

VIVIANE CRISTHYNE BINI BARBOSA

**UM NOVO ALGORITMO PARA RESOLVER
PROBLEMAS DE MINIMIZAÇÃO DE FUNÇÕES
NÃO LINEARES SUJEITA A RESTRIÇÕES
LINEARES DE IGUALDADE**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção e Sistemas da Pontifícia Universidade Católica do Paraná, como requisito parcial para obtenção do título de Mestre em Engenharia de Produção e Sistemas.

Área de concentração: Gerência de Produção e Logística
Orientador: Prof. Dr. Raimundo José Borges de Sampaio

Curitiba-PR

[2006]

VIVIANE CRISTHYNE BINI BARBOSA

**UM NOVO ALGORITMO PARA RESOLVER
PROBLEMAS DE MINIMIZAÇÃO DE FUNÇÕES
NÃO LINEARES SUJEITA A RESTRIÇÕES
LINEARES DE IGUALDADE**

Curitiba-PR

[2006]

Barbosa, Viviane Cristhyne Bini

Um novo algoritmo para resolver problemas de minimização de funções não lineares sujeita a restrições lineares de igualdade.

Curitiba, 2006.

Dissertação (Mestrado) - Pontifícia Universidade Católica do Paraná.
Programa de Pós-Graduação em Engenharia de Produção e Sistemas.

1. palavras-chave: NLP, conjugacidade, lagrangeano, hessiana.

TERMO DE APROVAÇÃO

VIVIANE CRISTYNE BINI BARBOSA

Um Novo Algoritmo para Resolver Problemas de Minimização de Funções Não Lineares Sujeita a Restrições Lineares de Igualdade

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre no Curso de Mestrado em Engenharia de Produção e Sistemas, Programa de Pós-Graduação em Engenharia de Produção e Sistemas, do Centro de Ciências Exatas e de Tecnologia da Pontifícia Universidade Católica do Paraná, pela seguinte banca examinadora:

Presidente da Banca

Prof. Dr. Raimundo José Borges de Sampaio (PPGEPS/PUCPR - Orientador)

Prof. Dr. Marco Antonio Cândido Barbosa (PUCPR - Membro Titular)

Prof. Dr. Luiz Carlos Matioli (UFPR - Membro Titular)

Curitiba, 29 de março de 2006

A minha irmã Thaty (*in
memorian*)

Agradecimentos

A minha Mãe, **Itália Marina Bini** por tudo.

Ao Prof. Raimundo J. B. de Sampaio, pela compreensão e dedicação com que orientou este trabalho, tornando possível a realização do mesmo.

A todos os professores, que contribuíram com seus conhecimentos, os quais foram essenciais para minha formação profissional.

A minha vovó, Cecília Lemos Bini e a minha madrinha, Neusa Cristina Bini Haiduk por todo o apoio.

Ao Rubens Deniz Conte, pelo companherismo, apoio e encorajamento.

Às colegas Ana Paula Delowski e Liara J. Minikovski.

Sumário

Lista de Tabelas	xiii
Lista de Siglas	xiv
Resumo	xvi
Abstract	xviii
1 Introdução	1
1.1 Desafio	1
1.2 Motivação	2
1.3 Proposta e Contribuição	2
1.4 Organização	3
2 Tópicos de Álgebra Linear	4
2.1 Espaços Vetoriais	4
2.2 Subespaços Vetoriais	5
2.2.1 Transformações Lineares	5
2.3 Bases	5
2.3.1 Conjunto Geradores	5
2.3.2 Independência Linear	6
2.3.3 Os Subespaços Fundamentais de uma Matriz	8

2.3.4	Dimensão	8
2.4	Normas de vetores	9
2.5	Produto Interno	9
2.5.1	Desigualdade de Schwartz	10
2.5.2	Norma Induzida Pelo Produto Interno	10
2.5.3	Matrizes Simétricas e Matrizes Hermetianas	10
2.5.4	Normas de Matrizes	11
2.5.5	Norma de Frobenius	12
2.6	Matriz de Projeção	12
2.7	Ortogonalidade	13
2.7.1	Projeções Ortogonais	14
2.7.2	Base Ortogonal	14
2.8	Fatoração de Matrizes	16
2.8.1	Fatoração LU	17
2.8.2	Fatoração de Cholesky	18
2.8.3	Fatoração QR	23
2.9	Autovalores e Autovetores	24
2.10	Teoria da Perturbação e Número de Condição	24
3	Tópicos em Análise do \mathbb{R}^n e Otimização	27
3.1	Funções de Classe C^k	28
3.1.1	Gradiente	28
3.1.2	Hessiana	28
3.1.3	Teorema de Taylor	29
3.1.4	Teorema Fundamental do Cálculo ou Teorema de Newton	29
3.1.5	Derivada Direcional	30
3.1.6	Direção de Descida	31
3.1.7	Aproximação de derivadas por diferenças finitas	32
3.2	Condições de Otimalidade	33

3.2.1	Teorema de Weierstrass	34
3.2.2	Conjuntos Convexos	34
3.2.3	Funções Convexas	34
3.2.4	Caracterização de um ponto de mínimo	36
3.2.5	Condições de Otimalidade para uma Função com Restrições Lineares de Igualdade	37
3.2.6	Convergência de Seqüências e Rapidez de Convergência	39
3.3	Funções Quadráticas	40
3.3.1	Propriedades Básicas de Funções Quadráticas	42
3.3.2	Minimização de Quadráticas Sobre Hiperplanos	48
3.3.3	Método de Direções Conjugadas	52
3.3.4	Método do Gradiente Conjugado	54
4	Uma Nova Abordagem para a Minimização de uma Função Não Linear Sujeita a Restrições Lineares de Igualdade	59
4.1	O Modelo Quadrático	61
4.1.1	Uma Abordagem do Espaço Nulo	63
4.1.2	Abordagem de Conjugacidade	65
4.2	Descrição do algoritmo	68
4.2.1	Obtenção de uma Aproximação para o Gradiente e a Hessiana	69
4.2.2	Inicialização	69
4.2.3	Iterações	70
4.2.4	Finalização	71
4.3	Algoritmo I - Versão QR de A^T	71
4.4	Algoritmo II - Versão $(B \ N)$	72
4.5	Algoritmo III - Versão B^{-1}	72
4.5.1	Obtenção da H -conjugacidade das direções	73
5	Experimentos Numéricos	79
6	Conclusões	82

Bibliografia	83
A Tabelas de Dados	86
B Tabelas de Resultados	89

Lista de Tabelas

A.1	Dados do problema com a função Wood	86
A.2	Dados do problema com a função Rosenbrock Singular Extendida	87
A.3	Dados do problema com a função Powell Singular Extendida	87
A.4	Dados Função Quadrática - $f(x) = \frac{1}{2}x^T Hx - h^T x$	88
B.1	Resultados Função Quadrática - Direções Conjugadas QR de A^T	90
B.2	Resultados Função Quadrática - Direções Conjugadas $(B N)$	90
B.3	Resultados Função Quadrática - Direções Conjugadas B^{-1}	91
B.4	Resultados Função Quadrática - <i>CG Reduzido</i>	91
B.5	Resultados Wood - Direções Conjugadas QR de A^T	92
B.6	Resultados Wood - CG Reduzido QR de A^T	92
B.7	Resultados Wood - Direções Conjugadas B_1^{-1}	93
B.8	Resultados Wood - CG Reduzido utilizando B_1^{-1}	93
B.9	Resultados Wood - Direções Conjugadas $(B_1 N)$	94
B.10	Resultados Wood - CG Reduzido $(B_1 N)$	94
B.11	Resultados Wood - Direções Conjugadas $(B_2 N)$	95
B.12	Resultados Wood - CG Reduzido $(B_2 N)$	96
B.13	Resultados Wood - Direções Conjugadas B_2^{-1}	97
B.14	Resultados Wood - CG Reduzido B_2^{-1}	98
B.15	Resultados Rosenbrock Singular Extendida - Direções Conjugadas QR A^T	99

B.16	Resultados Rosenbrock Singular Extendida - CG Reduzido QR de A^T . . .	100
B.17	Resultados Powell Singular Extendida - Direções Conjugadas QR de A^T . .	101
B.18	Resultados Powell Singular Extendida - CG Reduzido QR de A^T	102
B.19	Resultados Powell Singular Extendida - Direções Conjugadas B_1^{-1}	103
B.20	Resultados Powell Singular Extendida - CG Reduzido B_1^{-1}	104
B.21	Resultados Powell Singular Extendida - Direções Conjugadas $(B_1 N)$. . .	105
B.22	Resultados Powell Singular Extendida - CG Reduzido $(B_1 N)$	106
B.23	Resultados Powell Singular Extendida - Direções Conjugadas B_2^{-1}	107
B.24	Resultados Powell Singular Extendida - Direções Conjugadas $(B_2 N)$. . .	108

Lista de Siglas

GS	Gram-Schmidt Clássico
$M - GS$	Gram-Schmidt Modificado
CG	Gradiente Conjugado
\mathbb{R}	conjunto dos números reais
NLP	Programação Não Linear
QP	Programação Quadrática
$\nabla f(x)$	gradiente da função f no ponto x
$\nabla^2 f(x)$	hessiana da função f no ponto x
$Z^T \nabla f(x_k) Z$	hessiana reduzida no ponto x_k
g_k	aproximação para $\nabla f(x_k)$
H_k	aproximação para $\nabla^2 f(x_k)$
A	Matriz das restrições
b	Vetor b em \mathbb{R}^m
$\aleph(A)$	Espaço nulo de A
\mathbb{R}^n	Espaço dimensional n
$cond(A)$	número de condição de A
$f(x)$	Função objetivo
$\aleph(A)$	Espaço coluna de A
$\aleph(A^T)$	Espaço linha de A
V	Espaço vetorial real
LU	Fatoração LU
Q	Matriz ortogonal
QR	Decomposição QR
Z	Matriz cujas colunas geram $\aleph(A)$

B	Matriz básica
B^{-1}	Matriz inversa B
d	Vetor direção em \mathbb{R}^n
dz	Vetor direção em \mathbb{R}^{n-m}
N	Matriz não básica
x^*	solução ótima
x_B	Variáveis básicas
x_N	Variáveis não básicas
α_k	Tamanho do passo na iteração k
$Z^T H Z$	Matriz hessiana reduzida
$Z^T g$	Vetor gradiente reduzido
$\text{Im}(A)$	Conjunto imagem de A
I_{n-m}	Matriz Identidade de ordem $n - m$
\mathbb{C}	Conjunto dos números complexos

Resumo

Este trabalho trata do problema de minimizar uma função não linear sujeita a restrições lineares de igualdade,

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & Ax = b \end{aligned}$$

onde $f(x)$ é uma função duas vezes continuamente diferenciável, A é uma matriz $m \times n$, com $m < n$, de posto m , e b um vetor de \mathbb{R}^m . Geralmente um problema desse tipo é resolvido através de uma seqüência de problemas quadráticos, onde a função objetivo é uma aproximação quadrática do lagrangeano do problema, e nesse caso mais especificamente, uma aproximação quadrática da função a ser minimizada, e as restrições são as restrições originais do problema [20]. O algoritmo proposto gera uma seqüência de direções de busca no espaço nulo das restrições, e portanto, a seqüência gerada pelo algoritmo é sempre viável. Geralmente os algoritmos utilizados para resolver esse tipo de problema exigem que a cada iteração seja resolvido um novo sistema reduzido [13],

$$Z^T \nabla^2 f(x_k) Z p = -Z^T \nabla f(x_k)$$

onde as colunas de Z formam uma base para o espaço nulo de A , $\nabla^2 f(x)$ é a hessiana do lagrangeano, $\nabla f(x)$ é o gradiente do lagrangeano do problema e p obtido desse modo é uma direção de busca para o problema. Para a resolução desse sistema reduzido é necessário uma escolha adequada de Z [1], seja para evitar a introdução de mal condicionamento no sistema, seja para assegurar uma rápida convergência da seqüência produzida. Na abordagem por nós utilizada as direções de busca são obtidas diretamente no espaço nulo da matriz de restrições do problema, não havendo a necessidade de resolver o sistema reduzido a cada iteração do problema. O algoritmo utilizado tem dois passos iterativos

principais. No primeiro passo é utilizado um algoritmo do tipo Gradiente Conjugado Não Linear. Nesse passo o algoritmo encontra um novo vetor no núcleo da matriz das restrições que é conjugado com as direções anteriores e faz uma minimização unidimensional exata ao longo dessa direção, e assim sucessivamente. No segundo passo, de posse da solução encontrada no primeiro passo, resolve-se um novo subproblema quadrático e retorna ao primeiro passo.

Palavras-chave: NLP, conjugacidade, lagrangeano, hessiana.

Abstract

The aim of this work is to present a new approach to solve a nonlinear programming (NLP) problem with linear equality constraints, that is,

$$\begin{array}{ll} \textit{Minimize} & f(x) \\ \textit{subject to} & Ax = b \end{array}$$

where $f(x)$ is a twice continuously differentiable function, A is a $m - by - n$ matrix with $m < n$, of rank m , and $b \in \mathbb{R}^m$. In general, this kind of problem is solved using a sequence of quadratic subproblems, where the objective function is the quadratic Lagrangian approximation of the problem [20], and the constraints are the same from the original problem. In general, the algorithms used to solve this kind of problem require at each iteration the solution of a new reduced system, [13]. The choose of Z , a matrix whose columns span the null space of constraints is subtle, since it may bring ill-conditioning to the system to be solved, and thus affecting the behavior of the sequence generate by the algorithm [1]. The algorithm proposed here produces a sequence of search directions belonging to the null space of the constraints, and hence, the sequence generated is always feasible. In this approach, the search directions are obtained directly in the null space constraints matrix, so we do not have the need to solve the reduced system at each iteration. The algorithm used was divided in two steps. In the first step it uses a nonlinear conjugate gradient method. In this step the algorithm finds a new vector belonging to the null space of the constraints which is conjugate to previous directions and then it makes a one-dimensional exact minimization along this direction, and so on. In this step the algorithm finds the solution in at maximum k iterations, where k is the dimension of the null space of the constraints. In the second step a new quadratic approximation to the problem is defined and then it return to the first step.

Keywords: NLP, conjugacy, Lagrangean, Hessian.

Introdução

1.1 Desafio

Este trabalho trata do problema de programação não linear com restrições lineares de igualdade, visto aqui, como o problema de,

$$\begin{aligned} & \text{Minimizar} && f(x) \\ & \text{s.a} && Ax = b \end{aligned} \tag{1.1}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função duas vezes continuamente diferenciável, A é uma matriz $m \times n$ com posto m , e b é um vetor de \mathbb{R}^m .

Geralmente um problema desse tipo é modelado por uma seqüência de problemas de programação quadrática (QP). Em muitos casos, os algoritmos utilizados para resolver um problema baseado nesta idéia de resolver subproblemas quadráticos devem ter a priori, a necessidade de gerar uma seqüência de pontos viáveis. Esses são os casos em que a solução do problema está restrita a um conjunto de pontos que satisfazem alguns restrições. É o caso de resolver o problema de otimizar uma função sujeita a um conjunto de restrições impostas às variáveis.

O problema (1.1) exige que a seqüência de pontos gerada seja viável, pois a solução do problema deve ser um ponto que satisfaça um sistema de equações lineares de igualdade. Para tanto as direções de busca utilizadas pelo algoritmo devem estar no espaço nulo da matriz desse sistema, i.e., tais direções devem pertencer ao espaço nulo da matriz das restrições do problema. Geralmente, os algoritmos utilizados para resolver (1.1) exigem que a cada iteração seja resolvido um novo sistema reduzido, [11, 13, 18, 20, 21, 24],

$$Z^T \nabla^2 f(x_k) Z d = -Z^T \nabla f(x_k) \tag{1.2}$$

onde Z é a matriz cujas colunas formam uma base para o espaço nulo de A , $\nabla^2 f(x_k)$ é a hessiana da função, $\nabla f(x_k)$ é o gradiente da função e d obtido desse modo é uma direção de busca para o problema. As expressões $Z^T \nabla^2 f(x_k) Z$ e $Z^T \nabla f(x_k)$ do sistema (1.2) são, respectivamente, denominados matriz reduzida e gradiente reduzido.

Para a resolver esse sistema pode ser necessário diversas escolhas de Z [1] as quais, podem assegurar uma rápida convergência da seqüência. Um desafio para o problema de minimizar uma função não linear sujeita a restrições lineares de igualdade está na escolha de se utilizar um algoritmo para resolver o problema (1.1), o qual não envolva o cálculo da hessiana reduzida, ou seja, não resolva (1.2) a cada iteração k .

1.2 Motivação

Um algoritmo utilizado para resolver o problema (1.1), que não necessite resolver o sistema reduzido (1.2) no processo iterativo possui diferentes razões para ser estudado. Uma delas é que se a seqüência de aproximações obtida pelo algoritmo como soluções dos subproblemas, i.e., pelo problema de programação quadrática, for gerada de maneira tal que mantenha alguma propriedade de valor da matriz hessiana, torna-se de grande valor para a PNL.

Nesse contexto, convém lembrar também que, manter alguma particularidade, significa não inserir coisas que danifiquem tais particularidades. É importante estar ciente do fato que aproximações de boa qualidade para a solução de um subproblema podem gerar boas soluções para o problema original. A maneira como essa solução foi gerada implica na qualidade da mesma.

Uma segunda razão está baseada no fato de que: são muitas as aplicações do estudo da programação não linear, e.g., na engenharia de produção. Alguns exemplos são a aplicação desta no planejamento da produção, na alocação de recursos, nas análises de dados os quais envolvem problemas de quadrados mínimos. Na aplicação da PNL em engenharia de produção esse problema aparece em diversas áreas, ou diretamente [2, 10, 19], ou combinado com outras técnicas de otimização [26].

1.3 Proposta e Contribuição

Neste trabalho, é apresentado um novo algoritmo para resolver o problema (1.1), o qual não envolve o cálculo da hessiana reduzida, ou seja, o algoritmo não resolve (1.2) a cada iteração k do processo.

Tomando por base esses "requisitos" o trabalho tem por objetivo apresentar o desenvolvimento de um algoritmo baseado em idéias de conjugacidade. O algoritmo proposto trabalha diretamente no espaço nulo da matriz das restrições, gerando um conjunto H -conjugado de vetores, onde H_k é uma aproximação da hessiana da função a ser minimizada. O algoritmo tem dois passos iterativos principais. No primeiro passo é utilizado um algoritmo do tipo Gradiente Conjugado Não Linear [16]. Nesse passo o algoritmo encontra um novo vetor no núcleo da matriz das restrições que é conjugado com as direções anteriores e faz uma minimização unidimensional exata ao longo dessa direção, e assim sucessivamente. No segundo passo, de posse da solução encontrada no primeiro passo, cria-se um novo subproblema quadrático e retorna ao primeiro passo. A cada iteração é feita a checagem do critério de otimalidade.

De posse de uma base para o espaço nulo de A a qual designa o fator de mudança para análise e comparações das três versões do novo algoritmo, a atualização das direções no processo iterativo está baseada no processo é conhecido como Gram-Schmidt Modificado. Neste trabalho foi feita uma adaptação a este processo, para que a atualização da hessiana H_k na segunda fase do processo iterativo pudesse ser usada como o fator da atualização das direções na primeira fase iterativa.

O algoritmo foi testado para as funções conhecidas na literatura: Wood, Powell Singular Extendida, Rosenbrock Singular Extendida. Também foram realizados testes para a quadrática com a matriz de Hilbert como hessiana.

1.4 Organização

Este trabalho está organizado em 6 capítulos, sendo que o primeiro consiste na introdução. O segundo capítulo apresenta uma revisão dos conceitos básicos de álgebra linear. O terceiro capítulo traz tópicos em análise do \mathbb{R}^n e otimização. No quarto capítulo é descrito um novo algoritmo baseado na idéia de conjugacidade para minimização de uma função não linear sujeita a restrições lineares de igualdade. O quinto capítulo apresenta os resultados dos experimentos numéricos e os resultados obtidos com a aplicação do novo algoritmo. No sexto capítulo são apresentadas as conclusões obtidas com a análise dos resultados e as indicações para pesquisas futuras. As tabelas de dados e resultados encontram-se em anexo.

Tópicos de Álgebra Linear

2.1 Espaços Vetoriais

Definição 1 *Um espaço vetorial real V é um conjunto não vazio de vetores, munidos das operações de adição e multiplicação por escalar. Isso significa que para x, y vetores quaisquer pertencentes a V e α um número real, o vetor soma $x + y$ está em V (i.e., V é fechado para a operação de adição) e αx está em V (i.e., V é fechado para a operação de multiplicação por escalar). As operações de adição e multiplicação por escalar devem satisfazer as seguintes condições:*

- i. $x + y = y + x, \forall x, y \in V$;*
- ii. $x + (y + z) = (x + y) + z, \forall x, y, z \in V$;*
- iii. $\exists \mathbf{0} \in V$ tal que $x + \mathbf{0} = x$ e $\mathbf{0} + x = x, \forall x \in V$ ($\mathbf{0}$ é chamado de vetor nulo);*
- iv. $\forall x \in V, \exists -x \in V$ tal que $x + (-x) = \mathbf{0}$ e $(-x) + x = \mathbf{0}$;*
- v. $\alpha(x + y) = \alpha x + \alpha y, \forall x \in V$;*
- vi. $(\alpha + \beta)x = \alpha x + \beta x, \forall \alpha, \beta \in \mathbb{R}$ e $\forall x \in V$;*
- vii. $\alpha(\beta x) = (\alpha\beta)x, \forall \alpha, \beta \in \mathbb{R}$ e $\forall x \in V$;*
- viii. $1x = x, \forall x \in V$.*

2.2 Subespaços Vetoriais

O teorema a seguir, enuncia o que é um subespaço vetorial real.

Teorema 1 *Suponha V um espaço vetorial real e W um subconjunto não vazio de V . Então, W é um subespaço de V se, e somente se, as seguintes condições se cumprem:*

1. W é fechado para a operação de multiplicação por escalar, isto é, $\alpha w_1 \in W, \forall w_1 \in W$;
2. W é fechado para a operação de adição, isto é, $w_1 + w_2 \in W, \forall w_1, w_2 \in W$.

2.2.1 Transformações Lineares

Definição 2 *Sejam V e W dois espaços vetoriais. A transformação linear A de V (chamada de domínio de A) - em W (chamada de imagem de A), ou seja, $A : V \rightarrow W$ é uma correspondência que determina a cada vetor v em V um vetor $A(v)$ em W de tal maneira que:*

1. $A(v_1 + v_2) = A(v_1) + A(v_2)$ para todo vetor v_1, v_2 .
2. $A(\alpha v) = \alpha A(v)$ para todo vetor v e escalar α .

Equivalentemente, $A(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 A(v_1) + \alpha_2 A(v_2)$ para todo vetor v_1, v_2 e escalares α_1 e α_2 .

2.3 Bases

A definição de base está baseado em dois outros conceitos, são eles: conjuntos geradores e independência linear. Antes de definirmos o que é uma base faz-se necessária a compreensão de conjuntos geradores e independência linear.

2.3.1 Conjunto Geradores

Conjunto gerador é um conjunto de vetores que podem ser vetores linearmente independentes ou linearmente dependentes. Para possibilitar uma melhor compreensão sobre conjunto gerador faz-se necessário antes, citar o conceito de combinação linear.

Combinação Linear

Uma combinação linear dos vetores v_1, v_2, \dots, v_n é uma expressão da forma

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n,$$

onde os α_i 's são escalares.

Vetores Linearmente Dependentes e Vetores Linearmente Independentes

Um vetor v é dito linearmente dependente dos vetores v_1, v_2, \dots, v_n se e somente se v pode ser escrito como alguma combinação linear de v_1, v_2, \dots, v_n ; caso contrário, v é dito ser linearmente independente dos vetores v_1, v_2, \dots, v_n .

Com base nesses conceitos de combinação linear e vetores linearmente independentes, pode-se definir o que é um conjunto gerador.

Definição 3 *Seja S um conjunto de vetores v_1, v_2, \dots, v_n em V . S é dito gerador de algum subespaço V_0 de V se e somente se S é um subconjunto de V_0 e todo vetor v_0 em V_0 é linearmente independente dos vetores do conjunto S .*

2.3.2 Independência Linear

Definição 4 *Seja $L = \{v_1, v_2, \dots, v_k\}$ um conjunto não vazio de vetores.*

Suponha que

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k = 0$$

se $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$, então L é dito ser linearmente independente. Equivalentemente, L é linearmente dependente se e somente se existem escalares $\alpha_1, \alpha_2, \dots, \alpha_k$, não todos nulos, com

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k = 0.$$

O seguinte teorema formaliza que de fato as duas diferentes maneiras de definir o que é um conjunto linearmente independente são equivalentes. Dele também extraímos outros resultados úteis.

Teorema 2

1. Suponha que

$$L = \{v_1, v_2, \dots, v_k\}$$

com $k \geq 2$ e todo vetor $v_i \neq \mathbf{0}$, para $i = 1, 2, \dots, k$. Então, L é linearmente independente se e somente se ao menos um dos v_j é linearmente dependente dos vetores restantes v_i ($i \neq j$); em particular, L é linearmente dependente se e somente se ao menos um dos vetores v_j é linearmente dependente dos seus vetores precedentes v_1, v_2, \dots, v_{j-1} .

2. Qualquer conjunto contendo o vetor $\mathbf{0}$ é linearmente dependente.

3. $\{v\}$ é linearmente dependente se e somente se $v \neq \mathbf{0}$.

4. Suponha que v é linearmente dependente dos vetores do conjunto

$$L = \{v_1, v_2, \dots, v_k\},$$

e que v_j seja linearmente dependente sobre os outros vetores de L , isto é,

$$L'_j = \{v_1, v_2, \dots, v_{j-1}, v_{j+1}, \dots, v_k\}.$$

Então, v é linearmente dependente sobre L'_j .

5. Todo subconjunto de um conjunto linearmente independente é linearmente independente.

6. Suponha que L é um conjunto finito de vetores e que algum subconjunto L_0 de L é linearmente dependente. Então L é linearmente dependente.

A definição de base para um espaço vetorial V é dada a seguir.

Definição 5 Uma base para um espaço vetorial V é um conjunto gerador linearmente independente.

O seguinte teorema, refere-se a unicidade da representação do vetor quando esse pertencer a uma base.

Teorema 3 Seja $B = \{v_1, \dots, v_r\}$ uma base de um espaço vetorial V . Então, a representação de todo $v \in B$ é única: se $v = \alpha_1 v_1 + \dots + \alpha_r v_r$ e também $v = \alpha'_1 v_1 + \dots + \alpha'_r v_r$, então $\alpha_i = \alpha'_i$ para $1 \leq i \leq r$.

2.3.3 Os Subespaços Fundamentais de uma Matriz

Seja $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$ e $b \in \mathbb{R}^m$. Então, o sistema linear $Ax = b$ pode ser resolvido se e somente se o vetor b puder ser expresso como uma combinação das colunas de A . Esse conjunto que consiste em todas as combinações das colunas de A , denotado por $\mathfrak{R}(A)$ é um subespaço de \mathbb{R}^m . Quando $b = 0$ sempre existe a solução particular $x = 0$, mas podem existir infinitas outras soluções. O conjunto de soluções para $Ax = 0$ é ele mesmo um espaço vetorial - o espaço nulo de A . O espaço nulo de uma matriz consiste em todos os vetores x tais que $Ax = 0$, é denotado por $\mathfrak{N}(A)$ e é um subespaço de \mathbb{R}^n . O espaço das linhas de A é gerado pelas linhas de A , que é o espaço das colunas de A^T , denotado por $\mathfrak{R}(A^T)$. O espaço nulo de A^T contém todos os vetores y tais que $A^T y = 0$, e é escrito como $\mathfrak{N}(A^T)$.

Resumindo, os quatro subespaços fundamentais de uma matriz A são:

1. o espaço das colunas de A o qual é denotado por $\mathfrak{R}(A)$;
2. o espaço nulo de A o qual é denotado por $\mathfrak{N}(A)$;
3. o espaço das linhas de A o qual é denotado por $\mathfrak{R}(A^T)$;
4. o espaço nulo de A^T o qual é denotado por $\mathfrak{N}(A^T)$.

2.3.4 Dimensão

A escolha de uma base B para o espaço vetorial V não é única, i.e, existem possibilidades diferentes de escolher uma base para V , mas há um ponto em comum entre todas estas escolhas. É uma propriedade intrínseca ao próprio subespaço, a qual resulta do seguinte teorema:

Teorema 4 *Sejam B_1 e B_2 quaisquer duas bases para o espaço vetorial V . Então B_1 e B_2 contêm o mesmo número de vetores, o qual é característico de todas as bases de V . Esse número expressa o número de graus de liberdade do espaço e é chamado de dimensão de V .*

Diz-se que o posto de uma matriz é dado pelo número de suas linhas ou colunas linearmente independentes.

2.4 Normas de vetores

Em várias circunstâncias é importante ter alguma noção do “tamanho” de um vetor, quando refere-se ao tamanho e não ao número de suas componentes. Para isso, é preciso definirmos o que é uma norma.

Definição 6 Uma norma vetorial em V é um número não negativo, representado por $\|v\|$, associado a v satisfazendo as seguintes propriedades:

- i. $\|v\| > 0$ para $v \neq 0$, e $\|v\| = 0$ exatamente quando $v = 0$.
- ii. $\|\alpha v\| = |\alpha| \cdot \|v\|$ para qualquer escalar α e para todo vetor v .
- iii. $\|v + u\| \leq \|v\| + \|u\|$, para todo vetor v e u em V (Desigualdade Triangular).

Seja o vetor $v \in \mathbb{R}^n$, $v = (v_1, v_2, \dots, v_{n-1}, v_n)^T$. Cada uma das quantidades abaixo, denominadas de *norma - 1*, *norma - 2* e *norma - ∞* , respectivamente, define uma norma vetorial:

1. $\|v\|_1 = |v_1| + |v_2| + \dots + |v_{n-1}| + |v_n|$
2. $\|v\|_2 = \sqrt{v_1^2 + v_2^2 + \dots + v_{n-1}^2 + v_n^2}$
3. $\|v\|_\infty = \max_i |v_i|$, $i = 1, 2, \dots, n$.

2.5 Produto Interno

A definição a seguir, mostra que o conceito de ângulo entre dois vetores em subespaços vetoriais está contido no conceito do produto especial

$$\langle u, v \rangle = u_1 v_1 + u_2 v_2 + \dots + u_n v_n.$$

Desta forma, o produto especial acima define uma aplicação $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$.

Definição 7 Seja V um espaço vetorial real. Um produto interno em V é uma função $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ que designa para cada par ordenado de vetores $u, v \in V$ um número real, denotado por $\langle u, v \rangle$, satisfazendo as seguintes propriedades:

1. $\langle u, v \rangle = \langle v, u \rangle, \forall u, v \in V$.
2. $\langle \alpha u + \beta v, w \rangle = \alpha \langle u, w \rangle + \beta \langle v, w \rangle$ e $\langle w, \alpha u + \beta v \rangle = \alpha \langle w, u \rangle + \beta \langle w, v \rangle, \forall u, v, w \in V$
e $\forall \alpha, \beta \in \mathbb{R}$.
3. $\langle u, u \rangle > 0$ se $u \neq 0$, e $\langle u, u \rangle = 0$ se e somente se $u = 0$.

O ângulo θ entre dois vetores não nulos u e v é definido por seu cosseno:

$$\cos(\theta) = \frac{\langle u, v \rangle}{\langle u, u \rangle^{1/2} \langle v, v \rangle^{1/2}}.$$

Para as definições dadas mais adiante faz-se necessário introduzir um importante resultado conhecido como *Desigualdade de Schwartz*, o qual está enunciado a seguir.

2.5.1 Desigualdade de Schwartz

Teorema 5 *Seja V um espaço vetorial sobre \mathbb{R} , munido de um produto interno $\langle u, v \rangle$. Então*

$$|\langle u, v \rangle| \leq \|u\| \cdot \|v\|, \quad \forall u, v \in V.$$

2.5.2 Norma Induzida Pelo Produto Interno

Outro importante e também posteriormente utilizado é o conceito de norma induzida pelo produto interno. Este está definido logo a seguir.

Definição 8 *A norma induzida pelo produto interno em \mathbb{R}^n é uma função $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por*

$$\|v\| = \sqrt{\langle v, v \rangle}. \quad (2.1)$$

2.5.3 Matrizes Simétricas e Matrizes Hermetianas

Definição 9 *Uma matriz $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ para a qual $A^T = A$, tal que $(A)_{ij} = (A)_{ji}$ para todo $i, j = 1, 2, \dots, n$ é dita ser simétrica. Uma matriz $B : \mathbb{C}^n \rightarrow \mathbb{C}^n$ para a qual $B^H = B$, tal que $(B)_{ij} = \overline{(B)_{ji}}$ para todo $i, j = 1, 2, \dots, n$ é dita ser hermetiana. A primeira coluna da matriz B^H é o conjugado complexo da primeira linha de B e assim sucessivamente para as próximas colunas da matriz B^H .*

2.5.4 Normas de Matrizes

Sejam $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma transformação linear e $\|\cdot\|$ uma norma. Para todo vetor não-nulo $v \in \mathbb{R}^n$, o quociente $\frac{\|Av\|}{\|v\|}$ mede a ampliação causada pela transformação A , uma cota superior deste quociente mediria, portanto, o valor *supremo* de A .

Definição 10 *Seja A uma transformação linear de \mathbb{R}^n em \mathbb{R}^m . A norma induzida de A é definida por*

$$\|A\| = \sup_{\substack{v \neq 0 \\ v \in \mathbb{R}^n}} \left\{ \frac{\|Av\|}{\|v\|} \right\}.$$

Teorema 6 *Sejam $A, B, C : \mathbb{R}^n \rightarrow \mathbb{R}^m$ transformações lineares. Suponha que cada uma possua a norma definida por:*

$$\|\Omega\| = \sup_{\substack{v \neq 0 \\ v \in \mathbb{R}^n}} \left\{ \frac{\|\Omega v\|}{\|v\|} \right\}.$$

- i. $\|Av\| \leq \|A\| \cdot \|v\|$;
- ii. $\|A\| \geq 0$, $\|A\| = 0$ se e somente se $Av = 0$ para todo v ;
- iii. $\|kA\| = |k| \cdot \|A\|$ para todos os escalares reais k ;
- iv. $\|A + B\| \leq \|A\| + \|B\|$;
- v. $\|CA\| \leq \|C\| \cdot \|A\|$;
- vi. $\|I\| = 1$, onde $Iv = v$ para todo v em \mathbb{R}^n .

Definição 11 *Seja $A \in \mathbb{R}^{m \times n}$. Representaremos por $\|A\|_1$, $\|A\|_2$ e $\|A\|_\infty$ as normas correspondentes de A induzidas usando as normas vetoriais apropriadas tanto no domínio \mathbb{R}^n quanto no contradomínio \mathbb{R}^m . Isto é,*

$$\|A\|_1 = \max_{x \neq 0} \left\{ \frac{\|Ax\|_1}{\|x\|_1} \right\},$$

$$\|A\|_2 = \max_{x \neq 0} \left\{ \frac{\|Ax\|_2}{\|x\|_2} \right\},$$

$$\|A\|_\infty = \max_{x \neq 0} \left\{ \frac{\|Ax\|_\infty}{\|x\|_\infty} \right\}.$$

Teorema 7 *Seja A uma matriz $m \times n$. Então:*

- i. $\|A\|_1 = \max_j \sum_{i=1}^m |a_{ij}|$ (máximo das somas dos valores absolutos de colunas);*
- ii. $\|A\|_\infty = \max_i \sum_j |a_{ij}|$ (máximo das somas dos valores absolutos de linhas);*
- iii. $\|A\|_2 = \sqrt{\text{autovalor máximo de } A^H A}$.
= valor singular máximo de A .*

2.5.5 Norma de Frobenius

Outra norma de matriz de uso comum na álgebra linear é a norma de Frobenius. Essa norma não é induzida por norma de vetor e sua definição é dada pela expressão

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2},$$

onde a_{ij} representa o elemento da i -ésima linha e j -ésima coluna da matriz A .

2.6 Matriz de Projeção

Seja $A \in \mathbb{R}^{m \times n}$ e $b \in \mathbb{R}^m$, $b \notin \mathfrak{R}(A)$, ou seja, um vetor fora do espaço das colunas de A . A construção de uma linha perpendicular do ponto b até o $\mathfrak{R}(A)$ pode ser expressa em termos matriciais por $p = A(A^T A)^{-1} A^T b$. A matriz $A(A^T A)^{-1} A^T$ é a matriz de projeção. Assim, P denotada por:

$$P = A(A^T A)^{-1} A^T. \quad (2.2)$$

projeta qualquer vetor b no espaço $\mathfrak{R}(A)$.

A matriz de projeção P possui duas propriedades básicas:

p1. $P^2 = P$;

p2. $P^T = P$.

A imagem e o espaço nulo da projeção (2.2) é dado, respectivamente, pelos conjuntos

$$\begin{aligned} \text{Im}(P) &= \{w \in \text{Im}(P) : w = Pb, b \in \mathbb{R}^m\} \\ \mathfrak{N}(P) &= \{y \in \mathbb{R}^m : Py = 0\}. \end{aligned}$$

Sejam $w \in \text{Im}(P)$ e $b \notin \text{Im}(P)$. Se $z = (w - b)$ então $z \in \mathfrak{N}(P)$, pois $P(w - b) = (P^2b - Pb) \stackrel{(p1)}{=} Pb - Pb = 0$. Logo, o vetor $(w - b) \in \mathfrak{N}(P)$, i.e., $(I - P)b \in \mathfrak{N}(P)$. Se P é projeção então $(I - P)$ também é projeção a qual é chamada de *projeção complementar*. A projeção $(I - P)$ projeta um vetor qualquer de \mathbb{R}^m no espaço nulo de A .

Qualquer matriz simétrica com $P^2 = P$ representa uma projeção.

2.7 Ortogonalidade

É possível pensar em ortogonalidade como uma generalização do conceito de *perpendicularidade* em um espaço vetorial arbitrário munido de um produto interno. Para compreender o significado disso, considere o seguinte problema: dados uma reta r contendo a origem e um ponto Q não pertencente a r , encontre um ponto P em r mais próximo de Q .

A solução desse problema é caracterizada pela condição de que o vetor $QP = v - p$ é perpendicular ao $OP = p$. A solução p é caracterizada pela propriedade de que p é ortogonal a $PQ = v - p$. No contexto de espaços vetoriais munidos de produto interno, pode-se considerar problemas gerais de *quadrados mínimos*, em tais problemas é dado um vetor v em V e um subespaço W , quer-se encontrar um vetor em W mais próximo possível de v . Uma solução p para esse problema, tem que ser ortogonal a $v - p$. Uma descrição mais detalhada do problema de mínimos quadrados não será necessária neste trabalho, mais detalhes podem ser encontrados em [3].

Definição 12 *Seja V um espaço vetorial munido de um produto interno, $\langle \cdot, \cdot \rangle$, e $\| \cdot \|$ a norma induzida pelo produto interno, ver Def.(2.1), pg. 10.*

- i. *Dois vetores u e v são ortogonais se e somente se $\langle u, v \rangle = 0$.*
- ii. *Um conjunto de vetores é dito ser ortogonal se e somente se todo par de vetores do conjunto são ortogonais: $\langle u, v \rangle = 0$, para todo $u \neq v$ naquele conjunto.*
- iii. *Se um vetor não nulo u é usado para produzir $v = u/\|u\|$, tal que $\|v\| = 1$, então u foi normalizado para produzir o vetor normalizado v .*
- iv. *Um conjunto de vetores é dito ser ortonormal se e somente se o conjunto é ortogonal e $\|v\| = 1$ para todo v no conjunto.*

Pelas propriedades de produto interno, $\langle 0, v \rangle = \langle 0v, v \rangle = 0\langle v, v \rangle = 0$. Isto é, em qualquer espaço vetorial munido de um produto interno, o vetor $\mathbf{0}$ é ortogonal a todos os vetores desse espaço vetorial.

2.7.1 Projeções Ortogonais

Os conjuntos de vetores ortogonais introduzidos aqui são teoricamente e computacionalmente importantes. Daí, segue a necessidade de encontrar uma matriz ortogonal P tal que P projete algum vetor v que não pertence a um espaço ortogonal V , em V .

Teorema 8 *Sejam V um espaço vetorial munido de um produto interno e V_0 o subespaço de V gerado por um conjunto ortogonal*

$$S = \{v_1, v_2, \dots, v_q\}$$

de vetores não nulos. Defina a projeção ortogonal P_0 sobre V_0 como segue: $\forall v \in V$, faça

$$P_0v = \alpha_1v_1 + \dots + \alpha_qv_q,$$

onde

$$\alpha_i = \frac{\langle v_i, v \rangle}{\langle v_i, v_i \rangle}.$$

Assim,

1. $v - P_0v$ é ortogonal a todos os vetores $v_0 \in V_0$.
2. $P_0(u + v) = P_0u + P_0v, \forall u, v \in V$.
3. $P_0(\alpha v) = \alpha P_0v, \forall \alpha \in \mathbb{R}$ e $\forall v \in V$.

2.7.2 Base Ortogonal

Pode-se expressar um vetor v como uma combinação linear dos vetores de uma base ortogonal sem a necessidade de resolver equações para determinar os coeficientes dessa combinação. É possível simplesmente avaliar alguns produtos internos para obter os coeficientes diretamente. O teorema abaixo, resume esse fato.

Teorema 9 *Seja $B = \{v_1, v_2, \dots, v_q\}$ uma base ortogonal. Então a representação de um vetor v com respeito à base B pode ser escrita como segue:*

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_q v_q,$$

onde

$$\alpha_i = \frac{\langle v_i, v \rangle}{\langle v_i, v_i \rangle}.$$

Em geral, na otimização restrita, por exemplo, quando as restrições são do tipo $Ax = b$, necessita-se encontrar uma base para o espaço nulo de A , $\mathfrak{N}(A)$. Em alguns casos, pode ser conveniente encontrar uma base ortogonal para $\mathfrak{N}(A)$ [1]. Agora, surge a questão de como encontrar essa base ortogonal para o subespaço em questão. Um processo conhecido como *processo de ortogonalização de Gram-Schmidt* gera uma base ortogonal para o espaço. Duas formas de implementar esse processo serão apresentados mais adiante, são elas: Gram-Schmidt Clássico (C-GS) e o Gram-Schmidt Modificado (M-GS).

Processo de Ortogonalização de Gram-Schmidt

Dada uma base e um conjunto gerador S para um espaço vetorial V munido de um produto interno um dos processos usados para encontrar uma base ortogonal (ou ortonormal) para o espaço vetorial é o de Gram-Schmidt. O teorema seguinte enfatiza a seguinte idéia: o processo de Gram-Schmidt produz uma base ortogonal de um conjunto gerador e detecta se o conjunto gerador é linearmente independente.

Teorema 10 *Seja $S = \{v_1, v_2, \dots, v_q\}$ o gerador do espaço vetorial V munido de um produto interno. O processo de Gram-Schmidt é como segue. Sejam V_i o subespaço de V gerado por $\{v_1, v_2, \dots, v_i\}$ e P_i a projeção ortogonal sobre V_i ; se $V_i = \{0\}$, faça $P_i v = 0$ para todo v .*

1. *Define-se $u_1 = v_1$.*
2. *Para $2 \leq i \leq q$, define-se $u_i = v_i - P_{i-1} v_i$.*

As propriedades seguintes, cumprem-se para os vetores produzidos por esse processo:

- 1.1. *$B = \{u_1, \dots, u_q\}$ é um conjunto ortogonal gerador de V .*
- 1.2. *Para $1 \leq i \leq q$, $B_i = \{u_1, \dots, u_i\}$ é um conjunto ortogonal gerador de V_i , o subespaço gerado por $\{v_1, \dots, v_i\}$.*
- 1.3. *$u_i = 0$ se e somente se v_i é linearmente dependente dos vetores v_1, \dots, v_{i-1} .*
- 1.4. *Uma base ortogonal para V pode ser obtida de um conjunto gerador ortogonal B omitindo-se $u_i = 0$, se esse existir.*
- 1.5. *Se S é uma base para V , então B é uma base ortogonal para V .*

Algoritmo 1 C-GS

Dado um conjunto gerador v_1, \dots, v_q :

1. Defina-se $u_1 = v_1$.
2. Para $2 \leq i \leq q$, defina-se

$$u_i = v_i - \alpha_{1i}u_1 - \dots - \alpha_{i-1,i}u_{i-1},$$

onde

$$\alpha_{ji} = \frac{\langle u_j, v_i \rangle}{\langle u_j, u_j \rangle} \text{ se } u_j \neq 0 \text{ e } \alpha_{ji} = 0 \text{ se } u_j = 0.$$

Essa forma de implementar o processo de Gram-Schmidt é instável. No método de C-GS há uma grave perda de ortogonalidade entre os u_i 's calculados.

Algoritmo 2 M-GS

Dado um conjunto gerador v_1, \dots, v_q :

1. Defina-se $v_1^0, v_2^0, \dots, v_q^0$.
2. Para $1 \leq i \leq q$,

- 2.1. defina-se $u_i = v_i^{i-1}$;

- 2.2. defina-se $v_j^i = v_j^{i-1} - \frac{\langle u_i, v_j^{i-1} \rangle}{\langle u_i, u_i \rangle} u_i$ para $j = i + 1, \dots, q$.

No método de M-GS é feito um rearranjo dos cálculos feitos pelo C-GS o que traz vantagem ao procedimento computacional. Essa forma de implementar o processo de Gram-Schmidt é computacionalmente estável. Este algoritmo requer $2mn^2$ operações.

2.8 Fatoração de Matrizes

Em muitos problemas computacionais que utilizem álgebra linear, é útil representar uma matriz que possua alguma propriedade como um produto ou soma de matrizes. Assim, uma representação desse tipo é denominada *fatoração* ou *decomposição* da matriz original. Nesta seção, algumas das fatorações que são frequentemente utilizadas em problemas de otimização serão descritas na ordem dos seus graus de detalhes.

2.8.1 Fatoração LU

Para resolver um sistema linear não-singular

$$Ax = b, \quad (2.3)$$

uma aproximação direta é encontrar um sistema linear equivalente, no qual a matriz é do sistema seja triangular. Isto pode ser realizado gerando-se uma matriz não-singular M tal que MA seja triangular, desde que a solução do novo problema obtido

$$MAx = Mb$$

seja a mesma do problema (2.3).

Uma maneira sistemática para triangularizar A introduzindo zeros abaixo da diagonal é denominada *eliminação Gaussiana*. No k -ésimo estágio desse processo, as primeiras $k-1$ colunas foram transformadas na forma triangular superior. Para realizar o k -ésimo passo da eliminação, o elemento da j -ésima linha e k -ésima coluna será eliminado subtraindo-se um múltiplo m_{jk} , onde

$$m_{jk} = \frac{a_{jk}^{(k)}}{a_{kk}^{(k)}}, \quad j = k + 1, \dots, n.$$

O elemento $a_{kk}^{(k)}$ da matriz parcialmente reduzida é denominado de elemento pivô. Esse elemento tem uma regra especial porque o processo falhará se o pivô for zero em algum passo. Então $n - k$ números $\{m_{jk}\}$, $j = k + 1, \dots, n$ são referidos como os multiplicadores no k -ésimo estágio.

Depois da matriz ter sido reduzida a forma triangular superior, as mesmas transformações devem ser aplicadas para o vetor b de (2.3).

Os passos de eliminação podem ser descritos formalmente como pré-multiplicações de uma seqüência de matrizes elementares especiais na matriz original, cujo efeito é subtrair o multiplicador apropriado da linha pivô de todas as linhas abaixo dela.

Quando todos os pivôs são não-nulos, a eliminação Gaussiana é equivalente a aplicar uma seqüência especial $\{M_i\}$ de matrizes unitárias triangulares superiores a A pela esquerda, para obter uma matriz triangular superior U :

$$MA = M_{n-1} \dots M_2 M_1 A = U. \quad (2.4)$$

O formalismo de introduzir o conjunto de matrizes $\{M_i\}$ não significa que essas transformações são armazenadas como matrizes; ao passo que, apenas $n - i$ multiplicadores precisam ser armazenados para representar M_i .

Depois que A foi triangularizada, x é a solução do sistema triangular superior, ou seja,

$$MAx = Ux = Mb,$$

na qual o lado direito Mb é calculado aplicando a b as transformações que reduziram A .

Desde que M seja não-singular, (2.4) pode ser escrito como

$$A = M^{-1}U = M_1^{-1}M_2^{-1}\dots M_{n-1}^{-1}U. \quad (2.5)$$

A estrutura especial que as matrizes M_i possuem, indica que M^{-1} é uma matriz triangular inferior cuja j -ésima coluna simplesmente contém os multiplicadores usados no j -ésimo estágio da eliminação.

Definindo $M^{-1} = L$, de (2.5) tem-se uma representação de A como o produto de uma matriz triangular inferior unitária L e uma matriz triangular superior U :

$$A = LU.$$

Esta forma é denominada fatoração LU de A .

Resumidamente, para resolver $Ax = b$ pela eliminação Gaussiana, os passos são:

- i. calcular L e U tais que $A = LU$ (assumindo-se pivôs não nulos);
- ii. resolver $Ly = b$ (isto é equivalente a formar $y = Mb$);
- iii. resolver $Ux = y$.

Exceto para n pequeno, o principal esforço na eliminação Gaussiana surge de aproximadamente $\frac{1}{3}n^3$ operações para formar L e U , visto que dois sistemas triangulares nos passos (ii) e (iii) são resolvidos com aproximadamente n^2 operações.

2.8.2 Fatoração de Cholesky

Esse tipo de fatoração é utilizada para resolver o sistema (2.3) quando a matriz A é uma matriz simétrica e definida positiva.

A seguir, serão apresentadas as definições e alguns resultados a respeito de matrizes semidefinidas, definidas e indefinidas.

Matriz Definida Positiva

Definição 13 *Seja uma matriz simétrica $A \in \mathbb{R}^{n \times n}$. A matriz A é definida positiva se $x^T Ax > 0$ para todo vetor não-nulo x de \mathbb{R}^n .*

O teorema a seguir, estabelece algumas propriedades para que uma matriz seja ou não definida positiva.

Teorema 11 *Cada uma das cinco propriedades seguintes é uma condição necessária e suficiente para que uma matriz simétrica $A \in \mathbb{R}^{n \times n}$ seja definida positiva:*

1. $x^T Ax > 0$ para todo vetor x não nulo.
2. Todos os autovalores de A são positivos.
3. Todas as submatrizes triangulares superiores esquerdas A_k possuem determinantes positivos.
4. A possui um conjunto de n pivôs positivos.
5. Existe uma matriz R , com colunas linearmente independentes, tal que $A = R^T R$.

Demonstração. Suponha que x_i é um autovetor unitário associado ao autovalor λ_i . Então,

$$Ax_i = \lambda_i x_i \tag{2.6}$$

de modo que

$$x_i^T Ax_i = x_i^T \lambda_i x_i = \lambda_i, \tag{2.7}$$

pois $x_i^T x_i = 1$. Pela definição $x^T Ax > 0$, então, em particular, $x_i^T Ax_i > 0$ e, assim, $\lambda_i > 0$.

Agora, suponha $\lambda_i > 0$. Isso deve ser válido para todo vetor x e não somente para os autovalores de A . Como as matrizes simétricas possuem um conjunto completo de autovalores ortonormais [23, p. 371], é possível escrever qualquer vetor x como uma combinação linear, dada por:

$$c_1 x_1 + \cdots + c_n x_n.$$

Então,

$$Ax = c_1 Ax_1 + \cdots + c_n Ax_n = c_1 \lambda_1 x_1 + \cdots + c_n \lambda_n x_n. \quad (2.8)$$

Devido à ortogonalidade e à normalização $x_i^T x_i = 1$ e $x_i^T x_j = 0$, para $i \neq j$

$$x^T Ax = (c_1 x_1^T + \cdots + c_n x_n^T)(c_1 \lambda_1 x_1 + \cdots + c_n \lambda_n x_n) = c_1^2 \lambda_1 + \cdots + c_n^2 \lambda_n. \quad (2.9)$$

Se todo $\lambda_i > 0$, então a expressão (2.9) mostra que $x^T Ax > 0$. Portanto, a condição (2.7) implica na condição (2.6). As condições (2.8) e (2.9) e suas equivalências com a condição (2.9) serão provadas em três passos. Se (2.6) é verdadeiro então (2.8) também é. Primeiro, o determinante de qualquer matriz é o produto dos seus autovalores. Sabe-se que esses autovalores são positivos

$$\det(A) = \lambda_1 \cdot \lambda_2 \cdots \lambda_n > 0. \quad (2.10)$$

Para provar o mesmo resultado para todas as submatrizes A_k , verifica-se que, se A é definida positiva, então todas as A'_k s também são. Verificam-se os vetores cujas últimas $n - k$ componentes são nulas:

$$x^T Ax = \begin{bmatrix} x_k^T & 0 \end{bmatrix} \begin{bmatrix} A_k & * \\ * & * \end{bmatrix} \begin{bmatrix} x_k \\ 0 \end{bmatrix} = x_k^T A_k x_k. \quad (2.11)$$

Se $x^T Ax > 0$, $\forall x \neq 0$, então, em particular, $x_k^T A_k x_k, \forall x_k \neq 0$. Portanto, a condição (2.6) se verifica para A_k , e a submatriz permite os mesmos argumentos utilizados para A . Seus autovalores (que não são os mesmos λ_i) devem ser positivos, e seu determinante é o produto deles, de modo que os determinantes superiores esquerdos são positivos. Se (2.9) é verdadeira, então (2.10) também é. Existe uma relação direta entre os números $\det(A)$ e os pivôs. O k -ésimo pivô d_k é exatamente a razão entre $\det(A_k)$ e $\det(A_{k-1})$. Portanto, se os determinantes são todos positivos, então os pivôs também são e nenhuma permutação de linha é necessária para matrizes definidas positivas. Para verificar a condição (2.11), deve-se reconhecer $x^T R^T R x$ como a raiz quadrada de $\|R x\|^2$. Isto não pode ser negativo e também não pode ser nulo (a não ser que $x = 0$), porque R possui colunas linearmente independentes: se x é não nulo então $R x$ é não nulo. Portanto, $x^T Ax = \|R x\|^2$ é positivo e $A = R^T R$ é definida positiva. ■

Matriz Semidefinida Positiva

Definição 14 *Seja $A \in \mathbb{R}^{n \times n}$ uma matriz simétrica. A é dita ser semidefinida positiva se $x^T Ax \geq 0$ para todo vetor $x \in \mathbb{R}^n$.*

Cada uma das cinco propriedades seguintes é uma condição necessária e suficiente para que A seja semidefinida positiva:

1. $x^T Ax \geq 0$ para todo vetor x ;
2. Todos os autovalores de A satisfazem $\lambda_i \geq 0$;
3. Todas as submatrizes principais possuem determinantes não negativos;
4. A possui um conjunto de n pivôs $d_i \geq 0$;
5. Existe uma matriz R , possivelmente possuindo colunas linearmente dependentes, tal que $A = R^T R$.

Matriz Definida Negativa

Definição 15 *Seja $A \in \mathbb{R}^{n \times n}$ uma matriz simétrica. A é dita ser definida negativa se $x^T Ax < 0$ para todo vetor não nulo $x \in \mathbb{R}^n$.*

Cada uma das quatro propriedades seguintes é uma condição necessária e suficiente para que A seja definida negativa:

1. $x^T Ax < 0$ para todo vetor x ;
2. Todos os autovalores de A satisfazem $\lambda_i < 0$;
3. Todas as submatrizes triangulares superiores esquerdas A_k possuem determinantes negativos;
4. A possui um conjunto de n pivôs $d_i < 0$;

Matriz Semidefinida Negativa

Definição 16 *Seja $A \in \mathbb{R}^{n \times n}$ uma matriz simétrica. A é dita ser semidefinida negativa se $x^T Ax \leq 0$ para todo vetor $x \in \mathbb{R}^n$.*

Cada uma das quatro propriedades seguintes é uma condição necessária e suficiente para que A seja semidefinida negativa:

1. $x^T Ax \leq 0$ para todo vetor x ;
2. Todos os autovalores de A satisfazem $\lambda_i \leq 0$;
3. Todas as submatrizes principais possuem determinantes negativos;
4. A possui um conjunto de n pivôs $d_i \leq 0$;

Matrizes Indefinidas

As matrizes que não se enquadram nas definições anteriores são consideradas indefinidas.

Decomposição de Cholesky - $A = LL^T$

Uma matriz A simétrica e definida positiva de ordem n pode ser escrita como

$$A = LL^T, \quad (2.12)$$

onde L é uma matriz triangular inferior com elementos diagonais positivos. A matriz L^T é chamada de fator de Cholesky.

A fatoração de Cholesky pode ser calculada diretamente igualando elemento por elemento. Para a representação de $A = LL^T$, tem-se que:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} l_{11} & & & \\ l_{12} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{1n} & l_{2n} & \cdots & l_{nn} \end{pmatrix} \begin{pmatrix} l_{11} & l_{12} & \cdots & l_{1n} \\ & l_{22} & \cdots & l_{2n} \\ & & \ddots & \vdots \\ & & & l_{nn} \end{pmatrix} \quad (2.13)$$

Igualando os elementos que estão na primeira posição de cada matriz, tem-se:

$$a_{11} = l_{11}^2 \implies l_{11} = \sqrt{a_{11}}.$$

Igualando a primeira linha de ambos os lados em (2.13) segue que:

$$a_{12} = l_{11}l_{12}, \quad a_{13} = l_{11}l_{13}, \quad \dots, \quad a_{1n} = l_{11}l_{1n},$$

assim, todos os elementos da primeira linha de L^T podem ser calculados pois l_{11} já é conhecido.

Para elemento da segunda linha e segunda coluna, tem-se

$$a_{22} = l_{12}^2 + l_{22}^2.$$

Como l_{12} é conhecido, l_{22} pode ser calculado. Parte-se então para a segunda linha de L^T e assim sucessivamente.

A fatoração de Cholesky de (2.12) pode ser calculada com aproximadamente $\frac{1}{6}n^3$ multiplicações (e adições) e n raízes quadradas.

A fatoração de Cholesky é numericamente estável. Essa propriedade cumpre-se pela seguinte relação entre os elementos de L^T :

$$l_{12}^2 + l_{2k}^2 + \cdots + l_{kk}^2, \quad k = 1, \dots, n.$$

2.8.3 Fatoração QR

Suponha que v_1, v_2, \dots, v_q são vetores $p \times 1$ e considere a implementação tradicional do processo de Gram-Schmidt, pg. 16. Utilizando o algoritmo (1) em u_j , tem-se que

$$v_j = u_1\alpha_{1j} + u_2\alpha_{2j} + \dots + u_{j-1}\alpha_{j-1,j} + u_j \quad j = 1, 2, \dots, q.$$

Na notação matricial particionada, isso é justamente

$$[v_1 \ \cdots \ v_q] = [u_1 \ \cdots \ u_q] \begin{bmatrix} 1 & \alpha_{12} & \cdots & \alpha_{1q} \\ 0 & 1 & \cdots & \alpha_{2q} \\ \vdots & & \cdots & \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Sejam $Q_0 = [u_1 \ \cdots \ u_q]$ uma matriz $p \times p$ e R_0 a matriz triangular superior unitária $q \times q$. Desde que $A = [v_1 \ \cdots \ v_q]$ pode-se reescrevê-la como

$$A = Q_0 R_0. \quad (2.14)$$

Os u_i 's em Q_0 são construídos via algoritmo (1), e são, portanto, mutuamente ortogonais; isto é, Q_0 tem colunas ortogonais, algumas das quais devem ser iguais a 0. Seja Q e R as matrizes obtidas pela eliminação das colunas nulas de Q_0 e linhas correspondentes de R_0 , dividindo cada coluna não nula de Q_0 pela *norma* - 2, e multiplicando cada linha correspondente de R_0 pela mesma *norma* - 2. Então, (2.14) é substituída por

$$A = QR \quad (2.15)$$

com R triangular superior e Q tendo colunas ortonormais. A forma (2.14) é chamada de *decomposição QR não-normalizada* de A , enquanto (2.15) é a *decomposição QR normalizada*.

Para formalizar a descrição anterior, considere o seguinte teorema.

Teorema 12 *Seja $A \in \mathbb{R}^{p \times q}$, com posto k . Então:*

1. *A pode ser escrita em sua decomposição QR não normalizada como $A = Q_0 R_0$, onde:*

1.2. $Q_0 \in \mathbb{R}^{p \times p}$ e tem colunas ortogonais (das quais k são não nulos e $q - k$ são zeros) que geram o espaço coluna de A .

1.3. $R_0 \in \mathbb{R}^{q \times q}$, triangular superior unitária e não-singular.

1.4. A norma-2 da i -ésima coluna de Q_0 é igual a distância da i -ésima coluna de A para o espaço gerador pelas primeiras $i - 1$'s colunas de A .

2. A pode ser escrita em sua decomposição QR normalizada como $A = QR$, onde:

2.1. $Q \in \mathbb{R}^{p \times k}$ e tem colunas ortonormais que geram o espaço coluna de A .

2.2. $R \in \mathbb{R}^{k \times q}$, triangular superior, e tem posto k .

2.3. Se $k = q$, então $|\langle R \rangle_{ii}|$ é igual a distância da i -ésima coluna de A para o espaço gerado pelas primeiras $i - 1$'s colunas de A .

Demonstração. Ver [23], pg. 237. ■

2.9 Autovalores e Autovetores

Para qualquer $A \in \mathbb{R}^{n \times n}$, um escalar λ é um autovalor de A se a equação

$$Ax = \lambda x,$$

tem uma solução não nula x denominado de autovetor de A correspondente a λ . Visto que uma matriz B é singular se e somente se o determinante, $\det B$, é zero, segue que os autovalores de A são precisamente as n raízes (considerando a multiplicidade) do polinômio característico

$$\det(A - \lambda I) = 0,$$

onde I é a matriz identidade $n \times n$.

2.10 Teoria da Perturbação e Número de Condição

O sistema linear

$$Ax = b, \tag{2.16}$$

tem solução única para todo b somente se A é quadrada e não-singular. Antes de considerar como resolver (2.16), é interessante ver como a solução é afetada por pequenas perturbações no vetor b e nos elementos da matriz.

A solução exata de (2.16) para A não singular é dado por

$$x = A^{-1}b.$$

Suponha que o lado direito de (2.16) é perturbado por $b + \delta b$, e que a solução exata do sistema perturbado é $x + \delta x$, i. e.,

$$A(x + \delta x) = b + \delta b,$$

onde " δ " denota uma pequena perturbação em um vetor ou matriz. Por isso,

$$x + \delta x = A^{-1}(b + \delta b),$$

e visto que $x = A^{-1}b$,

$$\delta x = A^{-1}\delta b.$$

Para medir δx , utiliza-se a propriedade de vetor compatível e normas de matriz:

$$\|\delta x\| \leq \|A^{-1}\| \|\delta b\|, \quad (2.17)$$

cumprindo na igualdade para algum vetor δb . A perturbação na solução exata é ainda e tem limite superior igual a $\|A^{-1}\|$ vezes a perturbação em b .

Para determinar o efeito relativo desta mesma perturbação, observe que

$$\|b\| \leq \|A\| \|x\|. \quad (2.18)$$

Combinando as desigualdades (2.17) e (2.18) e reorganizando-as, segue

$$\frac{\|\delta x\|}{\|x\|} \leq \|A\| \|A^{-1}\| \frac{\|\delta b\|}{\|b\|}. \quad (2.19)$$

Se a matriz A em (2.16) é perturbada por δA , um processo similar ao que ocorreu com b , produz $(A + \delta A)(x + \delta x) = b$. Essa equação pode ser reescrita como

$$\delta x = -A^{-1}\delta A(x + \delta x),$$

tal que

$$\|\delta x\| \leq \|A^{-1}\| \|\delta A\| \|x + \delta x\|$$

ou

$$\frac{\|\delta x\|}{\|x + \delta x\|} \leq \|A^{-1}\| \|\delta A\|.$$

Quando a mudança $\|\delta A\|$ é considerada relativa, isto torna

$$\frac{\|\delta x\|}{\|x + \delta x\|} \leq \|A\| \|A^{-1}\| \frac{\|\delta A\|}{\|A\|}. \quad (2.20)$$

Em ambos (2.19) e (2.20), a mudança relativa na solução exata é limitado pelo fator $\|A\| \|A^{-1}\|$ multiplicado pela perturbação relativa aos dados. O número $\|A\| \|A^{-1}\|$ é definido como o *número de condição* de A e é denotado por $\text{cond}(A)$.

Como $\|I\| = 1$ para alguma norma induzida, e $I = AA^{-1}$, segue que

$$1 = \|I\| \leq \|A\| \|A^{-1}\|,$$

de modo que

$$\text{cond}(A) \geq 1$$

para toda matriz.

O número de condição de uma matriz A indica o *efeito máximo* de perturbações em b ou A na solução exata de (2.16). A matriz A é dita ser *mal condicionada* se $\text{cond}(A)$ é "grande" e A é dita ser *bem condicionada* se, ao contrário, $\text{cond}(A)$ é "pequeno".

3

Tópicos em Análise do \mathbb{R}^n e Otimização

Um problema de otimização começa com um conjunto de variáveis ou parâmetros independentes, e freqüentemente incluem condições que definem valores aceitáveis para as variáveis. Tais condições são denominadas *restrições* do problema. Outra componente essencial de um problema de otimização é uma medida do "benefício", denominado função objetivo, a qual depende sobre algum modo das variáveis. A solução de um problema de otimização é um conjunto de valores permitidos das variáveis para o qual a função objetivo assume um valor "ótimo". Em termos matemáticos, a otimização, usualmente, envolve *maximização ou minimização*; e.g., deseja-se maximizar o lucro ou benefício, ou minimizar as despesas ou perdas.

Este capítulo é devotado primeiramente uma revisão a funções de várias variáveis as quais são tomadas como funções objetivos nos problemas envolvendo otimização restrita ou irrestrita. Em um segundo momento, será apresentada as condições de otimalidade de um problema de programação não linear irrestrito e irrestrito.

As técnicas usadas para resolver problemas de programação não linear, em alguns casos, são baseadas em um modelo quadrático da função, e.g. *método de Newton* [13], [20], [9]. Existem duas justificativas para a escolha de um *modelo quadrático*: sua simplicidade e, mais importante, o sucesso e eficiência na prática de tais.

Também serão introduzidos neste capítulo os conceitos de convergência e rapidez (velocidade) da mesma caso esta ocorra.

Por último serão apresentados os conceitos e as propriedades das funções quadráticas. Depois da descrição das funções quadráticas, serão apresentados o método das direções conjugadas e o método do CG.

3.1 Funções de Classe C^k

Definição 17 A função $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ é dita ser da classe $C^{(k)}$ em D , isto é, $f \in C^{(k)}(D)$ se tem todas as derivadas contínuas até ordem k em D .

3.1.1 Gradiente

Considere uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$, suave. Se a primeira derivada parcial de f com respeito às n variáveis existem e são contínuas, o vetor coluna dessas n derivadas parciais é chamado de *gradiente* de f em x , $\nabla f(x)$.

Definição 18 Uma função contínua $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é dita ser diferencialmente contínua em $x \in \mathbb{R}^n$, se $\frac{\partial f}{\partial x_i}$ existe e é contínua, $i = 1, \dots, n$; o gradiente de f em x é então definido como

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix}.$$

A função f é dita ser continuamente diferenciável em $D \subset \mathbb{R}^n$, denotado por $f \in C^1(D)$, se é continuamente diferenciável para todo ponto em D .

3.1.2 Hessiana

Considere uma função contínua $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Se a primeira e a segunda derivadas parciais de f com respeito às n variáveis existem e é contínua, a matriz contendo essas derivadas parciais é chamada de *Hessiana* de f em x , $\nabla^2 f(x)$. Segue a seguinte definição:

Definição 19 Uma função continuamente diferenciável em \mathbb{R}^n $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é dita ser duas vezes continuamente diferenciável em $x \in \mathbb{R}^n$, se $\frac{\partial^2 f}{\partial x_i \partial x_j}$ existe e é contínua, $1 \leq i, j \leq n$; a Hessiana de f em x é então definida como a matriz quadrada de ordem n , cujo elemento da i – ésima linha e j – ésima coluna é

$$\nabla^2 f(x)_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}, \quad 1 \leq i, j \leq n.$$

A função f é duas vezes continuamente diferenciável em $D \subset \mathbb{R}^n$, denotado por $f \in C^2(D)$, se é continuamente diferenciável em todo ponto de D .

3.1.3 Teorema de Taylor

Os resultados de análises que são freqüentemente utilizados em otimização vêm do grupo dos *teoremas de Taylor ou teoremas do Valor Médio*. Esses teoremas são de fundamental importância pois eles mostram que, se a função e suas derivadas forem conhecidas em um único ponto, então podem-se calcular aproximações para a função em todos os pontos numa vizinhança imediata daquele ponto.

Teorema 13 *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Se $f \in C^{(k)}$, então existe um número real θ ($0 \leq \theta \leq 1$), tal que*

$$f(x+h) = f(x) + hf'(x) + \frac{1}{2}h^2f''(x) + \dots + \frac{1}{(k-1)!}h^{k-1}f^{(k-1)}(x) + \frac{1}{k!}h^k f^{(k)}(x+\theta h)$$

onde $f^{(k)}(x)$ denota a k -ésima derivada de f no ponto x .

A série de Taylor é uma ferramenta usada para aproximar uma função f ao redor de um ponto x_0 por um polinômio.

Definição 20 *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função de classe C^2 , i.e., $f \in C^2$. Então, a aproximação da função f próxima ao ponto x_0 pela série de Taylor é escrita como*

$$f(x_0+p) = f(x_0) + p^T \nabla f(x_0) + \frac{1}{2}p^T \nabla^2 f(x_0)p$$

3.1.4 Teorema Fundamental do Cálculo ou Teorema de Newton

Teorema 14 *Seja $f \in C^2$ e D um conjunto convexo aberto em \mathbb{R}^n . Se $x, y \in D$, então*

$$f(y) - f(x) = \int_0^1 \langle \nabla f(x + t(y-x)), y-x \rangle dt.$$

Demonstração. Considere uma função unidimensional $\phi(t) = f(x + t(y-x))$. O teorema fundamental do Cálculo afirma que

$$\phi(1) - \phi(0) = \int_0^1 \phi'(t) dt.$$

Visto que $\phi(1) = f(y)$, $\phi(0) = f(x)$ e

$$\phi'(t) = \langle \nabla f(x + t(y-x)), y-x \rangle,$$

tem-se que $\phi(1) - \phi(0) = \int_0^1 \langle \nabla f(x + t(y-x)), y-x \rangle dt$. ■

3.1.5 Derivada Direcional

Se $z = f(x, y) : f : \mathbb{R}^2 \rightarrow \mathbb{R}$ duas vezes continuamente diferenciável e se p um vetor unitário do \mathbb{R}^2 , $p = (p_1, p_2)^T$. Então, a derivada direcional de f em x na direção de p , é denotada por $\frac{\partial f}{\partial p}(x) = \frac{\partial f}{\partial x} \cdot p_1 + \frac{\partial f}{\partial y} \cdot p_2$.

Teorema 15 *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ continuamente diferenciável em um conjunto convexo $S \subset \mathbb{R}^n$. Então, para $x \in S$ e alguma perturbação não nula $p \in \mathbb{R}^n$, a derivada direcional de f em x na direção de p , definida por*

$$\frac{\partial f}{\partial p}(x) = \lim_{\varepsilon \rightarrow 0} \frac{f(x + \varepsilon p) - f(x)}{\varepsilon},$$

existe e é igual $\nabla f(x)^T p$. Para $x, x + p \in S$,

$$f(x + p) = f(x) + \int_0^1 \nabla f(x + tp)^T p \, dt = f(x) + \int_x^{x+p} \nabla f(z) \, dz, \quad (3.1)$$

e existe um $z \in (x, x + p)$ tal que

$$f(x + p) = f(x) + \nabla f(z)^T p. \quad (3.2)$$

Demonstração. Parametrizando f ao longo da linha formada pelo intervalo $(x, x + p)$ tem-se a função g de uma variável definida por

$$g : \mathbb{R} \rightarrow \mathbb{R}, \quad g(t) = f(x + tp).$$

Definindo $x(t) = x + tp$. Então, pela regra da cadeia, para $0 \leq t \leq 1$,

$$\frac{dg}{dt} = \sum_{i=1}^n \frac{\partial f(x(t))}{\partial x(t)_i} \frac{dx(t)_i}{dt} \quad (3.3)$$

$$\begin{aligned} &= \sum_{i=1}^n \frac{\partial f}{\partial x_i} p_i \\ &= \nabla f(x + tp)^T p. \end{aligned} \quad (3.4)$$

Fazendo $t = 0$ reduz-se (3.4) para

$$\frac{\partial f(x)}{\partial p} = \nabla f(x)^T p.$$

Pelo teorema fundamental do cálculo ou teorema de Newton (14),

$$g(1) = g(0) + \int_0^1 g'(t) dt$$

o qual, pela definição de g e (3.4), é equivalente a

$$f(x+p) = f(x) + \int_0^1 f(x+tp)^T p dt,$$

o que prova (3.1). Finalmente, pelo resultado desse teorema para funções de uma variável, tem-se que

$$g(1) = g(0) + g'(\xi), \quad \xi \in (0, 1)$$

e,

$$f(x+p) = f(x) + \nabla f(x + \xi p)^T p, \quad \xi \in (0, 1),$$

o que prova (3.2). ■

3.1.6 Direção de Descida

Definição 21 As direções $d \in \mathbb{R}^n$, tais que

$$\nabla f(x)^T d < 0$$

são chamadas direções de descida a partir de x .

Teorema 16 Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f \in C^1$ e $x \in \mathbb{R}^n$ tal que $\nabla f(x) \neq 0$, $d \in \mathbb{R}^n$, $\nabla f(x)^T d < 0$. Então, existe $\bar{\alpha} > 0$ tal que

$$f(x + \alpha d) < f(x), \quad \forall \alpha \in (0, \bar{\alpha}).$$

Demonstração. Considere $\phi : \mathbb{R} \rightarrow \mathbb{R}$,

$$\phi(\alpha) = f(x + \alpha d)$$

e

$$\phi(0) = f(x).$$

Pela regra da cadeia,

$$\phi'(\alpha) = \nabla f(x + \alpha d)^T d$$

$$\phi'(0) = \nabla f(x)^T d$$

e

$$\phi'(0) = \lim_{\alpha \rightarrow 0} \frac{\phi(0+\alpha) - \phi(0)}{\alpha}.$$

Para $0 < \alpha < \bar{\alpha}$, com $\bar{\alpha}$ suficientemente pequeno, o sinal de $\phi'(0)$ e $\phi(\alpha) - \phi(0)$ deve ser o mesmo, ou seja, negativo. Como $\phi'(0) = \nabla f(x)^T d < 0$, logo

$$\phi(\alpha) - \phi(0) < 0,$$

$$f(x + \alpha d) - f(x) < 0$$

$$f(x + \alpha d) < f(x).$$

■

3.1.7 Aproximação de derivadas por diferenças finitas

A aplicabilidade da expansão de uma função suave em Taylor, ver teorema (13) permitem que os valores das derivadas sejam aproximados, o que é melhor, computacionalmente, do que calculá-los analiticamente. Sejam $F : \mathbb{R}^n \rightarrow \mathbb{R}$, uma função com $F \in C^2$, e h_j o tamanho do intervalo da diferença finita associada com a j -ésima componente de x , e e_j o vetor com 1 na j -ésima posição e 0 nas demais. Então:

$$F(x + h_j e_j) = F(x) + h_j g(x)^T e_j + \frac{1}{2} h_j^2 e_j^T G(x + \theta_1 h_j e_j) e_j,$$

onde $0 \leq \theta_1 \leq 1$, o qual estima um limite superior para o erro de truncamento. Os termos $g(x)^T e_j$ e $e_j^T G(x + \theta_1 h_j e_j) e_j$ são apenas o j -ésimo elemento de g e o j -ésimo elemento da diagonal de $e_j^T G(x + \theta_1 h_j e_j)$ respectivamente. Dado:

$$g(x)_j = \frac{1}{h_j} (F(x + h_j e_j) - F(x)) + O(h_j), \quad (3.5)$$

a qual representa a expressão da diferença pela direita para $g(x)$. As fórmulas das diferenças pela esquerda e central para g são, respectivamente:

$$\frac{1}{h_j}(F(x) - F(x - h_j e_j)) \quad (3.6)$$

$$\frac{1}{2h_j}(F(x + h_j e_j) - F(x - h_j e_j)). \quad (3.7)$$

O erro de truncamento de (3.5) e (3.6) são de *ordem 1*, o qual difere de (3.7) que é de *ordem 2*. Também é possível obter a aproximação por diferenças finitas da derivada de segunda ordem $G(x)$. Logo, $G_{ij}(x)$ de *ordem 1* é dada por:

$$\frac{1}{h_i h_j}(F(x + h_j e_j + h_i e_i) - F(x + h_j e_j) - F(x + h_i e_i) + F(x)). \quad (3.8)$$

3.2 Condições de Otimalidade

Os problemas de otimização envolvem minimizar ou maximizar uma função objetivo, sujeita a um conjunto de restrições impostas nas variáveis. Em, essencialmente, todos os problemas de interesse as restrições serão expressas em termos de relações envolvendo as variáveis das funções contínuas. O problema geral a ser considerado é conhecido como *otimização linear restrita* e é expresso em termos matemáticos por:

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimizar}} & f(x) \\ \text{s.a.} & c_i(x) = 0, \quad i = 1, 2, \dots, m'; \\ & c_i(x) \geq 0, \quad i = m' + 1, \dots, m. \end{array} \quad (3.9)$$

Um ponto \hat{x} é dito ser viável se satisfaz todas as restrições do problema (3.9). O conjunto de todos os pontos viáveis é denominado região viável. Um problema para o qual não existe pontos viáveis é chamado de problema inviável.

Para selecionar um método eficiente para resolver um problema particular da forma (3.9), é necessário analisar e classificar o problema. Antes de considerar um método para resolver um problema é necessário definir uma "solução para o (3.9). Para tal solução existir considere os seguintes teoremas e as definições de conjuntos e funções convexas.

3.2.1 Teorema de Weierstrass

Teorema 17 *Seja S um conjunto não vazio, compacto, e seja $f : S \subset \mathbb{R}^n \rightarrow \mathbb{R}$ contínua em S . Então, o problema*

$$\min_{x \in S} f(x) \quad (3.10)$$

atinge seu mínimo, isto é, existe um minimizador para o problema (3.10).

Demonstração. Visto que f é contínua em S e S é fechado e limitado, f é limitado inferiormente em S . Consequentemente, visto que $S \neq \emptyset$, existe o menor limite inferior $\alpha \equiv \inf\{f(x) : x \in S\}$. Agora, seja $0 < \varepsilon < 1$, e considere o conjunto $S_k = \{x \in S : \alpha + \varepsilon^k\}$ para todo $k = 1, 2, \dots$. Pela definição de ínfimo, $S_k \neq \emptyset$ para todo k e então constrói-se uma seqüência de pontos $\{x_k\} \subseteq S$ selecionando um ponto $x_k \in S_k$ para todo $k = 1, 2, \dots$. Como S é limitado, existe uma subseqüência convergente $\{x_k\}_K \rightarrow \bar{x}$, indexado pelo conjunto K . Como S é fechado, tem-se $\bar{x} \in S$; pela continuidade da função f , desque que $\alpha \leq f(x_k) \leq \alpha + \varepsilon^k$ para todo k , tem-se que $\alpha = \lim_{k \rightarrow +\infty, k \in K} f(x_k) = f(\bar{x})$. Portanto, foi mostrado que existe uma solução $\bar{x} \in S$ tal que $f(\bar{x}) = \alpha = \inf\{f(x) : x \in S\}$, e então \bar{x} é um minimizador. ■

3.2.2 Conjuntos Convexos

Definição 22 *Um conjunto $S \subset \mathbb{R}^n$ é convexo se $\forall x, y \in S$ o segmento de reta que une x e y também está em S .*

Logo, o subconjunto S de \mathbb{R}^n é convexo se e somente se para todo x e y em S e todo λ com $0 \leq \lambda \leq 1$, o vetor $\lambda x + (1 - \lambda)y$ está também em S .

Teorema 18 *Seja S um conjunto convexo em \mathbb{R}^n . Faça $x^{(1)}, x^{(2)}, \dots, x^{(k)}$ pertencer a S . Se $\lambda_1, \lambda_2, \dots, \lambda_k$ são números não negativos cuja soma é unitária, então a combinação convexa $\sum_{i=1}^k \lambda_i x^{(i)}$ está também em S .*

3.2.3 Funções Convexas

A classe de funções chamadas de funções convexas incluem a classe de funções lineares que tem ampla variedade de aplicações.

Uma função $f(x)$ definida em um intervalo I da reta real é convexa se a corda que liga quaisquer pontos $(x_1, f(x_1))$ e $(x_2, f(x_2))$ no gráfico de $f(x)$ está sobre ou acima do gráfico de $f(x)$.

Se tal corda está sobre ou abaixo do gráfico de $f(x)$, então $f(x)$ é *côncava* em I .

Nota-se que se x_1 e x_2 são números reais, então um número u está entre x_1 e x_2 se e somente se existe λ com $0 \leq \lambda \leq 1$ tal que

$$u = \lambda x_1 + (1 - \lambda)x_2.$$

Se essa observação for aplicada para pontos no eixo x e no eixo y , então tem-se que uma função $y = f(x)$ é convexa em I se e somente se

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$$

para todo $x_1, x_2 \in I$ e todo $0 \leq \lambda \leq 1$. Se a desigualdade estrita mantém-se mesmo com x_1 e x_2 sendo pontos distintos em I , ou seja,

$$f(\lambda x_1 + (1 - \lambda)x_2) < \lambda f(x_1) + (1 - \lambda)f(x_2)$$

então $f(x)$ é estritamente convexa.

Definição 23 *Suponha que f é uma função definida em um conjunto convexo $S \subset \mathbb{R}^n$. Então:*

(i) *a função f é convexa em S se*

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y),$$

para todo $x, y \in S$ e todo λ com $0 \leq \lambda \leq 1$;

(ii) *a função f é estritamente convexa em S se*

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y),$$

para todo $x, y \in S$ com $x \neq y$ e todo λ com $0 < \lambda < 1$.

Teorema 19 *Todo minimizador local de uma função convexa f definida em um subconjunto S de \mathbb{R}^n é também um minimizador global. Todo minimizador local de uma função estritamente convexa f definida em um conjunto convexo $S \subset \mathbb{R}^n$ é o único minimizador global estrito de f em S .*

Demonstração. Suponha que x^* é um minimizador local para a função convexa f em S . Então existe um número positivo r tal que $f(x) \geq f(x^*)$ sempre que $x \in S$ e $\|x - x^*\| < r$.

Dado algum $y \in S$, quer provar-se que $f(y) \geq f(x^*)$. Para este fim, selecciona-se λ com $0 < \lambda < 1$ e tão pequeno que $x^* + \lambda(y - x^*) \in S$ e

$$\|x^* + \lambda(y - x^*) - x^*\| < r.$$

Por conseguinte,

$$f(x^*) \leq f(x^* + \lambda(y - x^*)) = f(\lambda y + [1 - \lambda]x^*) \leq \lambda f(y) + [1 - \lambda]f(x^*),$$

pois f é convexo em S . Nota-se que a última desigualdade é estrita se $y \neq x^*$ e f é estritamente convexa. Agora, subtrai-se $f(x^*)$ de ambos os lados da desigualdade anterior e divide-se o resultado por λ

$$\frac{f(x^*) - f(x^*)}{\lambda} \leq \frac{\lambda f(y) + [1 - \lambda]f(x^*) - f(x^*)}{\lambda},$$

para obter

$$0 \leq f(y) - f(x^*)$$

desigualdade estrita se f é estritamente convexa e $y \neq x^*$. Isto estabelece o resultado desejado. ■

3.2.4 Caracterização de um ponto de mínimo

Um ponto $\bar{x} \in \mathbb{R}^n$ é solução de (3.9) se \bar{x} é um ponto que pertença a região viável. Então, somente pontos viáveis podem ser ótimos.

Seja x^* um ponto viável para o problema (3.9), e defina $N(x^*, \delta)$ como sendo o conjunto dos pontos viáveis em uma vizinhança δ de x^* .

Definição 24 O ponto x^* é um mínimo local se existe $\delta > 0$ tal que:

1. $f(x)$ é definido em $N(x^*, \delta)$; e
2. $f(x^*) < f(y)$, $\forall y \in N(x^*, \delta)$, $y \neq x^*$.

3.2.5 Condições de Otimalidade para uma Função com Restrições Lineares de Igualdade

Considere o problema de minimizar uma função suave sujeita a um conjunto de restrições lineares de igualdade, i.e.,

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimizar}} & f(x) \\ \text{s.a} & Ax = b. \end{array} \quad (3.11)$$

A i -ésima linha da matriz $A \in \mathbb{R}^{m \times n}$ será denotada por a_i^T , a qual contém os coeficientes da i -ésima restrição linear:

$$a_i^T x = a_{i1}x_1 + \dots + a_{in}x_n = b_i.$$

Um ponto viável x^* é um mínimo local para (3.11) se e somente se $f(x^*) \leq f(x)$ para todo x viável em alguma vizinhança de x^* . Não existe ponto viável se as restrições são inconsistentes, então assume-se que b pertence ao range de A .

Pelo fato das restrições serem lineares, as propriedades de subespaços lineares tornam possível de estabelecer uma caracterização simples de todos os movimentos viáveis a partir de um ponto viável.

Sejam \bar{x} e \tilde{x} dois pontos viáveis do problema (3.11) e p uma direção viável tomada como $p = \bar{x} - \tilde{x}$. Então, p é ortogonal a todas as linhas de A , i.e., p satisfaz

$$Ap = 0. \quad (3.12)$$

Qualquer vetor p para o qual (3.12) se cumpre é chamada de direção viável com respeito às restrições lineares de igualdade do problema (3.11).

Qualquer passo de um ponto viável ao longo da direção p não viola as restrições, visto que $A(\hat{x} + \alpha p) = A\hat{x} = b$.

Seja $Z \in \mathbb{R}^{m \times n-m}$ uma base para o $\mathcal{N}(A)$, então $AZ = 0$, e toda direção viável pode ser escrita como uma combinação linear das colunas de Z . Portanto, se p satisfaz (3.12), p pode ser escrito como

$$p = Zp_z,$$

onde p_z é um vetor pertencente ao \mathbb{R}^{n-m}

Fazendo a expansão em Taylor teorema (13) de f em relação a x^* ao longo de uma direção viável $p = Z p_z$:

$$f(x^* + \epsilon Z p_z) = f(x^*) + \epsilon p_z^T Z^T \nabla f(x^*) + \frac{1}{2} \epsilon^2 p_z^T Z^T \nabla^2 f(x^* + \epsilon \theta p) Z p_z \quad (3.13)$$

onde θ satisfaz $0 \leq \theta \leq 1$, e ϵ é um escalar positivo. A expansão (3.13) mostra que $p_z^T Z^T \nabla f(x^*)$ é negativo, então toda vizinhança de x^* conterà pontos viáveis com valor da função estritamente inferior. Ainda, a condição necessária para que x^* seja um mínimo local da (3.11) é que $p_z^T Z^T \nabla f(x^*)$ anula-se para todo p_z , o que implica que

$$Z^T \nabla f(x^*) = 0. \quad (3.14)$$

O vetor $Z^T \nabla f(x^*)$ é chamado de gradiente projetado de f em x^* .

O resultado (3.14) implica que $\nabla f(x^*)$ pode ser uma combinação linear das linhas de A , isto é,

$$\nabla f(x^*) = \sum_{i=1}^m \alpha_i \lambda_i^* = A^T \lambda^*,$$

para algum vetor λ^* , que é chamado vetor multiplicador de Lagrange. Os multiplicadores de Lagrange são únicos somente se as linhas de A são linearmente independentes.

Como $Z^T \nabla f(x^*) = 0$, (3.13) torna-se

$$f(x^* + \epsilon Z p_z) = f(x^*) + \frac{1}{2} \epsilon^2 p_z^T Z^T \nabla^2 f(x^* + \epsilon \theta p) Z p_z$$

indicando que se a matriz $Z^T \nabla^2 f(x^*) Z$ é indefinida, toda vizinhança de x^* contém pontos viáveis com um valor estritamente inferior a f . Portanto, a segunda condição necessária para que x^* seja ótimo para (3.11) é que a matriz $Z^T \nabla^2 f(x^*) Z$, chamada de matriz Hessiana projetada de f em x^* , pode ser semidefinida positiva.

Condições necessárias para um mínimo do problema (3.11):

1. $Ax^* = b$;
2. $Z^T \nabla f(x^*) = 0$; ou, equivalentemente, $\nabla f(x^*) = A^T \lambda^*$;
3. $Z^T \nabla^2 f(x^*) Z$ é semidefinida positiva.

Condições suficientes para um mínimo do problema (3.11):

1. $Ax^* = b$;
2. $Z^T \nabla f(x^*) = 0$; ou, equivalentemente, $\nabla f(x^*) = A^T \lambda^*$;
3. $Z^T \nabla^2 f(x^*) Z$ é definida positiva.

3.2.6 Convergência de Seqüências e Rapidez de Convergência

A maioria dos métodos de otimização são iterativos, tal que uma seqüência infinita x_k de estimativas do ponto ótimo x^* é gerada. Mesmo provando, teoricamente, que esta seqüência converge para o limite do ponto requerido, o método será praticável somente se essa convergência ocorrer com certa rapidez. Nesta seção será discutido brevemente o que significa caracterizar a taxa, na qual, tal seqüência converge. Aqui, serão introduzidos os mais importantes e relevantes conceitos da teoria de convergência. Em todos os casos seguintes, assume-se que: a seqüência x_k converge para x^* , todos os valores x_k são distintos e $x_k \neq x^*$ para qualquer k .

Definição 25 *Dado um método iterativo que produza uma seqüência de pontos x_1, x_2, \dots a partir de um ponto inicial x_0 , quer-se saber se o método converge para uma solução x^* , e caso afirmativo, com que rapidez. Sejam $x^* \in \mathbb{R}^n$, $x_k \in \mathbb{R}^n$, $k = 1, 2, 3, \dots$. Então a seqüência $\{x_k\}$ converge com ordem τ quando τ é o maior número, tal que*

$$0 \leq \lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^\tau} < \infty.$$

Como o interesse está no valor de τ que ocorre no limite, esta constante τ é, algumas vezes, conhecida como taxa de convergência assintótica. Se $\tau = 1$, a seqüência x_k indica convergência linear, se $\tau = 2$, a seqüência é dita ter convergência quadrática.

Se a seqüência x_k tem ordem de convergência τ , a constante assintótica é o valor γ que satisfaz

$$\gamma = \lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^\tau}. \quad (3.15)$$

Quando $\tau = 1$, γ deve ser estritamente menor do que uma unidade para que a convergência ocorra.

Se a seqüência tem convergência linear, o comprimento de passo diminuído em $\|x_k - x^\|$ irá variar substancialmente com o valor da constante de erro assintótica. Assim, se γ em (3.15) é zero quando τ é feito igual a unidade, a este tipo associado de convergência é dado o nome especial de convergência superlinear.*

A definição (25), diz que, se existir uma constante $c \in [0, 1)$ e um inteiro $k \geq 0$ tal que para todo $k \geq \hat{k}$,

$$\|x_{k+1} - x^*\| \leq c \|x_k - x^*\|$$

então $\{x_k\}$ é linearmente convergente. No caso de existirem constantes $c \geq 0$, e $k \geq 0$ tais que $\{x_k\}$ converge para x^* e para todo $k \geq \hat{k}$,

$$\|x_{k+1} - x^*\| \leq c\|x_k - x^*\|^2,$$

então a convergência é quadrática. E ainda, para alguma seqüência $\{c_k\}$ que converge para o valor 0,

$$\|x_{k+1} - x^*\| \leq c_k\|x_k - x^*\|,$$

então $\{x_k\}$ converge superlinearmente para x^* .

Um método iterativo que convirjá para a resposta correta em uma certa taxa, contanto que seja inicializado perto o suficiente da resposta correta, é dito ser localmente convergente naquela taxa.

3.3 Funções Quadráticas

Seja $F : \mathbb{R}^n \rightarrow \mathbb{R}$ a função quadrática

$$F(x) = \frac{1}{2}x^T Ax - h^T x + c, \quad (3.16)$$

onde $A \in \mathbb{R}^{n \times n}$ é uma matriz simétrica, $h \in \mathbb{R}^n$ e $c \in \mathbb{R}$.

O gradiente de F em x é o vetor

$$\nabla F(x) = Ax - h. \quad (3.17)$$

Um ponto crítico de F é um ponto x tal que $\nabla F(x) = 0$. Além disso, x é um ponto crítico da função quadrática F se e somente se x é uma solução do sistema de equações lineares

$$Ax = h. \quad (3.18)$$

O sistema (3.18) pode, ou não, ter uma solução. No entanto, se A é não-singular existe exatamente uma solução, i.e.,

$$x_0 = A^{-1}h$$

e x_0 é o único ponto crítico de F . Se A é singular e x_0 é uma solução de (3.18), então toda solução de x de (3.18) é expressa na forma $x = x_0 + z$, onde $z \in \mathcal{N}(A)$, isto é, z um vetor tal que $Az = 0$. Em outras palavras, se x_0 é um ponto crítico de F , então todo ponto crítico de F difere de x_0 por um vetor $z \in \mathcal{N}(A)$.

Como F é uma função quadrática, tem-se a identidade

$$F(x + p) = F(x) + p^T(Ax - h) + \frac{1}{2}p^T Ap \quad (3.19)$$

Esta identidade é obtida substituindo x por $x + p$ em (3.16). Observe que (3.19) expressa o *Teorema de Taylor*, pg. 29, no ponto x . Reescrevendo (3.19) tem-se,

$$F(x + p) = F(x) + p^T \nabla F(x) + \frac{1}{2}p^T \nabla^2 F(x)p. \quad (3.20)$$

A matriz A da função quadrática é a *Hessiana* de F , isto é, $A = \nabla^2 F(x)$. Se x_0 é um ponto crítico de F , i.e., $\nabla F(x_0) = 0$, então substituindo x por x_0 e p por $x - x_0$ em (3.20), obtém-se a equação

$$F(x) = F(x_0) + \frac{1}{2}(x - x_0)^T A(x - x_0) \quad (3.21)$$

para F relativa a um ponto crítico de F . Geometricamente, essa equação nos diz que, quando A é não-singular, o ponto crítico $x_0 = A^{-1}h$ de F é o centro da superfície quadrática

$$F(x) = \gamma, \quad (3.22)$$

onde γ é uma constante.

Por (3.21), um ponto crítico x_0 de F é um ponto de mínimo de F se e somente se A é semidefinida positiva, ou seja, se e somente se a desigualdade $p^T Ap \geq 0$ se mantém para todo vetor p . Se $p^T Ap > 0$ sempre que $p \neq 0$, isto é, se A é definida positiva, então $x_0 = A^{-1}h$ é o único ponto de mínimo de F . Diz-se que F é definida positiva se A é definida positiva. As curvas de nível de (3.22) para uma função quadrática definida positiva são elipsóides com centro em $x_0 = A^{-1}h$. O problema de minimizar F é, portanto, equivalente ao problema de encontrar o centro geometricamente de um elipsóide. Tal fato conduz ao estudo da descrição geométrica do método das direções conjugadas que será discutido a partir daqui.

No estudo de um sistema linear $Ax = b$, define-se o *resíduo* do sistema como o vetor

$$r = h - Ax.$$

Quando A é simétrica, o vetor r é o gradiente negativo, ou seja,

$$r = -\nabla F(x) = h - Ax.$$

Segue que o resíduo $r(x)$ é a *máxima direção de descida* da função quadrática (3.16) no ponto x .

A preocupação é não somente minimizar F globalmente, mas também minimizar F sujeita a um conjunto de restrições lineares

$$q_i^T x = \rho_i \quad i = 1, \dots, n - k, \quad 1 \leq k < n, \quad (3.23)$$

onde q_1, \dots, q_{n-k} são vetores linearmente independentes. O conjunto de pontos satisfazendo (3.23) é denominado *Hiperplano*, denotado por H_k .

O teorema a seguir apresenta as condições de existência de pontos mínimos de uma função quadrática em hiperplanos.

Teorema 20 *Suponha que a hessiana A da função quadrática F (3.16) é definida positiva. Um ponto x^* em um hiperplano H_k minimiza F em H_k se e somente se o gradiente $\nabla F(x^*)$ é ortogonal a H_k . Existe um único ponto mínimo x^* de F em H_k . Um hiperplano H_k cortado por um elipsóide E_{n-1} de dimensão $(n - 1)$*

$$E_{n-1} : F(x) = \gamma$$

corta E_{n-1} em uma elipsóide de dimensão $(k - 1)$ cujo centro é o ponto mínimo x^ de F em H_k .*

Demonstração. Ver [16]. ■

3.3.1 Propriedades Básicas de Funções Quadráticas

Sejam $F : \mathbb{R}^n \rightarrow \mathbb{R}$ a função quadrática definida em (3.16), com A definida positiva, e $x^* = \arg \min F(x)$,

$$F(x) = F(x^*) + \nabla F(x^*)(x - x^*) + \frac{1}{2}(x - x^*)^T A(x - x^*)$$

então

$$F(x) = F(x^*) + \frac{1}{2}(x - x^*)^T A(x - x^*).$$

Defina o conjunto dos pontos

$$\begin{aligned} \{x \in \mathbb{R}^n : F(x) = \alpha\} &= \{x \in \mathbb{R}^n : \frac{1}{2}(x - x^*)^T A(x - x^*) = \alpha - F(x^*)\} \\ &= \{x \in \mathbb{R}^n : (x - x^*)^T A(x - x^*) = \gamma\}, \end{aligned}$$

onde

$$\gamma = 2(\alpha - F(x^*))$$

como

$$F_\gamma = (x - x^*)^T A(x - x^*) = \gamma.$$

Assim, $F_\gamma \neq \emptyset$ se $\alpha > F(x^*)$. F_γ é um hiperbolóide dado pela equação

$$(x - x^*)^T A(x - x^*) = \gamma,$$

logo F_γ tem dimensão $n - 1$. Então, $F_\gamma \subset \mathbb{R}^n$ e $\dim(F_\gamma) = n - 1$. Seja um hiperplano

$$H = \{x \in \mathbb{R}^n : a^T x = b\}$$

Isto é verificado do seguinte modo.

Para todo $x \in \mathbb{R}^n$, $x \in H$, se e somente se $x = x^* + z$ tal que $a^T x^* = b$, e $a^T z = 0$. Isso implica que $x \in H$ se e só se $x = x^* + z$, $x^* \in H$ e $z \in \aleph(a)$ o que implicando que $x \in H$ pode ser escrito como $x = x^* + \aleph(a)$.

Como $\text{Im}(a) = \{y \in \mathbb{R} : y = a^T x, x \in \mathbb{R}^n\}$ então, $\dim(\text{Im}(a)) = 1$. Logo, $\dim(\aleph(a)) = n - 1$. Então $\dim(H) = n - 1$.

Chamando $k = n - 1$ logo, se $F_\gamma \subset H$ então $\dim(F_\gamma \cap H) = k$ e se $F_\gamma \not\subset H$ então $\implies \dim(F_\gamma \cap H) = k - 1$.

Teorema 21 *Os pontos de mínimo de F sobre retas paralelas estão em um hiperplano H (variedade linear) que passa por $x^* = \arg \min(F(x))$. O hiperplano H é definido pela equação $p^T(Ax - h) = 0$ (ou $p^T \nabla F(x) = 0$), onde p é a direção das retas paralelas. O vetor Ap é normal ao hiperplano H .*

Demonstração. Seja $x_1 \in \mathbb{R}^n$, e seja a reta $x = x_1 + \alpha p$,

$$\begin{array}{ll} \text{minimizar} & F(x) \\ \text{s.a.} & x = x_1 + \alpha p \end{array}$$

implica em

$$\begin{array}{ll} \text{minimizar} & F(x) \\ \text{s.a.} & \alpha \in \mathbb{R} \end{array}.$$

Pela condição de otimalidade (pg. (38))

$$\alpha_1 = \arg \min(F(x_1 + \alpha p)) \iff p^T \nabla F(x_1 + \alpha_1 p) = 0.$$

Chamando $x_1^* = x_1 + \alpha_1 p$, então x_1^* é o ponto de mínimo de $F(x)$ se e somente se

$$p^T \nabla F(x_1^*) = 0.$$

Toma-se agora um novo ponto $x_2 \in \mathbb{R}^n$, $x_2 \neq x_1$, e usa-se o mesmo argumento. Então,

$$p^T \nabla F(x_2^*) = 0.$$

Observe que $\nabla F(x) = Ax - h$, então,

$$p^T (Ax_1^* - h) = 0 \implies p^T Ax_1^* = p^T h$$

e,

$$p^T (Ax_2^* - h) = 0 \implies p^T Ax_2^* = p^T h.$$

Logo, x_1^* e x_2^* pertencem ao hiperplano H definido por

$$H = \{x_1 \in \mathbb{R}^n : (Ap)^T x = p^T h\}.$$

Em particular,

$$\nabla F(x^*)^T p = 0 \implies x^* \in H.$$

Além disso, como x_1^* e $x_2^* \in H$, então $x_1^* - x_2^* \in H$, logo $(Ap)^T (x_1^* - x_2^*) = 0$, donde Ap é um vetor normal a H . ■

Segue o corolário.

Corolário 22 Dado um vetor não nulo $p \in \mathbb{R}^n$, sejam $x_1^*, x_2^* \in \mathbb{R}^n$ os pontos de mínimos de F sobre as retas L_1 e L_2 , com $L_1 \neq L_2$ que têm a mesma direção p . O vetor $q = x_1^* - x_2^*$ é conjugado do vetor p .

O seguinte teorema mostra que se certo ponto x_{k+1} é o ponto de mínimo de F sobre uma reta, dada por $x_k + \alpha p$, ou seja, uma reta na qual seja fixado um ponto ao "caminhar-se" a partir desse ponto em direção a p até certo comprimento, então dado esse passo em direção a p com um tamanho apropriado faz com que x_{k+1} possa se a escrito como $x_k + \alpha p$. O resultado do seguinte teorema, determina qual é o tamanho do passo dado na direção dessa reta para que a afirmação sobre x_{k+1} seja verdadeira.

Teorema 23 Seja x_2 o ponto de mínimo de F sobre a reta $x = x_1 + \alpha p$. Então, x_2 é dado por:

$$x_2 = x_1 + \alpha p,$$

com

$$\alpha = -\frac{p^T \nabla F(x_1)}{p^T A p}.$$

Demonstração.

$$\begin{aligned} x_2 = \arg \min_{x=x_1+\alpha p} (F(x)) &\iff x_2 = \arg \min_{\alpha} (F(x_1 + \alpha p)) \\ \iff p^T \nabla F(x_1 + \alpha p) = 0 &\iff p^T (A(x_1 + \alpha p) - h) = 0 \\ \iff p^T (Ax_1 - h + \alpha Ap) = 0 &\iff p^T \nabla F(x_1) + \alpha p^T A p = 0 \\ &\iff \alpha = -\frac{p^T \nabla F(x_1)}{p^T A p} \end{aligned}$$

■

Teorema 24 Seja $q \in \mathbb{R}^n$, $q \neq 0$. Para todo número real ρ o ponto de mínimo de F sobre o hiperplano H definido por $q^T x = \rho$ está sobre a reta

$$x = x_0 + \alpha A^{-1} q$$

onde x_0 é o ponto de mínimo de F na direção $p = A^{-1} q$.

Demonstração. Pelo teorema (21), pg. 43, se x^* é ponto mínimo de F sobre o hiperplano H definido por $q^T x = \rho$, então $\nabla F(x^*)$ deve ser ortogonal a H , logo $q \perp H$ e $\nabla F(x^*) \perp H$, e portanto, q é paralelo a $\nabla F(x^*)$, donde

$$\alpha q = \nabla F(x^*).$$

Logo, $\alpha q = Ax^* - h$, donde

$$Ax^* = h + \alpha q \implies x^* = A^{-1}h + \alpha A^{-1}q.$$

Se x_0 é o ponto de mínimo de F , então

$$\nabla F(x_0) = 0 \implies Ax_0 = h \implies x_0 = A^{-1}h,$$

logo

$$x^* = x_0 + \alpha A^{-1}q$$

e portanto, $x^* \in L = x_0 + \alpha A^{-1}q = x_0 + \alpha p$, onde $p = A^{-1}q$. Ademais, como $Ap = q \perp H$, então $Ap \perp H$ isso implica que p é conjugado de todo vetor de H . Assim, p é conjugado de H . ■

Teorema 25 *Os pontos de mínimo de F sobre hiperplanos de dimensão k estão em um hiperplano de dimensão $n - k$ que contém o ponto de mínimo de F . Em outras palavras, dado um conjunto de $n - k$ vetores linearmente independentes q_1, q_2, \dots, q_{n-k} , então para todo conjunto de números reais $\rho_1, \rho_2, \dots, \rho_{n-k}$, x^* é o ponto de mínimo de F sobre o hiperplano definido por:*

$$\begin{cases} q_1^T x = \rho_1 \\ q_2^T x = \rho_2 \\ \vdots \\ q_{n-k}^T x = \rho_{n-k} \end{cases} \quad (3.24)$$

está sobre o hiperplano H definido por

$$x = x^* + \alpha_1 A^{-1}q_1 + \dots + \alpha_{n-k} A^{-1}q_{n-k}.$$

Demonstração. Seja $Q = \begin{pmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_{n-k}^T \end{pmatrix}$, $\rho^T = (\rho_1, \rho_2, \dots, \rho_{n-k})^T$, $\dim(\text{Im}(Q)) = n - k$,

logo $\dim(\mathfrak{N}(Q)) = k$. O sistema (3.24) pode ser escrito como

$$Qx = \rho,$$

com a imagem e o espaço nulo de Q sendo, respectivamente, $\text{Im}(Q) = \{y : y = Qx, x \in \mathbb{R}^n\}$ e $\mathfrak{N}(Q) = \{u : Qu = 0\}$.

Sejam u_1, u_2, \dots, u_k um conjunto de vetores linearmente independentes tais que

$$q_i^T u_j = 0, \quad \forall i = 1, 2, \dots, n-k \quad e \quad \forall j = 1, 2, \dots, k$$

então, $u_j \in \mathfrak{N}(Q)$, $j = 1, 2, \dots, k$ e como $\dim(\mathfrak{N}(Q)) = k$, logo $\{u_1, u_2, \dots, u_k\}$ é uma base para $\mathfrak{N}(Q)$.

Seja $\bar{x} \in \mathbb{R}^n$ tal que $Q\bar{x} = \rho$, então $x \in H_F$ definido por (3.24) se e somente se $x = \bar{x} + u$, com $u \in \mathfrak{N}(Q)$ o que é equivalente a $x = \bar{x} + \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_k u_k$. H_F definido por (3.24) tem dimensão k , $\dim(H_F) = k$.

Agora, recorre-se ao problema de

$$\begin{aligned} & \text{minimizar} && F(x) \\ & \text{s.a.} && Qx = \rho \end{aligned} \quad (3.25)$$

Como x é solução de $Qx = \rho$, x pode ser escrito como $x = \bar{x} + U\beta$, onde $U = (u_1 \ u_2 \ \dots \ u_k)$. Desse modo, o problema (3.25) é equivalente a

$$\text{minimizar}_{\beta \in \mathbb{R}^k} F(\bar{x} + U\beta). \quad (3.26)$$

Assim,

$$\begin{aligned} \beta^* \text{ é solução de (3.26)} & \iff \nabla F(\bar{x} + U\beta^*)^T U = 0 \iff \\ & \iff \nabla F(\bar{x} + U\beta^*) \perp u_1, u_2, \dots, u_k \iff \\ & \iff \nabla F(\bar{x} + U\beta^*) \in [q_1, q_2, \dots, q_{n-k}] \iff \\ & \iff \exists \alpha_1, \alpha_2, \dots, \alpha_{n-k} \end{aligned}$$

tais que

$$\begin{aligned} \nabla F(\bar{x} + U\beta^*) & = \alpha_1 q_1 + \alpha_2 q_2 + \dots + \alpha_{n-k} q_{n-k} \iff \\ \iff A(\bar{x} + U\beta^*) - h & = \alpha_1 q_1 + \alpha_2 q_2 + \dots + \alpha_{n-k} q_{n-k} \iff \\ \iff (\bar{x} + U\beta^*) & = \\ & = A^{-1}h + \alpha_1 A^{-1}q_1 + \alpha_2 A^{-1}q_2 + \dots + \alpha_{n-k} A^{-1}q_{n-k} \end{aligned}$$

Logo,

$$(\bar{x} + U\beta^*) = A^{-1}h + \alpha_1 A^{-1}q_1 + \alpha_2 A^{-1}q_2 + \cdots + \alpha_{n-k} A^{-1}q_{n-k}$$

onde \bar{x}^* é o ponto de mínimo de F sobre H_F . ■

3.3.2 Minimização de Quadráticas Sobre Hiperplanos

Seja a quadrática definida por (3.16), pg. 40, onde $A \in \mathbb{R}^{n \times n}$ é uma matriz simétrica e definida positiva, e seja p_1, p_2, \dots, p_k um conjunto de vetores não nulos, conjugados, i.e.,

$$p_i^T A p_j = 0, \quad \forall i \neq j.$$

Teorema 26 *Seja H o hiperplano definido por $x \in H$, $x = x_1 + \alpha_1 p_1 + \cdots + \alpha_k p_k$, $\dim(H) = k$. O ponto de mínimo de F sobre H , x_{k+1} , é dado por*

$$x_{k+1} = x_1 + \bar{\alpha}_1 p_1 + \cdots + \bar{\alpha}_k p_k,$$

onde

$$\bar{\alpha}_i = \frac{p_i^T r_1}{p_i^T A p_i},$$

e

$$r_1 = -\nabla F(x_1) = h - A x_1.$$

$$r_{k+1} = r_1 - \bar{\alpha}_1 A p_1 - \cdots - \bar{\alpha}_k A p_k,$$

$$r_{k+1}^T p_i = 0, \quad \forall i = 1, 2, \dots, k.$$

$$F(x_{k+1}) = f(x_1) - \frac{1}{2} \left[\frac{(p_1^T r_1)^2}{p_1^T A p_1} + \cdots + \frac{(p_k^T r_1)^2}{p_k^T A p_k} \right].$$

Demonstração. Seja $P = (p_1 \ p_2 \ \dots \ p_k)$ uma matriz cujas colunas são os vetores p_1, p_2, \dots, p_k . Então P tem posto coluna igual a k , e portanto, $\text{posto}(P) = k$,

$$H = x_1 + P\alpha, \quad \alpha \in \mathbb{R}^k.$$

$$\underset{x \in H}{\text{minimizar}} \quad F(x) \quad \iff \quad \underset{\alpha \in \mathbb{R}^k}{\text{minimizar}} \quad F(x_1 + P\alpha)$$

onde o problema de otimização do lado direito é um problema de otimização irrestrita. $\bar{\alpha}$ minimiza $F(x_1 + P\alpha) \iff \nabla F(x_1 + P\bar{\alpha})^T = 0$. Então,

$$(i) \quad x_{k+1} = x_1 + \bar{\alpha}_1 p_1 + \cdots + \bar{\alpha}_k p_k;$$

$$(ii) \quad \nabla F(x_{k+1}) \perp p_1, p_2, \dots, p_k \implies \nabla F(x_{k+1}) \perp H$$

$$(iii) \quad \nabla F(x_{k+1}) = Ax_{k+1} - h = A(x_1 + P\bar{\alpha}) - h = Ax_1 - h + AP\bar{\alpha} = -r_1 + AP\bar{\alpha}$$

Então, $\forall i = 1, 2, \dots, k$

$$0 = p_i^T \nabla F(x_{k+1}) = -p_i^T r_1 + p_i^T AP\bar{\alpha}.$$

Donde, então,

$$0 = -p_i^T r_1 + \bar{\alpha}_i p_i^T Ap_i \implies \bar{\alpha}_i = \frac{p_i^T r_1}{p_i^T Ap_i}$$

$$\begin{aligned} (iv) \quad r_{k+1} &= -Ax_{k+1} + h = h - Ax_{k+1} \\ &= h - A(x_1 + \bar{\alpha}_1 p_1 + \cdots + \bar{\alpha}_k p_k) \\ &= h - Ax_1 + \bar{\alpha}_1 Ap_1 - \cdots - \bar{\alpha}_k Ap_k = r_1 - \bar{\alpha}_1 Ap_1 - \cdots - \bar{\alpha}_k Ap_k \end{aligned}$$

$$\begin{aligned} (v) \quad F(x_1) &= F(x_{k+1}) + \nabla F(x_{k+1})^T (x_1 - x_{k+1}) + \frac{1}{2} (x_1 - x_{k+1})^T A (x_1 - x_{k+1}) \\ &= F(x_{k+1}) + \nabla F(x_{k+1})^T (-P\bar{\alpha}) + \frac{1}{2} (-P\bar{\alpha})^T A (-P\bar{\alpha}) \\ &= F(x_{k+1}) + \frac{1}{2} \sum_{i=1}^k \bar{\alpha}_i^2 p_i^T Ap_i. \end{aligned}$$

Logo,

$$F(x_{k+1}) = F(x_1) - \frac{1}{2} \sum_{i=1}^k \frac{(p_i^T r_1)^2}{p_i^T Ap_i}.$$

■

Teorema 27 *Sejam p_1, p_2, \dots, p_m vetores não nulos, conjugados. Dado $x_1 \in \mathbb{R}^n$, sejam x_2, x_3, \dots, x_{m+1} pontos definidos recursivamente com a condição de que para cada k , $1 \leq k \leq m$, x_{k+1}*

$$\underset{x=x_k+\alpha p_k}{\text{minimiza}} F(x).$$

Então,

$$(i) \ x_{k+1} = x_k + \alpha_k p_k; \ r_{k+1} = r_k - \alpha_k A p_k, \text{ onde } \alpha_k = \frac{p_k^T r_k}{p_k^T A p_k}, \ r_k = -\nabla F(x_k).$$

$$(ii) \ x_{k+1}$$

$$\underset{x \in H}{\text{minimiza}} F(x),$$

onde $H = \{x_1 + \alpha_1 p_1 + \dots + \alpha_k p_k\}$.

$$(iii) \ p_k^T r_j = c_k, \ j \leq k \leq m$$

$$p_k^T r_j = 0, \ k < j \leq m + 1.$$

Demonstração. Para todo k tal que, $1 \leq k \leq m$, x_{k+1} é um vetor, o qual,

$$\underset{x=x_k+\alpha p_k}{\text{minimiza}} F(x)$$

se e somente se

$$\bar{\alpha} \text{ minimiza } F(x_k + \alpha p_k) \implies \nabla F(x_{k+1})^T p_k = 0.$$

Usando o fato de que

$$\nabla F(x_{k+1}) = A x_{k+1} - h$$

obtem-se que

$$\alpha_k = \frac{p_k^T r_k}{p_k^T A p_k}.$$

Pelo teorema (26), tem-se que

$$x_{k+1} = x_k + \alpha_k p_k, \ r_{k+1} = r_k - \alpha_k A p_k.$$

Então,

$$p_k^T r_{k+1} = p_k^T r_k - \alpha_k p_k^T A p_k = p_k^T r_k - \frac{p_k^T r_k}{p_k^T A p_k} p_k^T A p_k = 0.$$

$$\forall j = 1, 2, \dots, k-1, j \neq k$$

$$\begin{aligned} p_k^T r_{j+1} &= p_k^T (r_j - \alpha_j A p_j) \\ &= p_k^T r_j - \alpha_j p_k^T A p_j = p_k^T r_j, \end{aligned}$$

pois $p_k^T A p_j = 0, \forall j \neq k$. Então,

$$p_k^T r_1 = p_k^T r_2 = \dots = p_k^T r_{k-1} = p_k^T r_k.$$

Para $j = k, k+1, \dots, m$

$$p_k^T r_{j+1} = p_k^T (r_j - \alpha_j A p_j) = p_k^T r_j - \frac{p_k^T r_j}{p_k^T A p_j} p_k^T A p_j$$

$$j = k$$

$$p_k^T r_{k+1} = 0,$$

$$j > k$$

$$p_k^T r_{k+1} = p_k^T r_{k+2} = \dots = p_k^T r_m = 0.$$

Como $1 \leq k \leq m, p_k^T r_{k+1} = 0$. Então

$$\nabla F(x_{k+1}) \perp H \implies x_{k+1} \underset{x \in H}{\text{minimiza}} F(x).$$

■

Teorema 28 *Sejam p_1, p_2, \dots, p_m vetores linearmente independentes. Dado $x_1 \in \mathbb{R}^n$, sejam x_2, x_3, \dots, x_{m+1} pontos definidos por*

$$x_{k+1} = x_k + \alpha_k p_k = x_1 + \alpha_1 p_1 + \dots + \alpha_k p_k$$

onde $\alpha_i \neq 0, \forall i$. Se para $k = 1, 2, \dots, m, x_{k+1}$

$$\underset{x \in H_k}{\text{minimiza}} F(x),$$

onde $H_k = [p_1, p_2, \dots, p_k]$ então, os vetores p_1, p_2, \dots, p_m são conjugados.

Demonstração. Se x_{k+1} é tal que

$$\underset{x \in H_k}{\text{minimiza}} \quad F(x),$$

então $\nabla F(x_{k+1}) \perp p_1, p_2, \dots, p_k$, $k = 1, 2, \dots, m$.

Logo, $\forall j, j < k \leq m$

$$0 = p_j^T r_{k+1} = p_j^T r_k - \alpha_k p_j^T A p_k$$

o que implica que

$$\alpha_k p_j^T A p_k = 0.$$

Como $\alpha_k \neq 0$, então $p_j^T A p_k = 0$. Logo, $\forall j, j < k \leq m$, $p_j^T A p_k = 0$ o que implica que o conjunto $\{p_1, p_2, \dots, p_m\}$ é conjugado. ■

3.3.3 Método de Direções Conjugadas

Algoritmo 3 *Direções Conjugadas*

1. Dado $x_1 \in \mathbb{R}^n$, calcule $r_1 = h - Ax_1$. Selecione uma direção $p_1 \neq 0$.
2. Para $k = 1, 2, \dots, n$

$$x_{k+1} = x_k + \alpha_k p_k$$

$$\alpha_k = \frac{p_k^T r_k}{p_k^T A p_k}, \quad r_{k+1} = r_k - \alpha_k A p_k$$

$$\text{Selecione } p_{k+1} \text{ conjugado com } p_1, p_2, \dots, p_k \quad (p_k^T r_{k+1} = p_k^T r_k - \alpha_k p_k^T A p_k)$$

Teorema 29 *As direções p_1, p_2, \dots, p_k geradas pelo algoritmo (3) acima são mutuamente conjugadas e*

$$r_k = -\nabla F(x_k) = h - Ax_k \text{ é ortogonal a } p_1, p_2, \dots, p_{k-1}$$

e $p_k^T r_j = c$, onde c é uma constante e $j = 1, 2, \dots, k$, isto é,

$$(i) \quad p_k^T A p_j = 0, \quad \forall k \neq j$$

$$(ii) \quad p_j^T r_k = 0, \quad j = 1, 2, \dots, k-1$$

$$(iii) \quad p_k^T r_j = p_k^T r_k = \dots = p_k^T r_1$$

Demonstração. Como $r_{k+1} = r_k - \alpha_k A p_k$, então

$$p_k^T r_{k+1} = p_k^T r_k - \alpha_k p_k^T A p_k = 0.$$

Logo,

$$\begin{aligned} -r_{k+1} &= \nabla F(x_{k+1}) \perp p_k \implies x_{k+1} \underset{x=x_k+\alpha p_k}{\text{minimiza}} F(x) \iff \\ &\iff x_{k+1} \underset{x=x_1+\alpha_1 p_1+\dots+\alpha_k p_k}{\text{minimiza}} F(x) \implies \\ &\implies x_{k+1} \underset{x=x_1+\alpha_1 p_1+\dots+\alpha_k p_k}{\text{minimiza}} F(x) \implies \\ &\implies \nabla F(x_{k+1}) = -r_{k+1} \perp p_1, p_2, \dots, p_k. \end{aligned}$$

Observa-se que conclusão semelhante segue partindo-se de

$$\begin{aligned} r_k &= r_{k-1} - \alpha_{k-1} A p_{k-1} \implies p_{k-1}^T r_k = p_{k-1}^T r_{k-1} - \alpha_{k-1} A p_{k-1} \\ &= 0 \implies -r_k = \nabla F(x_k) \perp p_1, p_2, \dots, p_{k-1} \end{aligned}$$

ou seja, $\forall k = 1, 2, \dots, n$ e $\forall j = 1, 2, \dots, k-1$

$$p_j^T r_k = 0.$$

Então, $\forall j = 1, 2, \dots, k-1$

$$\begin{aligned} p_j^T r_{k+1} &= p_j^T r_k - \alpha_k p_j^T A p_k \implies p_j^T A p_k = 0 \implies \\ &\implies \text{as direções são mutuamente conjugadas.} \end{aligned}$$

Reescrevendo a equação do resíduo para $j+1$,

$$r_{j+1} = r_j - \alpha_j A p_j,$$

e fazendo $j = 1, 2, \dots, k-1$ obtém-se

$$p_k^T r_{j+1} = p_k^T r_j - \alpha_j p_k^T A p_j \implies p_k^T r_j = p_k^T r_{j+1}.$$

Fazendo $j = 1, 2, \dots, k-1$, obtém-se

$$p_k^T r_1 = p_k^T r_2 = \dots = p_k^T r_{j+1} = p_k^T r_k.$$

■

Teorema 30 O ponto x_{k+1} minimiza $F(x)$ sobre o hiperplano H definido por

$$x = x_1 + \alpha_1 p_1 + \cdots + \alpha_k p_k.$$

O hiperplano H corta o elipsóide $F(x) = \gamma$ ($\gamma \geq F(x_1)$) de dimensão $k - 1$, cujo centro é x_{k+1} . Os pontos x_1, x_2, \dots, x_{k+1} estão em H . Os pontos $x_{k+2}, x_{k+3}, \dots, x_{m+1} = x_0$ estão no hiperplano \bar{H} de dimensão $n - k$, definido por

$$p_j^T A(x - x_{k+1}) = 0, \quad j = 1, 2, \dots, k$$

conjugado com H .

Demonstração. Como a equação

$$F(x) = \gamma \iff (x - x_{k+1})^T A(x - x_{k+1}) = \gamma$$

introduz um vínculo, restam $n - 1$ graus de liberdade na qual x pode variar livremente, então $\dim(\text{Elipsóide}) = n - 1$. Por outro lado, $\dim(H_k) = k$. Então, $\dim(H_k \cap H_{n-1}) = k - 1$, se $H_k \not\subseteq H_{n-1}$. De fato, se $H_k \not\subseteq H_{n-1}$ então existe algum vetor $w \in H_k$, $w \notin H_{n-1} \implies \alpha w \notin H_{n-1}, \forall \alpha \in \mathbb{R}$. Então, $H_k \cap H_{n-1}$ tem no máximo dimensão $k - 1$.

Como x_{k+1} minimiza $F(x)$ sobre o hiperplano H , então

$$x_{k+1} = x_1 + \alpha_1 p_1 + \cdots + \alpha_k p_k \implies x_j \in H, \quad j = 1, 2, \dots, k + 1.$$

Além disso,

$$\begin{aligned} \forall x, x &= x_{k+1} + \alpha_{k+1} p_{k+1} + \cdots + \alpha_m p_m \implies \\ &\implies x - x_{k+1} \in [\{p_{k+1}, p_{k+2}, \dots, p_m\}]. \end{aligned}$$

Então, $x_{k+2}, x_{k+3}, \dots, x_{m+1} \in [\{p_{k+1}, p_{k+2}, \dots, p_m\}]$. Pela conjugacidade dos vetores p_1, p_2, \dots, p_m

$$p_j^T A(x - x_{k+1}) = 0, \quad j = 1, 2, \dots, k.$$

Em particular, x_{k+1} , cumpre essa condição. ■

3.3.4 Método do Gradiente Conjugado

O método do Gradiente Conjugado é um método de direção conjugada ver teorema 29, com uma propriedade muito especial: na geração do conjunto de vetores conjugados,

pode-se calcular um novo vetor p_k usando somente o vetor anterior p_{k-1} . Não é necessário conhecer todos os elementos anteriores a p_{k-1} , i.e., p_0, p_1, \dots, p_{k-2} do conjunto conjugado; p_k é pela forma de atualização, automaticamente, conjugado dos vetores p_0, p_1, \dots, p_{k-1} . Esta propriedade implica que o método requer menos cálculo e armazenagem.

Agora, os detalhes do método do Gradiente Conjugado (CG). Toda direção p_k é escolhida como uma combinação linear da máxima direção de descida, i.e., $-\nabla\phi(x_k)$ (a qual é dada pelo resíduo negativo $-r_k$) e direção anterior p_{k-1} . Escrevendo

$$p_k = -r_k + \beta_k p_{k-1} \quad (3.27)$$

onde o escalar β_k é determinado para exigir que p_{k-1} e p_k sejam conjugadas com respeito a A . Pré-multiplicando (3.27) por $p_{k-1}^T A$ e impondo a condição $p_{k-1}^T A p_k = 0$, tem-se que

$$\beta_k = \frac{r_k^T A p_{k-1}}{p_{k-1}^T A p_{k-1}}.$$

Escolhe-se a primeira direção p_0 como sendo a máxima direção de descida para o ponto inicial x_0 .

Igualmente ao método de direções conjugadas o método CG, representa minimizações sucessivas unidimensionais ao longo de cada direção. O algoritmo CG é expresso a seguir:

Algoritmo 4 CG

Dado x_0 ;

Faça $r_0 = Ax_0 - b$, $p_0 \leftarrow -r_0$, $k \leftarrow 0$;

Enquanto $r_k \neq 0$

$$\alpha_k \leftarrow -\frac{r_k^T p_k}{p_k^T A p_k};$$

$$x_{k+1} \leftarrow x_k + \alpha_k p_k;$$

$$r_{k+1} \leftarrow Ax_{k+1} - b;$$

$$\beta_{k+1} \leftarrow \frac{r_{k+1}^T A p_k}{p_k^T A p_k};$$

$$p_{k+1} \leftarrow -r_{k+1} + \beta_{k+1} p_k;$$

$$k \leftarrow k + 1;$$

Fim (enquanto)

A forma prática padrão do método do gradiente conjugado é dada pelo seguinte algoritmo.

$$\alpha_k = \frac{r_k^T r_k}{p_k^T A p_k}.$$

Segundo, tem-se que $\alpha_k A p_k = r_{k+1} - r_k$, logo

$$\beta_{k+1} = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}.$$

Algoritmo 5 *CG Prático*

Dado x_0 ;

Faça $r_0 = Ax_0 - b$, $p_0 \leftarrow -r_0$, $k \leftarrow 0$;

Enquanto $r_k \neq 0$

$$\alpha_k \leftarrow \frac{r_k^T r_k}{p_k^T A p_k};$$

$$x_{k+1} \leftarrow x_k + \alpha_k p_k;$$

$$r_{k+1} \leftarrow r_k + \alpha_k A p_k;$$

$$\beta_{k+1} \leftarrow \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k};$$

$$p_{k+1} \leftarrow -r_{k+1} + \beta_{k+1} p_k;$$

$$k \leftarrow k + 1;$$

Fim (enquanto)

O método CG é recomendado somente para problemas grandes; caso contrário, eliminação Gaussiana ou outro algoritmo de fatoração como decomposição de valor singular são preferidos, pois estes são menos sensíveis aos erros de arredondamento na implementação computacional.

Extensão para Problemas Não Quadráticos

O problema geral de minimização de uma função F em \mathbb{R}^n pode ser abordado, fazendo aproximações adequadas, por algoritmos de gradientes conjugados. Existem várias maneiras de atingir esse objetivo. A escolha depende parcialmente de quais propriedades de F são mais facilmente computáveis. Examina-se uma delas, a aproximação quadrática.

Na abordagem da aproximação quadrática são feitas as seguintes notações em x_k :

$$g_k \leftrightarrow \nabla F(x_k)$$

$$A \leftrightarrow \nabla^2 F(x_k)$$

e usando essas associações, reavaliadas a cada passo, todas as quantidades necessárias para implementar o algoritmo de gradientes conjugados básico podem ser obtidas. Se F for quadrática, essas associações são identidades, de maneira que o algoritmo geral obtido pelo uso delas é uma generalização do esquema dos gradientes conjugados. Isto é similar à filosofia implícita pelo método de Newton, onde a cada passo a solução de um problema geral é aproximada pela solução de um problema puramente quadrático através destas mesmas associações.

Quando aplicados a problemas não quadráticos, os métodos de gradientes conjugados usualmente não convergem em n passos. Após atingidos n passos não é mais possível encontrar novas direções conjugadas, já que estamos em \mathbb{R}^n e qualquer novo vetor encontrado será uma combinação linear dos vetores direção gerados anteriormente pelo algoritmo. Já que A – *conjugacidade* dos vetores direcionais no algoritmo puro de gradientes conjugados é dependente da direção inicial, que é o negativo do gradiente, o procedimento de recomeço parece ser o preferido. Sempre se inclui este procedimento de recomeço na implementação do algoritmo. O algoritmo geral de gradientes conjugados é então definido como segue:

Passo 1 - Começando com x_0 compute $g_0 = \nabla F(x_0)$ e faça $p_0 = -g_0$.

Passo 2 - Para $k = 0, 1, \dots, n - 1$:

(1). Faça $x_{k+1} = x_k + \alpha_k p_k$ onde

$$\alpha_k = \frac{-g_k^T p_k}{p_k^T \nabla^2 F(x_k) p_k}$$

(2). Calcule $g_{k+1} = \nabla F(x_k)$. A menos que $k = n - 1$, faça $p_{k+1} = -g_{k+1} + \beta_k p_k$ onde

$$\beta_k = \frac{g_{k+1}^T \nabla^2 F(x_k) p_k}{p_k^T \nabla^2 F(x_k) p_k}.$$

e repita (1).

Passo 3 - Substitua x_0 por x_k e volte para o Passo 1.

Um aspecto atrativo do algoritmo é que, assim como na forma pura do método de Newton, nenhuma busca linear é requerida em qualquer estágio. Além disso, o algoritmo converge num número finito de passos para um problema quadrático. Os aspectos indesejáveis são que $\nabla^2 F(x_k)$ deve ser avaliada a cada ponto, o que é freqüentemente pouco prático, e o algoritmo não é, nesta forma, globalmente convergente.

Método de Gradientes Conjugados Linear

O método de gradientes conjugados descrito anteriormente pode também ser visto como um método para resolver um conjunto de equações lineares [17], onde a matriz do sistema é simétrica e definida positiva. De fato, o método foi originalmente desenvolvido para esse propósito. Se o algoritmo de gradientes conjugados é aplicado para minimizar a função quadrática $\frac{1}{2}x^T Ax + b^T x$, onde A é simétrica e definida positiva, ele computa a solução do sistema

$$Ax = -b \quad (3.28)$$

que vem a ser a condição necessária e suficiente para que um ponto seja a solução ótima do problema de minimizar a quadrática. Quando usado para resolver um sistema linear, o algoritmo é conhecido como método de gradientes conjugados linear. A característica marcante do método de gradiente conjugado é que ele calcula a solução de um sistema linear usando somente produtos de matriz por vetor, e não requer os elementos da matriz explicitamente. Por exemplo, se A fosse dada por $R^T R$, não seria necessário formar este produto para aplicar o método de gradientes conjugados.

Para iniciar as iterações, adota-se a convenção que $\beta_{-1} = 0$, $p_{-1} = 0$. Dados x_0 e $g_0 = Ax_0 + b$, cada iteração inclui os próximos passos para $k = 0, 1, \dots, n - 1$:

$$\begin{aligned} p_k &= -g_k + \beta_{k-1} p_{k-1} \\ \alpha_k &= \frac{\|g_k\|_2^2}{p_k^T A p_k} \\ x_{k+1} &= x_k + \alpha_k p_k \\ g_{k+1} &= g_k + \alpha_k A p_k \\ \beta_k &= \frac{\|g_{k+1}\|_2^2}{\|g_k\|_2^2} \end{aligned}$$

Em teoria, o algoritmo de gradientes conjugados linear calculará a solução exata de (3.28) em um número finito de iterações. Em particular, o método tem a propriedade de, se for usada a aritmética exata, atingir a convergência em $m \leq n$ iterações, onde m é o número de autovalores distintos de A . Logo, o método pode convergir muito rapidamente se os autovalores de A forem agrupados com valores aproximadamente iguais. Entretanto, na prática, erros de arredondamento no cálculo das direções rapidamente provocam a perda da conjugacidade, e o método se comporta como um método iterativo. Daí a importância do processo de reinicialização quando da implementação do método.

4

Uma Nova Abordagem para a Minimização de uma Função Não Linear Sujeita a Restrições Lineares de Igualdade

Seja $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função duas vezes continuamente diferenciável, A uma matriz $m \times n$, com $m < n$, de posto m , e b um vetor de \mathbb{R}^m . Considere o seguinte problema

$$\begin{array}{ll} \text{Minimizar} & f(x) \\ \text{s.a.} & Ax = b \end{array} \quad (4.1)$$

Se x_k é um ponto viável das restrições do problema (4.1), então $Ax_k = b$. Logo, todo algoritmo do tipo $x_{k+1} = x_k + \alpha_k d_k$ com x_k viável, obriga que a direção de busca esteja no espaço nulo de A . De fato,

$$Ax_{k+1} = Ax_k + \alpha_k Ad_k,$$

e como x_{k+1} deve ser viável tem-se que para $\alpha_k \neq 0$,

$$Ad_k = 0,$$

donde então d_k deve estar no espaço nulo de A .

Agora, considere o problema de escolher a melhor direção sujeita à condição de que ela esteja no espaço nulo de A , ou seja,

$$\begin{array}{ll} \text{minimizar} & f(x_k + d) \\ \text{s.a.} & Ad = 0 \end{array} \quad (4.2)$$

ou seja, $d \in \mathfrak{N}(A)$, $d \neq 0$.

Ora, se $d \in \mathfrak{N}(A)$ então $d = Zp$, onde as colunas de $[Z]$ formam uma base para o espaço nulo de A . Então, o problema (4.2) pode ser colocado como o problema irrestrito

$$\underset{p \in \mathbb{R}^{n-m}}{\text{minimizar}} \quad f(x_k + Zp) . \quad (4.3)$$

A direção de máxima descida é obtida fazendo-se uma aproximação linear de f , e então buscando-se a solução de mínimo dessa aproximação

$$\underset{\|Zp\|=1}{\text{minimizar}} \quad f(x_k) + \nabla f(x_k)^T Zp , \quad (4.4)$$

cuja solução é,

$$Zp = -\frac{1}{\|\nabla f(x_k)\|} . \quad (4.5)$$

Observe que o sistema linear (4.4), em geral, não tem solução, dado que $\nabla f(x_k) \in \mathbb{R}^n$, e as colunas de Z geram apenas um subespaço próprio de \mathbb{R}^n . Para contornar o feito de que (4.5) não tem em geral solução, podemos sempre buscar a solução de quadrados mínimos que é dada pela solução do sistema de equações normais

$$\min_d \frac{1}{2} \|Zp + \nabla f(x_k)\|_2^2$$

cuja solução é dado pela solução do sistema de equações normais

$$Z^T Z = -Z^T \nabla f(x_k),$$

onde $Z^T Z$ é uma matriz simétrica, e Zp é o vetor unitário na direção de $\nabla f(x_k)$. Como Z tem posto coluna completo, $Z^T Z$ é uma matriz definida positiva, e portanto o sistema tem solução única, e p obtido desse modo é uma direção de descida para o problema (4.3).

No caso da obtenção da direção de Newton no $\mathfrak{N}(A)$, tem-se que resolver o seguinte problema

$$\begin{aligned} \underset{s.a.}{\text{minimizar}} \quad & f(x_k) + \nabla f(x_k)^T d + \frac{1}{2} d^T \nabla^2 f(x_k) d \\ & Ad = 0. \end{aligned}$$

Utilizando-se do fato de que $d = Zp$, o problema de minimização será,

$$\underset{p \in \mathbb{R}^{n-m}}{\text{minimizar}} \quad f(x_k) + \nabla f(x_k)^T Zp + \frac{1}{2} p^T Z^T \nabla^2 f(x_k) Zp,$$

cuja solução é a solução do sistema linear

$$Z^T \nabla^2 f(x_k) Z p = -Z^T \nabla f(x_k) \quad (4.6)$$

onde $Z^T \nabla^2 f(x_k) Z$ é simétrica, mas não necessariamente definida positiva. Contudo, as exigências sobre $\nabla^2 f(x_k)$, são apenas de que a hessiana de f seja uma matriz definida positiva no espaço nulo de A .

A obtenção de $\nabla^2 f(x_k)$ a cada iteração é computacionalmente muito caro, pois tem-se de obter n^2 derivadas de segunda ordem de f a cada iteração. Uma saída é aproximar $\nabla^2 f(x_k)$ por H_k , onde H_k é simétrica e definida positiva. Essa escolha assegura que a direção p é sempre uma direção de descida.

A descrição imediatamente anterior, descreve como encontrar uma solução ao problema (4.1), levando em conta a eliminação das restrições do mesmo, ou seja, "integrando" estas à função, basendo-se na idéia para qualquer direção que pertença ao espaço nulo de A um ponto viável somado desta também será viável. Logo, se for utilizada as três primeiras componentes da expansão de Taylor da função ao redor desse ponto o novo problema ficará restrito a que o produto da matriz A de (4.1) com um vetor d qualquer deve ser nulo. A solução para esse novo problema implica em ter de resolver o sistema reduzido (4.6), o qual, em alguns casos, acaba se tornando um processo ineficiente. Isso pelo fato de que tal processo é iterativo e o sistema (4.6) terá de ser resolvido mais de uma vez.

O novo algoritmo proposto, trabalha diretamente no espaço nulo da matriz das restrições, não existindo a necessidade de resolver o sistema reduzido (4.6) a cada iteração do algoritmo. Essa abordagem para resolver (4.1) baseia-se na idéia de conjugacidade. Para formalizar essa idéia utilizada no novo algoritmo proposto, faz-se necessário uma abordagem deste para o problema quadrático. A razão para tal é que o conceito de conjugacidade é estabelecido para problemas com funções quadráticas. O novo algoritmo adapta-se a esse modelo sendo necessárias apenas algumas adaptações.

4.1 O Modelo Quadrático

Um problema de programação quadrática (QP) surge, muitas vezes, como um subproblema em métodos de otimização restrita geral, razão pela qual deve ser sempre resolvido de maneira mais simples e barata possível.

Seja o problema de programação quadrática restrito dado por

$$\begin{aligned} \text{Minimizar } q(x) &= \frac{1}{2}x^T Gx - h^T x + c \\ \text{s.a. } Ax &= b \end{aligned} \quad (4.7)$$

onde G é uma matriz simétrica e definida positiva $n \times n$, h é um vetor de \mathbb{R}^n , A uma matriz $m \times n$ com $m < n$, de posto m , e b um vetor de \mathbb{R}^m . As condições de otimalidade de primeira ordem para o problema (4.7) são,

$$\begin{aligned}\nabla_x l(x^*, \lambda^*) &= 0 \\ Ax^* &= b.\end{aligned}$$

Como $l(x, \lambda) = \frac{1}{2}x^T Gx - h^T x + \lambda^T(b - Ax)$, então as condições anteriores resultam no sistema,

$$\begin{aligned}Gx^* - h - A^T \lambda^* &= 0 \\ Ax^* - b &= 0\end{aligned}\tag{4.8}$$

logo, as condições de otimalidade de primeira ordem são dadas pelo seguinte sistema linear,

$$\begin{pmatrix} G & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x^* \\ \lambda^* \end{pmatrix} = \begin{pmatrix} h \\ b \end{pmatrix}.\tag{4.9}$$

A matriz do sistema (4.7) é chamada de Karush-Kuhn-Tucker (KKT) e o sistema de equações lineares é chamado de sistema KKT.

Sob certas condições sobre as matrizes G e A , as condições de otimalidade de primeira ordem do sistema assume que existe uma única solução para o sistema (4.9), como pode-se observar no seguinte teorema,

Teorema 31 *Suponha que A seja uma matriz $m \times n$, com $m < n$, de posto completo m , e que G seja definida positiva no espaço nulo de A . Então, a matriz Karush-Kuhn-Tucker é não-singular e existe uma única solução (x^*, λ^*) que satisfaça o sistema (4.9).*

Demonstração. Sejam x e y dois vetores arbitrários tais que

$$\begin{pmatrix} G & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0.\tag{4.10}$$

Se $x = 0$, $A^T y = 0$, então $y = 0$ desde que A tenha posto completo. Se $x \neq 0$, $Ax = 0$, então, $x = Zw$ para qualquer vetor não nulo w , onde Z é uma matriz cujas colunas formam uma base para o espaço nulo de A , assim

$$0 = \begin{pmatrix} x \\ y \end{pmatrix}^T \begin{pmatrix} G & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = x^T Gx.$$

Segue que,

$$0 = x^T G x = w^T Z^T G Z w, \quad (4.11)$$

e tal w sendo zero, implicaria em, indo contra todas as hipóteses anteriores que $x = 0$. Assim, a equação (4.10) tem uma única solução se e somente se o vetor $[x \ y]^T = 0$. ■

Levando em consideração a não singularidade da matriz KKT, uma solução para o sistema (4.9) seria muito problemática do ponto de vista computacional visto que se tem de resolver um sistema quadrado aumentado de ordem m para $m + n$, sem saber com antecedência se as matrizes A e G satisfazem as condições do sistema acima. Além disso, do ponto de vista de otimização a matriz KKT é uma Hessiana de uma função indefinida, a qual é bem sabido ter problemas na manipulação [24]. Agora será explorado como resolver o sistema (4.8) para x^* e λ^* ótimos.

4.1.1 Uma Abordagem do Espaço Nulo

O método do espaço nulo, descrito a seguir, não exige a não singularidade da matriz de G , mas somente que as condições do Teorema 1 se cumpram, isto é, A deve ter posto completo e G deve ser definida positiva no espaço nulo de A .

Seja Z uma matriz de dimensões $n \times (n - m)$, cujas colunas formam uma base para o espaço nulo de A , e seja R uma matriz de dimensão $n \times m$ cujas colunas formam uma base para o espaço imagem de A^T . Observe que como A é uma matriz de dimensão $m \times n$, com $m < n$, e posto completo m , então uma escolha trivial para R é fazer $R = A^T$. Como toda solução x^* para o problema (4.9) é um ponto do \mathbb{R}^n , então ele pode ser escrito de maneira única como segue,

$$x^* = Ru + Zv,$$

onde $u \in R^m$ e $v \in R^{n-m}$. Além disso, como a matriz A tem posto completo m e a matriz $[R|Z]$ tem posto completo n , então $A[R \ | \ Z]$ tem posto completo m . Como $A[R \ | \ Z] = [AR \ | \ 0]$, então AR é uma matriz $m \times m$ e não singular. Agora, considere o sistema Karush-Kuhn-Tucker (4.8), o qual é dado pelas equações

$$\begin{aligned} Gx^* - A^T \lambda^* &= h \\ Ax^* &= b \end{aligned} \quad (4.12)$$

Como $x^* = Ru^* + Zv^*$, então temos que,

$$Ax^* = ARu^* + AZv^*.$$

donde se tem que,

$$ARu^* = b \quad \text{ou} \quad u^* = (AR)^{-1} b.$$

Multiplicando-se a primeira equação do sistema(4.12) por Z^T e usando o fato de que $x^* = Ru^* + Zv^*$, segue-se que,

$$Z^T GRu^* + Z^T GZv^* - Z^T A^T \lambda^* = Z^T h,$$

donde,

$$Z^T GRu^* + Z^T GZv^* = Z^T h,$$

e finalmente v^* é solução do sistema,

$$Z^T GZv^* = Z^T h - Z^T GRu^*,$$

ou seja,

$$v^* = (Z^T GZ)^{-1} [Z^T h - Z^T GRu^*].$$

Com a obtenção de u^* e v^* , x^* fica determinado como sendo,

$$x^* = R(AR)^{-1} b + Z(Z^T GZ)^{-1} [Z^T h - Z^T GRu^*].$$

Para determinar o valor de λ^* , multiplica-se a primeira equação de (4.12) por R^T , obtendo-se,

$$R^T Gx^* - R^T A^T \lambda^* = R^T h,$$

donde vem que,

$$\begin{aligned} (AR)^T \lambda^* &= R^T (Gx^* - h) \\ &= R^T \nabla f(x^*), \end{aligned}$$

e finalmente,

$$\lambda^* = (AR)^{-T} R^T \nabla f(x^*).$$

Observe que a obtenção de λ^* não é cara do ponto de vista computacional tendo em vista que a matriz AR já foi anteriormente fatorada para a obtenção da solução x^* . Portanto, λ^* pode ser obtido ao custo da ordem $O(n^2)$ operações.

Esse método, claramente, é bastante atraente se a dimensão do núcleo de A for pequeno. Por outro lado, como temos que resolver um sistema cuja matriz é $Z^T G Z$, é conveniente que Z seja escolhida ortogonal, para não aumentar o número de condição do sistema. Se Z é ortogonal então o número de condição da matriz $(Z^T G Z)$ não é maior do que o número de condição da matriz G , isto é,

$$\text{cond}(Z^T G Z) \leq \text{cond}(G) \cdot \text{cond}(Z)^2,$$

donde então,

$$\text{cond}(Z^T G Z) \leq \text{cond}(G),$$

ou seja, a condição do sistema reduzido, $Z^T G Z$, é pelo menos tão boa quanto a condição do sistema de G . As equações anteriores podem ser bastante simplificadas se introduzirmos nessa abordagem alguns elementos de conjugacidade.

4.1.2 Abordagem de Conjugacidade

A idéia fundamental por trás dessa abordagem de conjugacidade é construir uma matriz M tal que:

$$\begin{aligned} (i) \quad M^T G M &= I \\ (ii) \quad AM &= [0 \quad U]. \end{aligned}$$

Suponha que G , obtida desse modo, seja definida positiva e a matriz M é uma matriz cujas colunas são vetores G -conjugados, que pode ser construída do seguinte modo. Suponha M decomposta em blocos, como sendo,

$$M = [M_1 \quad M_2],$$

onde M_1 é uma matriz $n \times n - m$ e M_2 é uma matriz $m \times m$ tal que,

- (i) $AM_1 = 0$, isto é, as $m - n$ primeiras colunas de M devem estar no espaço nulo de A ;
- (ii) $AM_2 = U$, onde U é uma $m \times m$ matriz triangular superior.

(4.13)

Como A é uma matriz $m \times n$, $m < n$, com posto completo, isto é, $\text{posto}(A) = m$, fazemos a fatoração QR de A^T para obter,

$$A^T = QR,$$

onde Q é uma matriz ortogonal de dimensões $n \times n$ cujas primeiras m colunas de Q geram o espaço vetorial $\mathbb{R}(A^T)$, e as $n - m$ últimas colunas geram o núcleo de A , $\mathcal{N}(A)$. Então particionamos Q do seguinte modo,

$$Q = [Q_1 \quad Q_2].$$

onde Q_1 é uma matriz $m \times m$, e Q_2 é uma matriz $m \times n - m$, e correspondentemente particionamos R como,

$$R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix}.$$

onde R_1 é uma matriz $m \times m$ triangular superior, e 0 é a matriz nula de dimensão $n - m \times m$.

Definimos agora uma matriz de permutação P que permuta os blocos de Q do seguinte modo,

$$\bar{Q} = QP = [Q_2 \quad Q_1],$$

onde, $Q_1 = [q_j]$, para $j = 1, 2, \dots, m$ e $Q_2 = [q_j]$, para $j = m + 1, \dots, n$.

Agora seja X uma matriz definida como

$$\begin{aligned} X &= A^T(AA^T)^{-1}R^1 \\ &= (Q_1R_1)(R_1^TQ_1^TQ_1R_1)^{-1}R_1 \\ &= Q_1R_1R_1^{-1}R_1^{-T}R_1 \\ &= Q_1R_1^{-T}R_1. \end{aligned}$$

A matriz X tem as seguintes propriedades relevantes:

1. $AX = AA^T(AA^T)^{-1}R_1 = R_1$; isto é, AX é uma matriz triangular superior.
2. $Q_2^T X = Q_2^T Q_1 (R_1^{-T} R_1) = 0$; isto é, as colunas de Q_2 são ortogonais as colunas de X .
3. $X^T Q_2 = (R_1^T R_1^{-1}) Q_1^T Q_2 = 0$; isto é, as colunas de X são ortogonais as colunas de Q_2 .

Defina agora a matriz \tilde{Q} como $\tilde{Q} = [Q_2|X]$. \tilde{Q} é quadrada de posto completo n , portanto não singular. Logo, se G é uma matriz definida positiva de qualquer tamanho n . $\tilde{Q}^T G \tilde{Q}$, também é definida positiva, e podemos fazer sua fatoração de Cholesky

$$\tilde{Q}^T G \tilde{Q} = LL^T.$$

Seja $M = \tilde{Q}L^{-T}$, então

$$\begin{aligned} M^T G M &= (\tilde{Q}L^{-T})^T G (\tilde{Q}L^{-T}) \\ &= L^{-1} \tilde{Q}^T G \tilde{Q} L^{-T} \\ &= L^{-1} (LL^T) L^{-T} \\ &= I. \end{aligned}$$

Além disso,

$$\begin{aligned} AM &= A\tilde{Q}L^{-T} \\ &= A \left(Q_2 \mid X \right) L^{-T} \\ &= \left(0 \mid AX \right) L^{-T} \\ &= \left(0 \mid R_1 \right) L^{-T} \\ &= \left(0 \mid U \right) \end{aligned} \tag{4.14}$$

Portanto, M como definida acima, satisfaz as condições (i) e (ii) da equação (4.13). Por construção M pode ser particionada como sendo,

$$M = \left[Z \mid R \right],$$

onde Z é uma base para o espaço nulo de A e R é uma base para o espaço imagem de A^T . Tem-se então que,

$$\begin{aligned}
 I &= M^T G M \\
 &= \begin{pmatrix} Z & | & R \end{pmatrix}^T G \begin{pmatrix} Z & | & R \end{pmatrix} \\
 &= \begin{pmatrix} Z^T \\ R^T \end{pmatrix} G \begin{pmatrix} Z & | & R \end{pmatrix} \\
 &= \begin{pmatrix} Z^T \\ R^T \end{pmatrix} \begin{pmatrix} GZ & | & GR \end{pmatrix} \\
 &= \begin{pmatrix} Z^T GZ & Z^T GR \\ R^T GZ & R^T GR \end{pmatrix}
 \end{aligned}$$

donde então

$$\begin{aligned}
 Z^T GZ &= I \\
 Z^T GR &= 0 \\
 R^T GZ &= 0 \\
 R^T GR &= I.
 \end{aligned}$$

Essas relações obtidas facilitam sobremaneira a solução dos sistemas lineares obtidas na seção 4.1.1 pela abordagem do espaço nulo. De fato, a equação $(AR)u^* = b$, pode ser simplificada do seguinte modo,

$$AM = A \begin{bmatrix} Z & | & R \end{bmatrix} = \begin{bmatrix} AZ & | & AR \end{bmatrix} = \begin{bmatrix} 0 & | & U \end{bmatrix},$$

e portanto $(AR)u^* = b$ pode ser simplificado para $Uu^* = b$, que é um sistema triangular superior. A equação $(Z^T GZ)v^* = Z^T h - Z^T GRu^*$ fica reduzida a $v^* = Z^T h$, e finalmente, $(AR)^T \lambda^* = R^T \nabla f(x^*)$, fica reduzida a $U^T \lambda^* = R^T \nabla f(x^*)$.

A partir da próxima seção será descrito um novo algoritmo, também baseado na idéia de conjugacidade que generaliza as idéias acima descritas, no sentido de que são relaxadas as condições de que $M^T G M = I$, e $AM = \begin{bmatrix} 0 & | & U \end{bmatrix}$ das seções anteriores.

4.2 Descrição do algoritmo

Nesta seção serão detalhadas três possibilidades de se encontrar uma base para o espaço nulo de A . Esse procedimento é o fator de mudança para análise e comparações dos

algoritmos implementados. De posse dos detalhes de cada uma das três possibilidades, visto aqui como três diferentes versões do novo algoritmo, será descrito como é feita a atualização da direção no primeiro processo iterativo dos algoritmos. A maneira como essa direção é atualizada é comum a todas as versões do algoritmo, assim como os passos de atualização do ponto, do tamanho do passo, do gradiente e da hessiana.

4.2.1 Obtenção de uma Aproximação para o Gradiente e a Hessiana

A aproximação H_k para a hessiana da função é calculada utilizando a aproximação por diferenças finitas, ver (3.8) pg. 33. Assim, H_k , a aproximação para hessiana da função no ponto x_k , foi implementada da seguinte maneira:

$$(H_k)_{ij} = \frac{f(x_k+h_i e_i+h_j e_j)-f(x_k+h_i e_i)-f(x_k+h_j e_j)+f(x_k)}{h_j h_i},$$

onde e_j e e_i são as j -ésima e i -ésima colunas da matriz identidade da mesma ordem de H_k e h_j e h_i são números positivos suficientemente pequenos usados como intervalo das diferenças. Para o experimento $h_{i,j} = 10^{-7}$, $j, i = 1, 2, \dots, n$ ($n = \text{ordem da matriz } H_k$). A avaliação analítica do gradiente da função não é um processo simples e nem barato computacionalmente. Para contornar este problema, o gradiente da função no ponto x_k , $\nabla f(x_k)$, foi calculado também por uma aproximação por diferenças finitas, ver (3.5) pg. 32:

$$(g_k)_j = \frac{f(x_k+h_j e_j)-f(x_k)}{h_j},$$

onde e_j é a j -ésima coluna da matriz identidade da mesma ordem de H_k e h_j é um número positivo suficientemente pequeno, para o experimento $h_j = 10^{-7}$, $j = 1, 2, \dots, n$.

4.2.2 Inicialização

A inicialização do algoritmo depende da escolha de um ponto inicial viável x_0 , isto é, $Ax_0 = b$, cuja escolha pode ser feita por qualquer método de solução de sistemas lineares. Em algumas funções foi utilizado o seguinte esquema para encontrar um ponto inicial.

Seja A a matriz dada por (4.1) tal que por simplicidade, as m primeiras colunas de A sejam linearmente independentes, i.e., formem uma matriz básica B , e as $n - m$ colunas restante formem uma matriz não básica N , ou seja, N é uma matriz que possui

vetores linearmente dependentes. Portanto, a matriz A pode ser particionada da seguinte maneira:

$$A = \begin{pmatrix} B & N \end{pmatrix}. \quad (4.15)$$

Segue que

$$Ax = b \Leftrightarrow (B \ N) \begin{pmatrix} x_B \\ x_N \end{pmatrix} = b \quad (4.16)$$

o que implica

$$Bx_B + Nx_N = b \quad (4.17)$$

ou ainda,

$$Bx_B = b - Nx_N.$$

Fazendo $x_N = 0$, implica em resolver o sistema $Bx_B = b$, o qual tem solução x_{B_1} . O ponto inicial x_0 é dado por:

$$\begin{pmatrix} x_{B_1} \\ 0 \end{pmatrix}$$

De posse de x_0 , calculam-se aproximações para hessiana e gradiente da função utilizando a aproximação por diferenças finitas.

A direção d_0 é uma direção inicial que pertence ao espaço nulo de A .

O comprimento de passo inicial é dado por: $\alpha_0 = -\frac{g_0^T d_0}{d_0^T H d_0}$.

4.2.3 Iterações

A cada iteração do Passo 1

Passo 1.(1). calcula-se o novo ponto: $x_{k+1} = x_k + \alpha_k d_k$;

Passo 1.(2). calcula-se o novo valor da função: f_{k+1} ;

Passo 1.(3). calcula-se a nova aproximação do gradiente: g_{k+1} ;

Passo 1.(4). calcula-se o novo tamanho de passo: α_{k+1}

$$\alpha_{k+1} = -\frac{g_{k+1}^T d_k}{d_k^T H d_k}; \quad (4.18)$$

Passo 1.(5). calcula-se a nova direção d_{k+1} ;

Como o espaço nulo de A possui $n - m$ vetores linearmente independentes, i.e., a $\dim \mathfrak{N}(A) = n - m$, então (1), (2), (3), (4) e (5) são repetidos $n - m$ vezes. Ao final dessas repetições é feito a atualização do segundo passo iterativo. Logo em seguida, o algoritmo verifica se o ponto encontrado nesse passo atende o critério de otimalidade. Se verifica, a solução encontrada é a solução ótima, caso contrário, o ponto encontrado como solução no Passo 1 algoritmo vai Passo 2 o qual será estabelecido logo a seguir.

A cada iteração do Passo 2

Passo 2.(1). calcula-se a nova aproximação da hessiana para o ponto: $x_{k+1} = x_k + \alpha_k d_k$.

Ao final dessa atualização de H_k volta-se a (1), (2), (3), (4) e (5) do Passo 1.

4.2.4 Finalização

O algoritmo termina quando $\|d^T g_k\| < \epsilon$, onde d é um conjunto de vetores H -conjugadas para o espaço nulo de A (usada tolerância $\epsilon = 10^{-7}$), isto é, quando o algoritmo verifica o critério de otimalidade, ou quando um número máximo de iterações for alcançado.

4.3 Algoritmo I - Versão QR de A^T

Seja A a matriz dada por (4.1). A forma de implementação do *Algoritmo I* leva em consideração a decomposição QR da matriz A^T . Assim, uma base para o espaço nulo de A é obtida como descrito a seguir.

Faz-se a decomposição QR de A^T , ou seja,

$$A^T = QR = (Q_1:Q_2) \begin{pmatrix} R_1 \\ 0 \end{pmatrix},$$

donde, tem-se que (Q_1) gera $\mathfrak{R}(A^T)$ e (Q_2) gera $\mathfrak{N}(A^T)$.

As colunas de Q_2 formam uma base para o núcleo de A . Seja

$$Q_2 = (q_1 \quad q_2 \quad \dots \quad q_{n-m}).$$

Faça $Z = Q_2$. Logo, $z_1 = q_1, z_2 = q_2, \dots, z_{n-m} = q_{n-m}$ pertencem ao espaço nulo de A , pois $Az_1 = Az_2 = \dots = Az_{n-m} = 0$.

4.4 Algoritmo II - Versão $(B \ N)$

Seja A a matriz dada por (4.1) tal que por simplicidade, as m primeiras colunas de A sejam linearmente independentes, i.e., formem uma matriz básica B , e as $n - m$ colunas restante formem uma matriz não básica N , ou seja, N é uma matriz que possui vetores linearmente dependentes. Portanto, a matriz A pode ser particionada como visto em (4.15) e o sistema $Ax = b$ fica substituído por (4.17). Logo, se tem de resolver

$$Bx_B = b - Nx_N. \quad (4.19)$$

Fazendo $x_N = 0$ em (4.19), implica em resolver o sistema $Bx_B = b$, o qual tem solução x_{B_1} . O vetor x_0 é dado por:

$$\begin{pmatrix} x_{B_1} \\ 0 \end{pmatrix}$$

Fazendo $x_N = e_1$, onde e_1 é o primeiro vetor canônico, implica em resolver o sistema $Bx_B = b - Ne_1$, o qual tem solução x_{B_2} . O vetor x_1 é dado por:

$$\begin{pmatrix} x_{B_2} \\ e_1 \end{pmatrix}$$

Seja $z_1 = x_1 - x_0$. Para $j = 2, 3, \dots, n - m$

Construa o conjunto de vetores Z tal que as colunas sejam dadas por:

$$z_j = x_j - x_{j-1},$$

Os vetores v_j , $j = 1, 2, 3, \dots, n - m$ obtidos pertencem ao espaço nulo de A pois, se $Ax_0 = b$ e $Ax_1 = b$, fazendo $z_1 = x_1 - x_0$, então $Az_1 = A(x_1 - x_0) = b - b = 0$. Logo, para todo $j = 1, 2, 3, \dots, n - m$, $z_j = x_j - x_{j-1}$ tem-se $Az_j = 0$ que é equivalente a dizer que $z_j \in \mathfrak{N}A, \forall j$.

4.5 Algoritmo III - Versão B^{-1}

Seja A a matriz dada por (4.1) tal que as m primeiras colunas de A sejam linearmente independentes, i.e., formem uma matriz básica B , e as $n - m$ colunas restante formem uma matriz não básica N . Portanto, a matriz A e o vetor x podem ser particionados como feito em (4.16). Dessa forma de particionar a matriz resulta o seguinte sistema :

$$Bx_B + Nx_N = b. \quad (4.20)$$

A matriz B é uma matriz não-singular, a qual possui inversa. Logo, pré-multiplicando-se B^{-1} em (4.20), obtém-se

$$x_B + B^{-1}Nx_N = B^{-1}b.$$

Portanto, qualquer solução x de $Ax = b$, tem a forma

$$x = \begin{pmatrix} B^{-1}b - B^{-1}Nx_N \\ x_N \end{pmatrix},$$

Logo, pode-se escrever x como

$$x = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix} b + \begin{pmatrix} -B^{-1}N \\ I_{n-m} \end{pmatrix} x_N. \quad (4.21)$$

Desse modo, uma base Z para o espaço nulo é dada pela matriz $n \times n - m$ de (4.21),

$$Z = \begin{pmatrix} -B^{-1}N \\ I_{n-m} \end{pmatrix}.$$

Pode-se escolher vetores z_i , $i = 1, 2, 3, \dots, n - m$ pertencentes ao espaço nulo de forma que z_i seja igual a i -ésima coluna de Z . Desse modo, para qualquer vetor z_i , $i = 1, 2, 3, \dots, n - m$ tem-se $Az_i = 0$.

A atualização das direções no Passo 1 dos algoritmos está baseada em um mesmo processo, o qual é conhecido como Gram-Schmidt Modificado, ver algoritmo (2) pg. 16. Foi necessária uma adaptação a este processo, pois tal apenas ortogonaliza vetores e o interesse é H_i -ortogonalizar, onde H_i pode ser visto como uma aproximação para a hessiana da função calculada no ponto x_{n-m} , ponto onde o algoritmo encontra uma solução na $n - m$ -ésima iteração do CG , i.e., no fim do primeiro estágio iterativo do algoritmo. O índice i é visto aqui como a i -ésima iteração do segundo processo iterativo. Logo, ao invés da utilização do produto interno usual, foi usado um produto interno em H_i .

O fato das direções tomadas por base inicial serem formadas por colunas que pertencem ao espaço nulo de A , faz com que apenas a H_i -conjugacidade seja atualizada. Evitando o erro (problema) de que uma nova direção seja tomada como uma combinação linear das outras direções.

4.5.1 Obtenção da H -conjugacidade das direções

A atualização da H -conjugacidade nas três versões do novo algoritmo é calculada como descrito a seguir.

Seja Z uma base para o espaço nulo de A . Assim,

$$Z = \begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ z_1 & z_2 & z_3 & \cdots & z_{n-m-1} & z_{n-m} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Para $j = 1$, d_j , faça

$$d_1 = z_1.$$

onde $\langle x, y \rangle_H = x^T H y$ é um produto interno em H . Observe que, como $z_1 \in \mathfrak{N}(A)$

$$A d_1 = A z_1 = 0$$

logo, $d_1 \in \mathfrak{N}(A)$.

Para $i = j + 1, \dots, n - m$, atualize os próximos z'_i s

$$\bar{z}_i = z_i - \frac{\langle d_j, z_i \rangle_H}{\langle d_j, d_j \rangle_H} d_j$$

Então, como d_1 é conhecido,

$$\begin{aligned} \bar{z}_2 &= z_2 - \frac{\langle d_1, z_2 \rangle_H}{\langle d_1, d_1 \rangle_H} d_1 \\ \bar{z}_3 &= z_3 - \frac{\langle d_1, z_3 \rangle_H}{\langle d_1, d_1 \rangle_H} d_1 \\ &\vdots \\ \bar{z}_{n-m} &= z_{n-m} - \frac{\langle d_1, z_{n-m} \rangle_H}{\langle d_1, d_1 \rangle_H} d_1 \end{aligned}$$

Obtenha d_2 , tal que $d_1^T H d_2 = 0$. Para $j = 2$ tem-se

$$d_2 = \bar{z}_2$$

Observe que, como $z_2 \in \mathfrak{N}(A)$ e $d_1 \in \mathfrak{N}(A)$

$$A d_2 = A z_2 - \frac{\langle d_1, z_2 \rangle_H}{\langle d_1, d_1 \rangle_H} A d_1 = 0$$

logo, $d_2 \in \mathcal{N}(A)$ e

$$d_1^T H d_2 = d_1^T H z_2 - \frac{\langle d_1, z_2 \rangle_H}{\langle d_1, d_1 \rangle_H} d_1^T H d_1 = d_1^T H z_2 - \frac{d_1^T H z_2}{d_1^T H d_1} d_1^T H d_1 = 0$$

Assim, como d_2 é conhecido, atualize $\bar{z}_{j+1}, \bar{z}_{j+2}, \dots, \bar{z}_{n-m}$

$$\begin{aligned} \bar{\bar{z}}_3 &= \bar{z}_3 - \frac{\langle d_2, \bar{z}_3 \rangle_H}{\langle d_2, d_2 \rangle_H} d_2 \\ \bar{\bar{z}}_4 &= \bar{z}_4 - \frac{\langle d_2, \bar{z}_4 \rangle_H}{\langle d_2, d_2 \rangle_H} d_2 \\ &\vdots \\ \bar{\bar{z}}_{n-m} &= \bar{z}_{n-m} - \frac{\langle d_2, \bar{z}_{n-m} \rangle_H}{\langle d_2, d_2 \rangle_H} d_2. \end{aligned}$$

e obtenha d_3 , tal que $d_1^T H d_3 = 0$ e $d_2^T H d_3 = 0$. Para $j = 3$ tem-se

$$d_3 = \bar{\bar{z}}_3.$$

Observe que, como $\bar{z}_3, d_2 \in \mathcal{N}(A)$

$$A d_3 = A \bar{z}_3 - \frac{\langle d_2, \bar{z}_3 \rangle_H}{\langle d_2, d_2 \rangle_H} A d_2 = 0$$

logo, $d_3 \in \mathcal{N}(A)$ e

$$\begin{aligned} d_1^T H d_3 &= d_1^T H \bar{z}_3 - \frac{\langle d_2, \bar{z}_3 \rangle_H}{\langle d_2, d_2 \rangle_H} d_1^T H d_2 = d_1^T H \bar{z}_3 - \frac{d_1^T H \bar{z}_3 d_1^T}{d_2^T H d_2} d_1^T H d_2 = 0 \\ d_2^T H d_3 &= d_2^T H \bar{z}_3 - \frac{\langle d_2, \bar{z}_3 \rangle_H}{\langle d_2, d_2 \rangle_H} d_2^T H d_2 = d_2^T H \bar{z}_3 - \frac{d_2^T H \bar{z}_3}{d_2^T H d_2} d_2^T H d_2 = 0. \end{aligned}$$

Assim, sucessivamente, pode-se ir gerando direções que são mutuamente H -conjugadas e que estão no espaço nulo de A . Um algoritmo para obter os vetores H -conjugados é dado a seguir,

Algoritmo 6 *M-GS H-Conjugado*

Para $j = 1, 2, 3, \dots, n - m - 1, n - m$

$$d_i \leftarrow z_i;$$

Para $i = j + 1, \dots, n - m$

$$z_i \leftarrow z_i - \frac{d_j^T H z_i}{d_j^T H d_j} d_j;$$

Fim

Fim

O novo algoritmo a ser apresentado nesse trabalho resolve o problema (4.1), pg. 59, conquanto satisfaz as restrições impostas ao mesmo e avalia se a condição de otimalidade de primeira ordem se verifica. O pseudocódigo desse algoritmo é dado por:

Algoritmo 7 *Algoritmo Básico - Direções Conjugadas*

Dado um ponto inicial viável x_0 ($Ax_0 = b$).

Avalie o valor da função $f_0 = f(x_0)$.

Avalie o valor do gradiente $g_0 = g(x_0)$.

Avalie o valor da hessiana $H_0 = H(x_0)$.

Faça $d_0 \in \mathcal{N}(A)$.

Calcule α_0 de acordo com eq.(4.18), pg.70.

$d = d_0$.

$i = 0$

Enquanto $\|d^T g_k\| \neq 0$

$k = 0$.

Enquanto $k < n - m$

Atualize o ponto $x_{k+1} = x_k + \alpha_k d_k$.

Avalie $f_{k+1} = f(x_{k+1})$.

Avalie $g_{k+1} = g(x_{k+1})$.

Encontre uma direção de busca d_{k+1} de modo que d_{k+1} seja H_i -conjugada com d_k .

Atualize α_{k+1} de acordo com eq.(4.18), pg.70.

$k = k + 1$.

Deixe d formar o conjunto de direções de busca H_i -conjugado.

Fim

$$\bar{x} = x_{k+1}$$

$$\text{Avalie } H_{i+1} = H(\bar{x}).$$

$$i = i + 1$$

Fim

$$x^* = \bar{x}$$

O próximo algoritmo usa o método do *CG* para resolver o sistema reduzido $Z^T H_i Z d_z = -Z^T g_k$.

Algoritmo 8 *Algoritmo Básico - Gradiente Conjugado Reduzido*

Dado um ponto inicial viável x_0 ($Ax_0 = b$).

Avalie o valor da função $f_0 = f(x_0)$.

Avalie o valor do gradiente $g_0 = g(x_0)$.

Avalie o valor da hessiana $H_0 = H(x_0)$.

Encontre uma base Z para o espaço nulo de A .

Calcule α_0 de acordo com eq.(4.18), pg.70.

Faça $X = Z^T H_0 Z$.

$i = 0$

Enquanto $\|Z^T g_k\| \neq 0$

$k = 0$.

Enquanto $k < n - m$

Resolver o sistema $Z^T H_i Z d_z = -Z^T g_k$

Encontre uma direção de busca d_k fazendo $d_k = Z d_z$

Atualize o ponto $x_{k+1} = x_k + \alpha_k d_k$.

Avalie $f_{k+1} = f(x_{k+1})$.

Avalie $g_{k+1} = g(x_{k+1})$.

Atualize α_{k+1} de acordo com eq.(4.18), pg.70.

$k = k + 1$.

Fim

$$\bar{x} = x_{k+1}$$

$$\text{Avalie } H_{i+1} = H(\bar{x}).$$

$$\text{Avalie } X = Z^T H_{i+1} Z.$$

$$i = i + 1$$

Fim

$$x^* = \bar{x}$$

Experimentos Numéricos

Para analisar o comportamento do algoritmo foram efetuados vários experimentos numéricos. Na implementação do algoritmo, os problemas de minimização com restrições lineares são definidos como:

$$\begin{aligned} & \text{Minimizar} && f(x) \\ & \text{s.a.} && Ax = b \end{aligned} \tag{5.1}$$

Na apresentação das tabelas contendo os resultados obtidos se utiliza a seguinte simbologia:

1. *Iterações*: número da iteração do laço principal do algoritmo;
2. $|Z^T g|$: no caso dos algoritmos de direções conjugadas *norma* – 2 do produto interno entre a base Z formada iterativamente e o vetor gradiente, e o caso dos algoritmos de gradiente conjugado reduzido, Z é uma base para o espaço nulo de A ;
3. $f(x^*)$: valor da função objetivo na solução obtida;
4. $\frac{\|x_{k+1}-x^*\|}{\|x_k-x^*\|}$: é a taxa de convergência a cada iteração k . No caso da função de Powell a taxa de convergência foi calculada em relação a um novo ponto encontrado pelo algoritmo quando a exigência sobre $|Z^T g|$ era imediatamente inferior a 10^{-7} ;
5. *Tempo(s)*: tempo de CPU em segundos;
6. *1200* na coluna de *iterações*: significa que o programa parou por ter atingido o número máximo de iterações permitido.

Em todos os casos foi utilizado aproximação por diferenças finitas para as hessianas das funções objetivo. O ponto inicial utilizado para o problema está na região viável do mesmo.

As funções utilizadas pelos algoritmos nos experimentos numéricos são funções clássicas na PNL e estão descritas a seguir. A dimensão de algumas das funções é ajustável via um parâmetro simples usado na função, no caso 'n'.

1. Função Wood

(a) $n = 4$;

(b) $f(x) = 100(x_1^2 - x_2)^2 + (1 - x_1)^2 + 90(x_3^2 - x_4)^2 + (1 - x_3)^2 +$
 $+10.1((1 - x_2)^2 + (1 - x_4)^2) + 19.8(1 - x_2)(1 - x_4)$;

(c) Mínimo global: $x^* = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$;

(d) $f(x^*) = 0$.

3. Função Rosenbrock Singular Extendida

(a) $n =$ um número múltiplo de 2;

(b) para $i = 1, \dots, n/2$

1. $f_{2i-1}(x) = 10(x_{2i} - x_{2i-1}^2)$;

2. $f_{2i}(x) = (1 - x_{2i-1})$;

3. $f(x) = \sum_{i=1}^{n/2} (f_{2i-1}(x)^2 + f_{2i}(x)^2)$;

(c) Mínimo global: $x^* = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$;

(d) $f(x^*) = 0$.

4. Função Powell Singular Extendida

(a) $n =$ um número inteiro múltiplo de 4;

(b) para $i = 1, \dots, n/4$

1. $f_{4i-3}(x) = x_{4i-3} + 10x_{4i-2}$;

$$2. f_{4i-2}(x) = \sqrt{5}(x_{4i-1} - 10x_{4i});$$

$$3. f_{4i-1}(x) = (x_{4i-2} - 2x_{4i-1})^2;$$

$$4. f_{4i}(x) = \sqrt{10}(x_{4i-3} - x_{4i})^2;$$

$$5. f(x) = \sum_{i=1}^{n/4} (f_{4i-3}(x)^2 + f_{4i-2}(x)^2 + f_{4i-1}(x)^2 + f_{4i}(x)^2);$$

$$(c) \text{ M\u00ednimo global: } x^* = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix};$$

$$(d) f(x^*) = 0.$$

Todos os algoritmos foram implementados em MatLab.

O objetivo destes testes realizados \u00e9 comparar o resultado obtido pelo algoritmo de CG Reduzido o qual resolve o sistema reduzido a cada itera\u00e7\u00e3o do la\u00e7o central do algoritmo.

O conjunto de restri\u00e7\u00f5es utilizadas para resolver o problema (5.1), p\u00e1gina 79, est\u00e3o neste trabalho em anexo. Para Wood os dados encontram-se na tabela (A.1), p\u00e1gina 86. Os dados utilizados para a fun\u00e7\u00e3o Rosenbrock Singular Extendida est\u00e3o na tabela (A.2), p\u00e1gina 87 e para a fun\u00e7\u00e3o Powell os mesmos encontram-se na tabela (A.3), p\u00e1gina 87.

Conclusões

A abordagem mais promissora encontrada na literatura para resolver problemas de programação não linear com restrições lineares de igualdade usa como idéia básica reduzir a dimensão do espaço no qual o problema está imerso. Essa abordagem, contudo, não está isenta de problemas, principalmente aqueles dependentes da escolha da base para o espaço nulo das restrições. De fato, a escolha de uma base não ortogonal pode introduzir mal condicionamento nos subproblemas que são resolvidos a cada iteração, e nos casos onde a matriz Hessiana é esparsa, os sistemas reduzidos apresentam intenso fenômeno de "fill in". O método de direções conjugadas apresentado neste trabalho contorna em parte esses dois problemas. Portanto, as idéias apresentadas aqui tem vantagens sobre a abordagem descrita, de reduzir a dimensão do espaço, quando a Hessiana da função a ser minimizada é esparsa e quando a escolha da base para representar o espaço nulo das restrições não é ortogonal.

Nesse trabalho a H -ortogonalização dos vetores foi feita baseada no método de Gram-Schmidt modificado (M-GS). Fica em aberto para trabalhos futuros o uso de outros métodos de ortogonalização de vetores como, por exemplo, rotação de Givens, o que poderá trazer mais eficiência ao algoritmo. Outra questão em aberto é o quanto e de que maneira a escolha da base para o espaço nulo das restrições impacta na convergência do algoritmo. De todo modo os resultados numéricos oriundos dos testes são bastante promissores quanto às potencialidades da abordagem apresentada quando a Hessiana da função é mal condicionada ou é esparsa.

Referências Bibliográficas

- [1] BASSANI, Carmen. *Estudo da influência da escolha da base para o núcleo das restrições na minimização de uma função não linear sujeita a restrições linear*. Dissertação (mestrado) – Pontifícia Universidade Católica do Paraná, Curitiba, 2004.
- [2] BAUMANN, J, et. al. *Nonlinear programming applied to calibrating thermal and fluid models to test data*. 8° Annual IEEE Symposium, março 2002, p. 77-82.
- [3] BJÖRCK, Ake. *Numerical methods for least squares problems*. Philadelphia : SIAM, 1996.
- [4] BORGES, C.L.T., FALCÃO, D.M. e COUTINHO, A.L.G.A.: *Utilização de método tipo gradiente conjugado na aceleração do fluxo de potência em computação vetorial*. XIV SNPTEE Seminário nacional de Produção e Transmissão de Energia Elétrica, Belém-PA, 1997.
- [5] COLLEMAN, T.F.: *Linearly constrained optimization and projected preconditioned conjugate gradients*, in Proceedings of the Fifth SIAM Conference on Applied Linear Algebra, SIAM: Philadelphia, pp.118 - 122, 1994.
- [6] COLLEMAN, T.F. and VERMA, A.: *A preconditioned conjugate gradient approach to linear equality constrained minimization*, Computer Science Department and Cornell Theory Center, Cornell University, Ithaca, New York, USA, pp. 61 - 72, 2001.
- [7] COLLEMAN, T.F., LIU, Jianguo; YUAN, Wei: *A New Trust-Region Algorithm for Equality Constrained Optimization*, Computational Optimization and Applications, pp. 177 - 199, 2002.

- [8] DELOWSKI, Ana Paula. *Um novo algoritmo baseado em conjugacidade para resolver o problema de minimização de uma função não linear sujeita a restrições lineares de igualdade*. Dissertação (mestrado) – Pontifícia Universidade Católica do Paraná, Curitiba, 2005.
- [9] DENNIS, J. E.; SCHNABEL, R. B. *Numerical methods for unconstrained optimization and nonlinear equations*. SIAM, Philadelphia, 1996.
- [10] FACÓ, J. L. D. *Nonlinear Programming Solutions for Optimal Control Problems*. Mathematical Programming in Rio, 2003, Búzios - RJ. Proceedings of Mathematical Programming in Rio, 2003. v. I. p. 62-66.
- [11] FLETCHER, Roger: *Practical methods of optimization*. A Wiley Interscience Publication, Chichester, N.Y.: Wiley, 2nd ed, 1987.
- [12] GILBERT, J.R.; HEATH, M.: *Computing a sparse basis for the nullspace*, SIAM J. Alg. & Disc. Meth., vol. 8, pp.446 - 459, 1987.
- [13] GILL, Philip E.; MURRAY, Walter; WRIGHT, Margareth H . *Practical Optimization*. Academic Press, 1981.
- [14] GILL, P., MURRAY, D. Ponceleon e SAUNDERS, M.: *Preconditiones for indefinite systems arising in optimization*. SIAMJ. Matrix Anal. Appl., pp. 292-311, 1992.
- [15] GOLUB, Gene H.; Charles F. Van Loan: *Matrix computation*. The Johns Hopkins University Press Ltd., London, 3 edição, 1996.
- [16] HESTENES, Magnus R. *Applications of Mathematics: Conjugate Direction Methods in Optimization*. Springer-Verlag, New York, 1980.
- [17] HESTENES, Magnus R.; STIEFEL, E. *Methods of Conjugate Gradient for Solving Linear Systems*. J. Res. Nat. Bur. Standards 49, 1942, p. 409-436.
- [18] LUENBERG, David G. *Introduction to linear and nonlinear programming*. Addison-Wesley Publishing Company, Massachusetts, 1973.
- [19] MARSDEN, James R.; PINGRY, David E.; WHINSTON, Andrew. *Application of nonlinear programming to water quality control.*, Water, Air, & Soil Pollution, Volume 2, Número 2, Páginas 155 - 169.

-
- [20] MARTINEZ, J.M.; SANTOS, S.A. *Métodos Computacionais de Otimização*, XX Colóquio Brasileiro de Matemática, IMPA, 1995.
- [21] NASH, S.G.; SOFER, A. *Linear and Nonlinear Programming*. McGrawHill, Singapore, 1996.
- [22] NASH, S.G. e SOFER, A.: *Preconditioning Reduced Matrices*. SIAM. Matrix Anal. Appl., pp. 47-68, 1996.
- [23] NOBLE, Ben; DANIEL, James W. *Applied Linear Algebra*. 3ªed. Prentice-Hall, New Jersey, 1988.
- [24] NOCEDAL, J.; WRIGHT, Stephen J. *Numerical Optimization*. Springer, New York, 1999.
- [25] STRANG, Gilbert. *Linear Algebra and its applications*. Hartcour Brace, San Diego, 3 edição, 1988.
- [26] WAH, Benjamin W.; CHEN, Yi-Xin. *Constrained genetic algorithms and their applications in nonlinear constrained optimization*. 12th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'00), p. 286.

A

Tabelas de Dados

Este anexo contém os dados utilizados para resolver o problema,

$$\begin{array}{ll} \textit{Minimizar} & f(x) \\ \textit{s.a.} & Ax = b \end{array}$$

os quais estão dispostos nas tabelas a seguir.

Para a função quadrática foram utilizadas matriz com dados aleatórios, onde apenas os elementos da hessiana são dados pela representação de Hilbert.

A	b	x^*	$f(x^*)$
2 1 -3 1	1	1	0
0 1 1 0	2	1	
		1	
		1	

Tabela A.1: Dados do problema com a função Wood

A										b	x^*	$f(x^*)$
											1	0
											1	
											1	
											1	
10	2	-1	4	6	12	-3	1	2	-1	32	1	
0	10	1	0	-1	1	1	0	0	1	13	1	
0	0	1	1	0	-1	1	0	0	1	3	1	
											1	
											1	
											1	

Tabela A.2: Dados do problema com a função Rosenbrock Singular Extendida

A												b	x^*	$f(x^*)$
													2,7327	33,523
													-0,1972	
													0,5258	
													2,2684	
3	1	-1	2	0	7	1	-2	-6	5	1	-3	10	2,4555	
0	3	0	-2	1	4	0	8	0	1	-1	0	12	-0,1170	
0	1	0	5	4	0	1	0	6	0	1	0	21	1,8708	
													-0,0725	
													0,0581	
													-0,1168	
													-0,4516	

Tabela A.3: Dados do problema com a função Powell Singular Extendida

Matriz Hessiana H	Matriz A	Vetor h	Vetor b
10×10	5×10	10×1	5×1
10×10	2×10	10×1	2×1
17×17	6×17	17×1	6×1

Tabela A.4: Dados Função Quadrática - $f(x) = \frac{1}{2}x^T Hx - h^T x$

B

Tabelas de Resultados

Este anexo contém alguns resultados obtidos, utilizando o software MatLab, para as funções Wood $n = 4$, Rosenbrock Extendida com $n = 10$ e Powell Singular Extendida com $n = 12$ e quadrática com Hessiana de Hilbert.

Iterações	$\ Z^T \nabla f(x_k)\ $	$f(x^*)$	Tempo(s)
5	$7,7415 \cdot 10^{-12}$	$-3,6163 \cdot 10^3$	0,03
8	$8,0787 \cdot 10^{-8}$	$-9,7915 \cdot 10^6$	0,02
11	$6,7843 \cdot 10^{-6}$	$-1,9536 \cdot 10^{+10}$	0,09

Tabela B.1: Resultados Função Quadrática - Direções Conjugadas QR de A^T

Iterações	$\ Z^T \nabla f(x_k)\ $	$f(x^*)$	Tempo(s)
5	$9,7275 \cdot 10^{-12}$	$-3,6163 \cdot 10^3$	0,03
8	$3,5253 \cdot 10^{-8}$	$-9,7915 \cdot 10^6$	0,03
11	$1,9314 \cdot 10^{-6}$	$-1,9536 \cdot 10^{+10}$	0,12

Tabela B.2: Resultados Função Quadrática - Direções Conjugadas ($B \mid N$)

Iterações	$\ Z^T \nabla f(x_k)\ $	$f(x^*)$	Tempo(s)
5	$2,1881 \cdot 10^{-12}$	$-3,6163 \cdot 10^3$	0,03
8	$3,4131 \cdot 10^{-8}$	$-9,7915 \cdot 10^6$	0,031
11	$1,7800 \cdot 10^{-6}$	$-1,9536 \cdot 10^{+10}$	0,09

Tabela B.3: Resultados Função Quadrática - Direções Conjugadas B^{-1}

Iterações	$\ Z^T \nabla f(x_k)\ $	$f(x^*)$	Tempo(s)
7	$2,0807 \cdot 10^{-12}$	$-3,6163 \cdot 10^3$	0,05
28	$3,5640 \cdot 10^{-9}$	$-9,7915 \cdot 10^6$	0,06
276	$9,2509 \cdot 10^{-6}$	$-1,9536 \cdot 10^{+10}$	0,14

Tabela B.4: Resultados Função Quadrática - *CG Reduzido*

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1}-x^*\ }{\ x_k-x^*\ }$	Tempo(s)
1	$1,4829 \cdot 10^{+2}$	$3,0953 \cdot 10^{+2}$	—	
2	$3,0645 \cdot 10^{-1}$	$2,2111 \cdot 10^{+2}$	0,8069	
3	$4,5324 \cdot 10^{-1}$	$0,6341 \cdot 10^{+2}$	0,7117	
4	$2,7742 \cdot 10^{-1}$	$0,2695 \cdot 10^{+2}$	0,8429	
5	$2,6120 \cdot 10^{-1}$	$0,1153 \cdot 10^{+2}$	0,6843	
6	$1,3477 \cdot 10^{-1}$	$0,0319 \cdot 10^{+2}$	0,5379	
⋮	⋮	⋮	⋮	0,04
24	$1,8454 \cdot 10^{-6}$	$5,84 \cdot 10^{-10}$	0,5223	
25	$9,8425 \cdot 10^{-7}$	$1,83 \cdot 10^{-10}$	0,5122	
26	$5,2493 \cdot 10^{-7}$	$6,2 \cdot 10^{-11}$	0,4922	
27	$2,7996 \cdot 10^{-7}$	$2,4 \cdot 10^{-11}$	0,4498	
28	$1,4931 \cdot 10^{-7}$	$1,1 \cdot 10^{-11}$	0,3478	
29	$0,7963 \cdot 10^{-7}$	$6 \cdot 10^{-12}$	0,2849	

Tabela B.5: Resultados Wood - Direções Conjugadas QR de A^T

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1}-x^*\ }{\ x_k-x^*\ }$	Tempo(s)
1	$1,5748 \cdot 10^{+2}$	$3,0909 \cdot 10^{+2}$	—	
2	$2,6593 \cdot 10^{+2}$	$2,2289 \cdot 10^{+2}$	0,8014	
3	$2,0581 \cdot 10^{+2}$	$4,4468 \cdot 10^{+1}$	0,6440	
4	$0,4521 \cdot 10^{-1}$	$1,5014 \cdot 10^{-1}$	0,0760	0,05
5	$1,6416 \cdot 10^{-3}$	$1,7866 \cdot 10^{-8}$	0,0003	
6	$0 \cdot 10^{-11}$	$3 \cdot 10^{-12}$	0,0103	

Tabela B.6: Resultados Wood - CG Reduzido QR de A^T

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1}-x^*\ }{\ x_k-x^*\ }$	Tempo(s)
1	$3,4522 \cdot 10^{+2}$	$3,0953 \cdot 10^{+2}$	–	
2	$1,3273 \cdot 10^{-1}$	$2,2355 \cdot 10^{+2}$	0,8069	
3	$1,5939 \cdot 10^{-1}$	$0,6593 \cdot 10^{+2}$	0,7117	
4	$2,1284 \cdot 10^{-1}$	$0,4403 \cdot 10^{+2}$	0,8429	
5	$1,5483 \cdot 10^{-1}$	$0,2710 \cdot 10^{+2}$	0,6843	
6	$1,2498 \cdot 10^{-1}$	$0,1754 \cdot 10^{+2}$	0,5379	
⋮	⋮	⋮	⋮	0,10
74	$2,2240 \cdot 10^{-7}$	$0,7485 \cdot 10^{-10}$	0,5223	
75	$1,8277 \cdot 10^{-7}$	$0,5415 \cdot 10^{-10}$	0,5122	
76	$1,5019 \cdot 10^{-7}$	$0,3968 \cdot 10^{-10}$	0,4922	
77	$1,2340 \cdot 10^{-7}$	$0,2950 \cdot 10^{-10}$	0,4498	
78	$1,0137 \cdot 10^{-7}$	$0,2230 \cdot 10^{-10}$	0,3478	
79	$0,8327 \cdot 10^{-7}$	$0,1716 \cdot 10^{-10}$	0,3156	

Tabela B.7: Resultados Wood - Direções Conjugadas B_1^{-1}

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1}-x^*\ }{\ x_k-x^*\ }$	Tempo(s)
1	$3,4557 \cdot 10^{+2}$	$3,0909 \cdot 10^{+2}$	–	
2	$6,4729 \cdot 10^{+2}$	$2,2290 \cdot 10^{+2}$	0,8013	
3	$4,7909 \cdot 10^{+2}$	$4,6096 \cdot 10^{+1}$	0,6596	
4	$2,2939 \cdot 10^{+1}$	$4,3901 \cdot 10^{-2}$	0,0124	0,06
5	$3,5580 \cdot 10^{-3}$	$9,84 \cdot 10^{-10}$	0,0001	
6	$0 \cdot 10^{-12}$	$3 \cdot 10^{-12}$	0	

Tabela B.8: Resultados Wood - CG Reduzido utilizando B_1^{-1}

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1}-x^*\ }{\ x_k-x^*\ }$	Tempo(s)
1	$3,4803 \cdot 10^{+2}$	$3,0909 \cdot 10^{+2}$	–	
2	$2,6100 \cdot 10^{-2}$	$1,8084 \cdot 10^{+2}$	0,7198	
3	$3,5524 \cdot 10^{-4}$	$0,6058 \cdot 10^{+2}$	0,8647	
4	$1,4766 \cdot 10^{-4}$	$0,5825 \cdot 10^{+2}$	0,9654	
5	$3,0631 \cdot 10^{-4}$	$0,5605 \cdot 10^{+2}$	0,9863	
6	$3,6381 \cdot 10^{-4}$	$0,5345 \cdot 10^{+2}$	0,9789	
⋮	⋮	⋮	⋮	0,83
607	$1,1067 \cdot 10^{-7}$	$6,28 \cdot 10^{-10}$	0,8269	
608	$1,0821 \cdot 10^{-7}$	$6 \cdot 10^{-10}$	0,7938	
608	$1,0581 \cdot 10^{-7}$	$5,74 \cdot 10^{-10}$	0,7442	
610	$1,0345 \cdot 10^{-7}$	$5,49 \cdot 10^{-10}$	0,6615	
611	$1,0115 \cdot 10^{-7}$	$5,24 \cdot 10^{-10}$	0,5961	
612	$0,9890 \cdot 10^{-7}$	$5,01 \cdot 10^{-10}$	0,5794	

Tabela B.9: Resultados Wood - Direções Conjugadas ($B_1 | N$)

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1}-x^*\ }{\ x_k-x^*\ }$	Tempo(s)
1	$4,7731 \cdot 10^{+2}$	$3,0909 \cdot 10^{+2}$	–	
2	$9,5688 \cdot 10^{+2}$	$2,2324 \cdot 10^{+2}$	0,8010	
3	$7,3839 \cdot 10^{+2}$	$3,9390 \cdot 10^{+1}$	0,5985	
4	$3,5504 \cdot 10^{+2}$	$1,9609 \cdot 10^{+1}$	0,8099	0,06
5	$0,2591 \cdot 10^{+1}$	$1,0130 \cdot 10^{-1}$	0,0833	
6	$1,4667 \cdot 10^{-3}$	$7,348 \cdot 10^{-9}$	0,0002	
7	$0 \cdot 10^{-12}$	$3 \cdot 10^{-12}$	0	

Tabela B.10: Resultados Wood - CG Reduzido ($B_1 | N$)

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1}-x^*\ }{\ x_k-x^*\ }$	Tempo(s)
1	$1,9949 \cdot 10^{+1}$	$3,0909 \cdot 10^{+2}$	–	
2	$4,4025 \cdot 10^{-2}$	$2,9308 \cdot 10^{+2}$	0,6557	
3	$4,6518 \cdot 10^{-2}$	$2,7621 \cdot 10^{+2}$	0,4156	
4	$5,0286 \cdot 10^{-2}$	$2,5722 \cdot 10^{+2}$	0,7162	
5	$1,2371 \cdot 10^{-1}$	$2,0799 \cdot 10^{+2}$	0,3442	
6	$1,3021 \cdot 10^{-1}$	$1,4525 \cdot 10^{+2}$	0,6940	
⋮	⋮	⋮	⋮	0,04
19	$7,7924 \cdot 10^{-7}$	$2,2730 \cdot 10^{-11}$	0,4493	
20	$5,1534 \cdot 10^{-7}$	$9,9401 \cdot 10^{-12}$	0,4463	
21	$3,4072 \cdot 10^{-7}$	$4,3400 \cdot 10^{-12}$	0,4402	
22	$2,2521 \cdot 10^{-7}$	$0,1900 \cdot 10^{-13}$	0,4272	
23	$1,4882 \cdot 10^{-7}$	$0,0830 \cdot 10^{-13}$	0,3968	
24	$0,9831 \cdot 10^{-7}$	$0,0360 \cdot 10^{-14}$	0,3172	

Tabela B.11: Resultados Wood - Direções Conjugadas ($B_2 \mid N$)

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1} - x^*\ }{\ x_k - x^*\ }$	Tempo(s)
1	$3,6853 \cdot 10^{+2}$	$3,0909 \cdot 10^{+2}$	—	
2	$5,1939 \cdot 10^{+2}$	$2,2365 \cdot 10^{+2}$	0,7980	
3	$4,1461 \cdot 10^{+2}$	$0,6862 \cdot 10^{+2}$	0,7152	
4	$0,2855 \cdot 10^{+2}$	$0,0033 \cdot 10^{+2}$	0,0915	
5	$0,0713 \cdot 10^{+2}$	$8,6987 \cdot 10^{-1}$	0,5235	
6	$0,0527 \cdot 10^{+2}$	$0,1029 \cdot 10^{-1}$	0,3343	
⋮	⋮	⋮	⋮	0,03
22	$5,2841 \cdot 10^{-4}$	$9,7447 \cdot 10^{-11}$	0,3175	
23	$1,1794 \cdot 10^{-4}$	$2,2667 \cdot 10^{-12}$	0,5571	
24	$9,9347 \cdot 10^{-5}$	$3,0483 \cdot 10^{-12}$	0,31756	
25	$0,2086 \cdot 10^{-6}$	$8,3433 \cdot 10^{-12}$	0,5570	
26	$1,6535 \cdot 10^{-7}$	$9,5389 \cdot 10^{-13}$	0,3175	
27	$0,3691 \cdot 10^{-7}$	$0,2611 \cdot 10^{-14}$	0,5570	

Tabela B.12: Resultados Wood - CG Reduzido ($B_2 \mid N$)

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1}-x^*\ }{\ x_k-x^*\ }$	Tempo(s)
1	$1,9949 \cdot 10^{+1}$	$3,0909 \cdot 10^{+2}$	–	
2	$1,5262 \cdot 10^{-2}$	$3,0282 \cdot 10^{+2}$	0,6557	
3	$1,4186 \cdot 10^{-2}$	$2,9767 \cdot 10^{+2}$	0,7448	
4	$1,3530 \cdot 10^{-2}$	$2,9335 \cdot 10^{+2}$	1,3283	
5	$1,3230 \cdot 10^{-2}$	$2,8960 \cdot 10^{+2}$	4,6344	
6	$1,3270 \cdot 10^{-2}$	$2,8626 \cdot 10^{+2}$	0,9167	
⋮	⋮	⋮	⋮	0,06
38	$5,0439 \cdot 10^{-6}$	$4,3297 \cdot 10^{-8}$	0,0612	
39	$2,2842 \cdot 10^{-6}$	$8,8800 \cdot 10^{-9}$	0,0939	
40	$1,0344 \cdot 10^{-6}$	$1,8210 \cdot 10^{-9}$	0,1117	
41	$4,6841 \cdot 10^{-7}$	$3,7300 \cdot 10^{-10}$	0,1115	
42	$2,1209 \cdot 10^{-7}$	$7,7000 \cdot 10^{-11}$	0,1013	
43	$0,9602 \cdot 10^{-7}$	$1,6000 \cdot 10^{-11}$	0,0986	

Tabela B.13: Resultados Wood - Direções Conjugadas B_2^{-1}

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1} - x^*\ }{\ x_k - x^*\ }$	Tempo(s)
1	$3,8845 \cdot 10^{+2}$	$3,0909 \cdot 10^{+2}$	—	
2	$3,8603 \cdot 10^{+2}$	$2,2133 \cdot 10^{+2}$	0,8043	
3	$3,5070 \cdot 10^{+2}$	$0,0054 \cdot 10^{+1}$	0,7158	
4	$0,0743 \cdot 10^{+2}$	$2,2645 \cdot 10^{-2}$	0,0411	
5	$0,0152 \cdot 10^{+2}$	$0,8686 \cdot 10^{-2}$	0,2034	
6	$2,9820 \cdot 10^{-1}$	$0,3307 \cdot 10^{-3}$	0,1957	
⋮	⋮	⋮	⋮	0,02
11	$0,8414 \cdot 10^{-3}$	$6,9154 \cdot 10^{-12}$	0,2000	
12	$0,1641 \cdot 10^{-3}$	$2,6296 \cdot 10^{-13}$	0,1950	
13	$3,1996 \cdot 10^{-5}$	$0,9999 \cdot 10^{-14}$	0,1950	
14	$0,6239 \cdot 10^{-5}$	$0,3802 \cdot 10^{-15}$	0,1950	
15	$1,2167 \cdot 10^{-7}$	$1,4457 \cdot 10^{-17}$	0,1950	
16	$0,2373 \cdot 10^{-7}$	$0,0550 \cdot 10^{-17}$	0,1951	

Tabela B.14: Resultados Wood - CG Reduzido B_2^{-1}

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1}-x^*\ }{\ x_k-x^*\ }$	Tempo(s)
1	$6,3820 \cdot 10^{+2}$	$1,9499 \cdot 10^{+3}$	–	
2	$3,3831 \cdot 10^{-1}$	$1,7417 \cdot 10^{+3}$	0,9336	
3	$7,7555 \cdot 10^{-1}$	$1,2187 \cdot 10^{+3}$	0,8854	
4	$6,2543 \cdot 10^{-1}$	$5,5224 \cdot 10^{+2}$	0,9106	
5	$3,0454 \cdot 10^{-1}$	$0,9974 \cdot 10^{+2}$	0,9101	
6	$1,7813 \cdot 10^{-1}$	$0,3100 \cdot 10^{+2}$	0,9886	
⋮	⋮	⋮	⋮	107,40
1168	$1,1411 \cdot 10^{-5}$	$1,4074 \cdot 10^{-9}$	0,8985	
1169	$1,1091 \cdot 10^{-5}$	$1,3808 \cdot 10^{-9}$	0,7559	
1170	$0,1078 \cdot 10^{-6}$	$1,3553 \cdot 10^{-9}$	0,6832	
1171	$0,1047 \cdot 10^{-6}$	$1,3306 \cdot 10^{-9}$	0,5082	
1172	$1,0177 \cdot 10^{-7}$	$1,3069 \cdot 10^{-9}$	0,4168	
1173	$0,9888 \cdot 10^{-7}$	$1,2842 \cdot 10^{-9}$	0,4657	

Tabela B.15: Resultados Rosenbrock Singular Extendida - Direções Conjugadas $QR A^T$

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1}-x^*\ }{\ x_k-x^*\ }$	Tempo(s)
1	$6.7723 \cdot 10^{+2}$	$1,9498 \cdot 10^{+3}$	—	
2	$2.5936 \cdot 10^{+2}$	$1,6666 \cdot 10^{+2}$	0,7540	
3	$2.5927 \cdot 10^{+2}$	$1,6661 \cdot 10^{+3}$	0,9991	
4	$2.5919 \cdot 10^{+2}$	$1,6657 \cdot 10^{+2}$	0,9991	
5	$2.5910 \cdot 10^{+2}$	$1,6652 \cdot 10^{+2}$	0,9991	
6	$2.5902 \cdot 10^{+2}$	$1,6647 \cdot 10^{+2}$	0,9991	
⋮	⋮	⋮	⋮	104,47
1195	$2,2333 \cdot 10^{+2}$	$1,2016 \cdot 10^{+2}$	0,8571	
1196	$2,2331 \cdot 10^{+2}$	$1,2012 \cdot 10^{+2}$	0,8333	
1197	$2,2329 \cdot 10^{+2}$	$1,2009 \cdot 10^{+2}$	0,7999	
1198	$2,2328 \cdot 10^{+2}$	$1,2005 \cdot 10^{+2}$	0,7499	
1199	$2,2326 \cdot 10^{+2}$	$1,2002 \cdot 10^{+2}$	0,6666	
1200	$2,2325 \cdot 10^{+2}$	$1,1998 \cdot 10^{+2}$	0,4999	

Tabela B.16: Resultados Rosenbrock Singular Extendida - CG Reduzido QR de A^T

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1}-x^*\ }{\ x_k-x^*\ }$	Tempo(s)
1	$6,0252 \cdot 10^{+2}$	$6,6689 \cdot 10^{+3}$	–	
2	$0,1294 \cdot 10^{+1}$	$2,0005 \cdot 10^{+3}$	0,6700	
3	$8,8303 \cdot 10^{-1}$	$1,5271 \cdot 10^{+2}$	0,4026	
4	$3,0048 \cdot 10^{-1}$	$0,5688 \cdot 10^{+2}$	0,7206	
5	$2,4483 \cdot 10^{-1}$	$0,4218 \cdot 10^{+2}$	0,7255	
6	$1,2716 \cdot 10^{-1}$	$0,3772 \cdot 10^{+2}$	0,7293	
⋮	⋮	⋮	⋮	7,24
16	$5,9193 \cdot 10^{-5}$	$3,3522 \cdot 10^{+1}$	0,2488	
17	$1,2866 \cdot 10^{-5}$	$3,3522 \cdot 10^{+1}$	0,2302	
18	$3,6761 \cdot 10^{-6}$	$3,3522 \cdot 10^{+1}$	0,2463	
19	$0,7072 \cdot 10^{-6}$	$3,3522 \cdot 10^{+1}$	0,1892	
20	$1,1981 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,1468	
21	$0,2535 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,1321	

Tabela B.17: Resultados Powell Singular Extendida - Direções Conjugadas QR de A^T

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1}-x^*\ }{\ x_k-x^*\ }$	Tempo(s)
1	$3,5531 \cdot 10^{+3}$	$6,6689 \cdot 10^{+3}$	—	
2	$0,1854 \cdot 10^{+3}$	$0,1535 \cdot 10^{+3}$	0,3072	
3	$0,1529 \cdot 10^{+2}$	$3,7877 \cdot 10^{+2}$	0,4634	
4	$2,2716 \cdot 10^{-1}$	$3,3523 \cdot 10^{+1}$	0,6163	
5	$1,6936 \cdot 10^{-5}$	$3,3523 \cdot 10^{+1}$	0,7586	
6	$1,3906 \cdot 10^{-5}$	$3,3523 \cdot 10^{+1}$	0,8058	
⋮	⋮	⋮	⋮	1,45
14	$0,1411 \cdot 10^{-6}$	$3,3523 \cdot 10^{+1}$	0,4665	
15	$0,1388 \cdot 10^{-6}$	$3,3523 \cdot 10^{+1}$	0,5569	
16	$0,1431 \cdot 10^{-6}$	$3,3523 \cdot 10^{+1}$	0,4970	
17	$0,1471 \cdot 10^{-6}$	$3,3523 \cdot 10^{+1}$	0,6095	
18	$1,2912 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,6246	
19	$0,9951 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,5098	

Tabela B.18: Resultados Powell Singular Extendida - CG Reduzido QR de A^T

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1}-x^*\ }{\ x_k-x^*\ }$	Tempo(s)
1	$2,3757 \cdot 10^{+3}$	$6,6689 \cdot 10^{+3}$	–	
2	$6,4832 \cdot 10^{-1}$	$2,0587 \cdot 10^{+3}$	0,7100	
3	$6,0932 \cdot 10^{-1}$	$0,3282 \cdot 10^{+3}$	0,5001	
4	$6,3649 \cdot 10^{-1}$	$0,1366 \cdot 10^{+3}$	0,5402	
5	$2,9301 \cdot 10^{-1}$	$0,7546 \cdot 10^{+2}$	0,6972	
6	$1,5461 \cdot 10^{-1}$	$0,5139 \cdot 10^{+2}$	0,6951	
⋮	⋮	⋮	⋮	12,12
33	$1,9015 \cdot 10^{-6}$	$3,3522 \cdot 10^{+1}$	0,5239	
34	$1,0437 \cdot 10^{-6}$	$3,3522 \cdot 10^{+1}$	0,5209	
35	$5,9904 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,5009	
36	$2,8346 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,4423	
37	$1,7204 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,3649	
38	$0,8511 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,3256	

Tabela B.19: Resultados Powell Singular Extendida - Direções Conjugadas B_1^{-1}

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1}-x^*\ }{\ x_k-x^*\ }$	Tempo(s)
1	$1,2392 \cdot 10^4$	$6,6689 \cdot 10^3$	—	
2	$6,7601 \cdot 10^2$	$0,1535 \cdot 10^3$	0,3072	
3	$0,7621 \cdot 10^2$	$0,4963 \cdot 10^2$	0,4638	
4	$0,4070 \cdot 10^2$	$0,3836 \cdot 10^2$	0,6145	
5	$0,2421 \cdot 10^2$	$0,3559 \cdot 10^2$	0,7551	
6	$0,2306 \cdot 10^2$	$3,5041 \cdot 10^2$	0,7955	
⋮	⋮	⋮	⋮	7,30
91	$2,2727 \cdot 10^{-5}$	$3,3522 \cdot 10^1$	0,9371	
92	$2,1336 \cdot 10^{-5}$	$3,3522 \cdot 10^1$	0,4368	
93	$1,9067 \cdot 10^{-5}$	$3,3522 \cdot 10^1$	0,8956	
94	$0,1563 \cdot 10^{-6}$	$3,3522 \cdot 10^1$	0,9995	
95	$1,5577 \cdot 10^{-7}$	$3,3522 \cdot 10^1$	0,7649	
96	$0,8886 \cdot 10^{-7}$	$3,3522 \cdot 10^1$	0,7554	

Tabela B.20: Resultados Powell Singular Extendida - CG Reduzido B_1^{-1}

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1}-x^*\ }{\ x_k-x^*\ }$	Tempo(s)
1	$2,3757 \cdot 10^{+3}$	$6,6689 \cdot 10^{+3}$	–	
2	$4,6418 \cdot 10^{-1}$	$2,2423 \cdot 10^{+3}$	0,7219	
3	$4,0803 \cdot 10^{-1}$	$0,4803 \cdot 10^{+3}$	0,5528	
4	$4,0306 \cdot 10^{-1}$	$0,2671 \cdot 10^{+3}$	0,7994	
5	$0,2487 \cdot 10^{+1}$	$2,0039 \cdot 10^{+2}$	1,0660	
6	$0,9975 \cdot 10^{-1}$	$1,1026 \cdot 10^{+2}$	0,8048	
⋮	⋮	⋮	⋮	27,94
162	$0,3981 \cdot 10^{-6}$	$3,3522 \cdot 10^{+1}$	0,6489	
163	$0,2649 \cdot 10^{-6}$	$3,3522 \cdot 10^{+1}$	0,6235	
164	$0,1972 \cdot 10^{-6}$	$3,3522 \cdot 10^{+1}$	0,5746	
165	$1,0319 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,5953	
166	$1,0680 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,3969	
167	$0,5972 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,3721	

Tabela B.21: Resultados Powell Singular Extendida - Direções Conjugadas ($B_1 \mid N$)

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1}-x^*\ }{\ x_k-x^*\ }$	Tempo(s)
1	$1,5314 \cdot 10^4$	$6,6689 \cdot 10^3$	—	
2	$8,8263 \cdot 10^2$	$0,1560 \cdot 10^3$	0,3010	
3	$0,5086 \cdot 10^2$	$0,4849 \cdot 10^2$	0,4515	
4	$2,7223 \cdot 10^{-1}$	$0,3567 \cdot 10^2$	0,5542	
5	$0,2679 \cdot 10^{-5}$	$0,3377 \cdot 10^2$	0,6272	
6	$0,7168 \cdot 10^{-5}$	$0,3336 \cdot 10^2$	0,7465	
⋮	⋮	⋮	⋮	29,14
378	$0,2296 \cdot 10^{-5}$	$3,2550 \cdot 10^1$	0,6002	
379	$0,4192 \cdot 10^{-5}$	$3,2549 \cdot 10^1$	0,5414	
380	$0,3364 \cdot 10^{-5}$	$3,2549 \cdot 10^1$	0,8281	
381	$3,2733 \cdot 10^{-6}$	$3,2549 \cdot 10^1$	0,8697	
382	$3,2851 \cdot 10^{-7}$	$3,2549 \cdot 10^1$	0,6303	
383	$0,9750 \cdot 10^{-7}$	$3,2549 \cdot 10^1$	0,6002	

Tabela B.22: Resultados Powell Singular Extendida - CG Reduzido ($B_1 \mid N$)

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1}-x^*\ }{\ x_k-x^*\ }$	Tempo(s)
1	$5,7921 \cdot 10^{+4}$	$3,8478 \cdot 10^{+4}$	–	
2	$1,3281 \cdot 10^{-1}$	$4,1772 \cdot 10^{+3}$	0,6633	
3	$2,0648 \cdot 10^{-1}$	$2,8420 \cdot 10^{+3}$	0,8876	
4	$2,3270 \cdot 10^{-1}$	$1,9977 \cdot 10^{+3}$	0,9218	
5	$2,1678 \cdot 10^{-1}$	$1,3692 \cdot 10^{+3}$	0,9005	
6	$1,9163 \cdot 10^{-1}$	$9,5443 \cdot 10^{+2}$	0,8930	
⋮	⋮	⋮	⋮	22,43
114	$4,635 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,6319	
115	$3,138 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,6101	
116	$2,303 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,5825	
117	$1,530 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,5104	
118	$1,021 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,3900	
119	$0,594 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,3526	

Tabela B.23: Resultados Powell Singular Extendida - Direções Conjugadas B_2^{-1}

Iteração	$\ Z^T \nabla f(x_k)\ $	$f(x_k)$	$\frac{\ x_{k+1}-x^*\ }{\ x_k-x^*\ }$	Tempo(s)
1	$5,7921 \cdot 10^{+4}$	$3,8478 \cdot 10^{+4}$	—	
2	$8,3898 \cdot 10^{-1}$	$4,5689 \cdot 10^{+3}$	0,6545	
3	$9,2109 \cdot 10^{-1}$	$3,4414 \cdot 10^{+3}$	0,9231	
4	$1,3090 \cdot 10^{-1}$	$2,9142 \cdot 10^{+3}$	0,9427	
5	$1,2746 \cdot 10^{-1}$	$2,5126 \cdot 10^{+3}$	0,9453	
6	$1,0727 \cdot 10^{-1}$	$2,2256 \cdot 10^{+3}$	0,9603	
⋮	⋮	⋮	⋮	57,387
300	$4,125 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,6887	
301	$3,166 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,6698	
302	$2,286 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,6151	
303	$1,639 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,5655	
304	$1,185 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,4191	
305	$0,871 \cdot 10^{-7}$	$3,3522 \cdot 10^{+1}$	0,3865	

Tabela B.24: Resultados Powell Singular Extendida - Direções Conjugadas ($B_2 \mid N$)