Mas por (1.7)

$$T_{2k+1}(x) = 2^{2k}x^{2k+1} + ... + (-1)^k(2k+1)x$$

com $a_0 = a_2 = a_4 = ... = 0$. Logo

$$q_r(z) = \frac{2^{2k}}{(-1)^k(2k+1)}\left(\frac{z}{L}\right)^k + ... + 1$$

que é um polinômio de grau k com $q_r(0) = 1$. Assim $q_k \in \mathcal{P}$. Substituindo a expressão (3.12) em (3.11), temos que

$$f(x^k) \leq \tfrac{L}{2(2k+1)^2}\left\|x^0\right\|^2 \max_{0\leq z\leq L}\left(T_{2k+1}\left(\tfrac{z}{L}\right)\right)^2. \tag{3.13}$$

Usando o Teorema 3.6, $\max_{0\leq x\leq 1}|T_k(x)| = 1$, temos o resultado. □

A estimativa 3.12 é ótima para minimização de funções quadráticas através de métodos que fazem uso apenas de informação de até primeira ordem da função. Isto é, para qualquer método de primeira ordem é possível exibir uma função quadrática tal que o limite 3.12 é atingido, o que significa que ele é um limitante superior de complexidade.

Capítulo 4

Testes Computacionais

Os testes computacionais são ferramentas que nos ajudam a investigar a eficiência dos algoritmos na resolução dos problemas de acordo com o critério escolhido. Além disso, os critérios de parada também são importantes pois podem constatar pontos estacionários ou que o algoritmo fracassou, ou seja, atingiu seu número máximo de iterações e não encontrou solução para o problema.

Os algoritmos de gradientes conjugados discutidos no Capítulo 2 foram programados em Matlab 7.10 (R2010a).

Os testes computacionais foram divididos em duas baterias de problemas testes. Inicialmente para calibrar os algoritmos trabalhamos com funções quadráticas com Hessiana aleatória. Posteriormente comparamos as variantes dos métodos de gradientes conjugados de Fletcher-Reeves e de Polak-Ribière com buscas unidimensionais exata e inexata para funções não quadráticas incluindo os problemas sugeridos em [11].

4.1 Funções quadráticas

Programação da matriz definida positiva

Para testar os métodos, foi necessário criar um programa que gerasse funções quadráticas distintas. Para isso, usamos a fatoração QR e a diagonalização de matrizes simétricas.

Com o objetivo de gerar uma matriz A definida positiva com autovalores mínimo (m) e máximo (M), geramos inicialmente um vetor D com elementos aleatórios entre M e m. Para tanto, usando o comando "rand", geramos um vetor d com n valores aleatórios entre 0 e 1 e consideramos para todo $i = 1, ..., n$

$$D_i = m + \frac{(d_i - d_{min})}{(d_{max} - d_{min})}(M - m)$$

onde d_{min} e d_{max} são o menor e o maior valor entre as componentes de d. Notemos que

$$0 \leq \frac{(d_i - d_{min})}{(d_{max} - d_{min})} \leq 1.$$

Assim, quando esta expressão atinge seu maior valor, temos $D_i = M$. E quando atinge seu menor valor temos $D_i = m$. Para os testes, consideramos $m = 1$ e diversos testes com M entre 2 e 10000. Para construirmos a matriz diagonal de ordem n com os autovalores de A, usamos o comando "$diag(D)$", que gera uma matriz cuja diagonal principal é preenchida pelos elementos do vetor D.

Para gerar uma matriz ortogonal, criamos uma matriz $V \in \mathbb{R}^{n\times n}$ com valores aleatórios e utilizamos o comando do Matlab "$[Q, RR] = qr(V)$" que gera uma matriz ortogonal Q e uma matriz triangular superior RR.

Finalmente temos que

$$A = Qdiag(D)Q^T.$$

Testes

Consideramos funções quadráticas $f : \mathbb{R}^n \to R$ da forma

$$f(x) = \frac{1}{2}x^T Ax.$$

Comparamos o Algoritmo 2.6 com as três variantes do parâmetro β_k, ou seja, o dado por (2.18), proposto por Fletcher-Reeves e o (2.17) devido a Polak-Ribière. Note que para funções quadráticas, estas expressões coincidem entre si.

Para os nossos algoritmos, o ponto inicial x^0 é qualquer ponto pertencente ao $\mathbb{R}^n$, gerado aleatoriamente pelo comando "rand". O comprimento de passo foi calculado explicitamente por (2.7).

O critério de parada que usamos foi a norma do gradiente menor que 10^{-6}.

A Fig. 4.1 mostra a variação da norma do gradiente para minimização de uma função quadrática em R^n com $n = 1000$. Contudo, através da análise do gráfico podemos perceber que a função foi minimizada em menos de 350 iterações.

Com isso podemos concluir que para funções quadráticas os algoritmos têm o mesmo comportamento, ou seja, coincidem.

4.2 Funções não quadráticas

Comparamos o Algoritmo dos Gradientes Conjugados (2.6) nas duas variantes propostas por Fletcher e Reeves e Polak e Ribière, com as subvariantes diferindo entre si

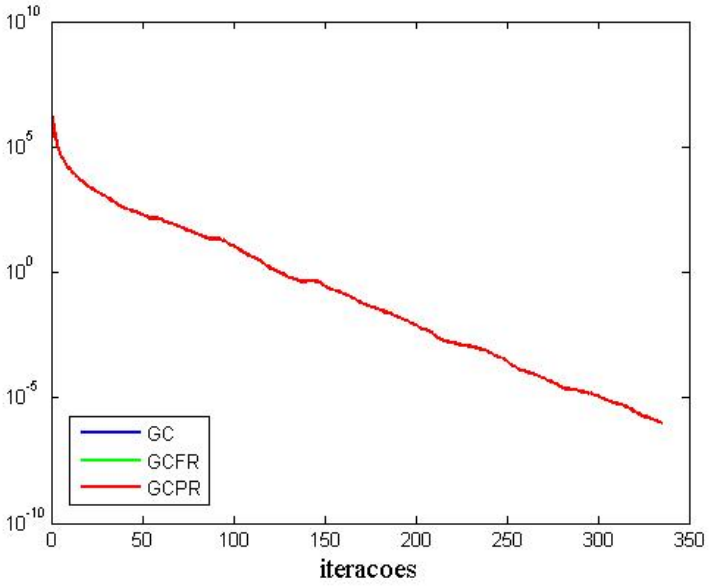

Figura 4.1: Variação da norma do gradiente.

pela busca realizada por Armijo ou Seção Áurea para o cálculo de α_k. Assim, para fins de comparação temos quatro algoritmos:

- $[GCFR - AU]$ - o Algoritmo de Gradientes Conjugados de Fletcher-Reeves com Seção Áurea;
- $[GCPR - AU]$ - o Algoritmo de Gradientes Conjugados de Polak-Ribière com Seção Áurea;
- $[GCFR - AR]$ - o Algoritmo de Gradientes Conjugados de Fletcher-Reeves com busca de Armijo;
- $[GCPR - AR]$ - o Algoritmo de Gradientes Conjugados de Polak-Ribière com busca de Armijo.

O critério de parada foi de 10^{-5} para a norma do gradiente. Caso o critério de parada não tenha sido atingido, em 10000 iterações, consideramos que o método falhou.

Como discutimos na Seção (2.4), nada garante que os métodos de gradientes conjugados minimizam uma função não quadrática em n passos, logo, para manter o sentido de conjugação entre as direções geradas, reiniciaremos o algoritmo a cada n iterações, fazendo $\beta_k = 0$.

Funções Testadas

Consideramos a bateria de problemas proposta em [11], que consiste de 34 funções. As primeiras 19 delas têm dimensão fixa, variando de 2 a 11, a vigésima função permite variação da dimensão da função de 1 até 31 e as outras funções, de 20 a 34 podem ter qualquer dimensão.

As funções são dadas como a soma do quadrado de m funções, ou seja, dadas $f_i : \mathbb{R}^n \to \mathbb{R}$ para $i = 1, ..., m$ temos

$$f(x) = \sum_{i=1}^{m} f_i^2(x).$$

O exemplo a seguir ilustra uma das funções desse banco de problemas.

Exemplo 4.1 *A função de Rosenbrock é definida em* $\mathbb{R}^2$ *como a soma de* $m = 2$ *funções:*

$$f_1(x) = 10(x_2 - x_1^2) \qquad e \qquad f_2(x) = 1 - x_1.$$

Logo, a função f *a ser minimizada é*

$$f(x) = f_1^2(x) + f_2^2(x) = 100x_1^4 + 100x_2^2 - 200x_2x_1^2 + x_1^2 - 2x_1 + 1.$$

O ponto inicial dado é $x^0 = (-1.2, 1)$. *O minimizador da função é o ponto* $x^* = (1, 1)$.

Além deste banco de funções, implementamos as seguintes funções:

$$\begin{aligned}
f_1(x) &= 2x_1^3 - 3x_1^2 - 6x_1x_2(x_1 - x_2 - 1); \\
f_2(x) &= x_1^3 + x_2^3 - 3x_1x_2; \\
f_3(x) &= sen(x_1)\ sen(x_2) + e^x; \\
f_4(x) &= x_1^2 - x_1x_2 + 2x_2^2 - 2x_1 + \tfrac{2}{3}x_2 + e^{x_1+x_2}; \\
f_5(x) &= 2(x_2 - x_1^2)^2 + (1 - x_1)^2; \\
f_6(x) &= 100(x_2 - x_1^2)^2 + (1 - x_1)^2; \\
f_7(x) &= 2(x_1^2 + \textstyle\sum (x_i + x_{i+1})^2) + \tfrac{1}{2}\|x\|^2; \\
f_8(x) &= \tfrac{1}{2}\|x\|^2 + \textstyle\sum_{i=1}^{n} i\ (e^{x_i} - x_i - 1).
\end{aligned}$$

onde as sete primeiras estão definidas em $\mathbb{R}^2$ e a última em $\mathbb{R}^n$. Em nosso teste, consideramos n = 2,10,25,50,100,500,1000,2000,5000.

Assim consideramos no total um banco de 42 funções, algumas das quais com liberdade na escolha da dimensão do problema. Assim, no total resolvemos 170 problemas com dimensão entre 2 e 300.

Análise dos Resultados

Na Fig.4.2 comparamos os quatro algoritmos e temos a variação da norma do gradiente da Função de Rosenbrock Estendida em função do número de iterações em $\mathbb{R}^{1000}$.

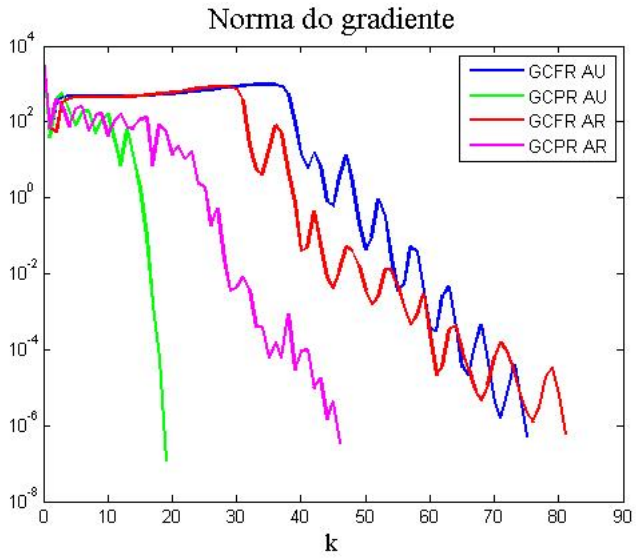

Figura 4.2: Variação da norma do gradiente.

O método de gradientes conjugados GCPR-AU com β_k dado por (2.17) devido a Polak-Ribière e busca exata para o cálculo do comprimento do passo α_k foi o mais eficiente para minimizar esta função.

Para poder tirar alguma conclusão mais geral do desempenho dos quatro métodos consideramos o gráfico de perfil de desempenho como proposto em [4].

Este gráfico consiste de uma ferramenta utilizada em Otimização para comparar o desempenho de t métodos de um conjunto T, quando aplicados para resolver s problemas de um conjunto S, usando critérios como tempo computacional, número de avaliações de função ou número de iterações.

Como os quatro métodos são análogos, utilizamos como critério de análise o número de iterações gasto para atingir o critério de parada.

A Fig. 4.3 mostra o gráfico de perfil de desempenho dos quatro métodos. O eixo vertical representa a porcentagem de problemas resolvidos e o horizontal indica o fator multiplicativo do número de iterações de um determinado algoritmo em relação ao melhor algoritmo.

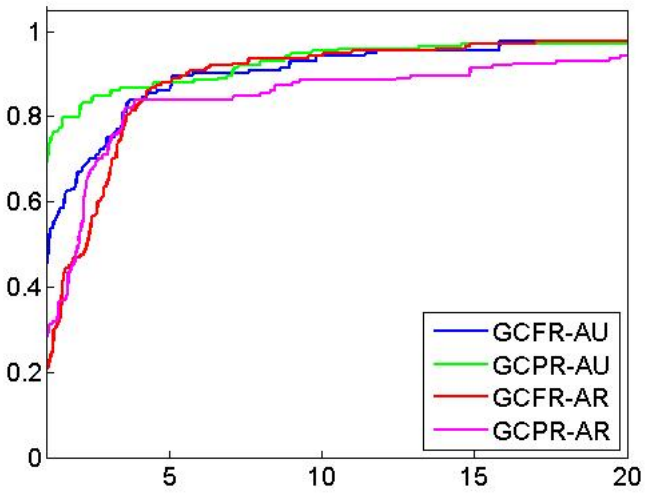

Figura 4.3: Gráfico de perfil de desempenho.

Da Fig. 4.3 temos algumas interpretações:

- O método GCPR-AU ganha em 70% dos problemas, o método GCFR-AU ganha em 46% dos problemas, o método GCPR-AR ganha em 29% dos problemas e o método GCFR-AR ganha em 21% dos problemas.

- A soma da porcentagem de problemas resolvidos por todos os algoritmos é superior a 100% pois quando dois métodos resolvem o problema, independente do número de iterações que cada um leva, é contabilizado para os dois.

- GCPR-AU resolve 80% dos problemas gastando não mais que o dobro do número de iterações do melhor algoritmo; enquanto os demais algoritmos podem gastar até 4 vezes mais o menor número de iterações.

- Alguns problemas não foram resolvidos por nenhum dos métodos. Fazendo uma análise mais detalhada identificamos que os problemas 26, 33 e 34 não foram resolvidos.

- A ordem de desempenho dos métodos, colocada do melhor desempenho ao pior, é: GCPR-AU, GCFR-AU, GCPR-AR e GCFR-AR.

Conclusão

Neste trabalho, com o objetivo de discutirmos o método dos gradientes conjugados para minimização irrestrita, apresentamos uma breve revisão e exposição de tópicos de Álgebra Linear e Cálculo. Falamos das direções conjugadas e de suas principais propriedades. Provamos que o método de gradientes conjugados minimiza uma função quadrática definida em $\mathbb{R}^n$ em até n passos com complexidade ótima.

Apresentamos o algoritmo de gradientes conjugados com duas variantes propostas por Polak e Ribière e por Fletcher e Reeves, expondo a coincidência dos algoritmos para funções quadráticas.

Essas variantes permitem a extensão do método de gradientes conjugados para funções não quadráticas. Para tanto incluimos uma busca unidirecional exata, tipo Seção Áurea, ou inexata, tipo Armijo. Constatamos através de testes numéricos que, conforme o pressuposto inicial, o algoritmo de Polak e Ribière com busca exata tem o melhor desempenho.

Referências Bibliográficas

[1] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, Massachusetts, 1995.

[2] R. L. Burden and J. D. Faires. *Numerical Analysis*. Ohio, United States, 2008.

[3] A. R. Conn, N. I. M. Gould, and Ph. L. Toint. *Trust-Region Methods*. MPS-SIAM Series on Optimization, SIAM, Philadelphia, 2000.

[4] E. D. Dolan and J. J. Moré. Benchmarking optimization software with performance profiles. *Mathematical Programming*, 91:201–213, 2002.

[5] L. M. Elias. Minimização de funções quadráticas. 2010.

[6] A. Friedlander. *Elementos de Programação Não-Linear*. Unicamp.

[7] A. Izmailov and M. Solodov. *Otimização: Condições de Otimalidade, Elementos de Análise Convexa e Dualidade*, volume 1. IMPA, Rio de Janeiro, 2005.

[8] Steven J. Leon. *Álgebra Linear com Aplicações*. Rio de Janeiro, 1999.

[9] E. L. Lima. *Curso de Análise*, volume 1. IMPA, Rio de Janeiro, Brasil, 1981.

[10] D. G. Luenberger. *Linear and Nonlinear Programming*. Addison - Wesley Publishing Company, New York, 1986.

[11] J. J. Moré, B. S. Garbow, and K. E. Hillstrom. Testing unconstrained optimization software. 1981.

[12] Y. Nesterov. *Intruductory Lectures on Convex Optimization - A basic course*. Kluwer Academic Publishers, Boston, 2004.

[13] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer Series in Operations Research. Springer-Verlag, 1999.

[14] B. T. Polyak. *Introduction to Optimization, Optimization Software*. New York, 1987.

[15] A. Ribeiro and E. W. Karas. *Notas de aula para Disciplina de Otimização I, disponível em http://people.ufpr.br/ ademir.ribeiro/ensino/livro/livro.pdf*. Curitiba, 2010.