

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM
COMPUTACIONAL

Hernando José Rocha Franco

**Método de Direções Interiores ao Epígrafo - IED para Otimização Não
Diferenciável e Não Convexa via Dualidade Lagrangeana: Estratégias para
Minimização da Lagrangeana Aumentada**

Tese apresentada ao Programa de Pós-graduação em Modelagem Computacional da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do grau de Doutor em Modelagem Computacional.

Orientador: Prof. D.Sc. Wilhelm Passarella Freire

Coorientador: Prof. D.Sc. Luis Paulo da Silva Barra

Juiz de Fora

2018

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Franco, Hernando José Rocha.

Método de Direções Interiores ao Epígrafo - IED para Otimização Não Diferenciável e Não Convexa via Dualidade Lagrangeana: Estratégias para Minimização da Lagrangeana Aumentada / Hernando José Rocha Franco. -- 2018.

112 f.

Orientador: Wilhelm Passarella Freire

Coorientador: Luis Paulo da Silva Barra

Tese (doutorado) - Universidade Federal de Juiz de Fora, ICE/Engenharia. Programa de Pós-Graduação em Modelagem Computacional, 2018.

1. Dualidade Lagrangeana Aumentada. 2. Otimização Não Diferenciável. 3. Otimização Não Convexa. I. Freire, Wilhelm Passarella , orient. II. Barra, Luis Paulo da Silva, coorient. III. Título.


Hernando José Rocha Franco

Método de Direções Interiores ao Epígrafo - IED para Otimização Não Diferenciável e Não Convexa via Dualidade Lagrangeana: Estratégias para Minimização da Lagrangeana Aumentada

Tese apresentada ao Programa de Pós-graduação em Modelagem Computacional da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do grau de Doutor em Modelagem Computacional.

Aprovada em 7 de Junho de 2018.

BANCA EXAMINADORA



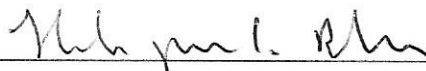
Prof. D.Sc. Wilhelm Passarella Freire - Orientador
Universidade Federal de Juiz de Fora



Prof. D.Sc. Luis Paulo da Silva Barra - Coorientador
Universidade Federal de Juiz de Fora



Prof. D.Sc. Afonso Celso de Castro Lemonge
Universidade Federal de Juiz de Fora



Prof. D.Sc. Hélio José Corrêa Barbosa
Universidade Federal de Juiz de Fora



Prof. D.Sc. José Herskovits Norman
Universidade Federal do Rio de Janeiro



Prof. D.Sc. Dênis Emanuel da Costa Vargas
Instituto Federal de Educação, Ciência e Tecnologia do Sudeste MG

*Dedico este trabalho aos meus familiares:
minha esposa Cláudia e minha filha Fernanda,
meus pais Francisco (in memoriam) e Maria José,
meus irmãos Juliana e Aristóteles.
Ao tio Totinho, em esperança.*

AGRADECIMENTOS

A Deus, que me permitiu mais uma significativa experiência de vida.

Aos meus orientadores, prof. Wilhelm e prof. Luis Paulo, por me oportunizarem a entrada neste doutorado. Agradeço-lhes também pela confiança, pelo apoio no desenvolvimento deste trabalho e pelo convívio desses anos.

Aos colegas do PPGMC, pela amizade e pela troca de conhecimentos em momentos tão diversos do curso. À Camila Fernandes, colega do núcleo de estudos em matemática computacional na primeira fase do curso. Em especial, ao Tales Lima, pelo companheirismo e contribuição com os algoritmos genéticos.

Aos professores, que nas diferentes fases deixaram suas contribuições neste estudo, bem como colaboraram para o meu aperfeiçoamento profissional.

Ao IF Sudeste MG, por apoiar a qualificação docente. Em particular, aos meus pares do Câmpus Avançado Bom Sucesso, que foram solidários e incentivadores nesta trajetória.

Ao Ignácio, pelo constante incentivo.

A todos que, ao longo desses quatro anos e meio, deixaram contribuições que me ajudaram a avançar as etapas do doutorado.

Obrigado!

“Se, a princípio, a ideia não é absurda, então não há esperança para ela.”

Albert Einstein.

RESUMO

A teoria clássica de otimização presume a existência de certas condições, por exemplo, que as funções envolvidas em um problema desta natureza sejam pelo menos uma vez continuamente diferenciáveis. Entretanto, em muitas aplicações práticas que requerem o emprego de métodos de otimização, essa característica não se encontra presente. Problemas de otimização não diferenciáveis são considerados mais difíceis de lidar. Nesta classe, aqueles que envolvem funções não convexas são ainda mais complexos. O *Interior Epigraph Directions* (IED) é um método de otimização que se baseia na teoria da Dualidade Lagrangeana e se aplica à resolução de problemas não diferenciáveis, não convexas e com restrições. Neste estudo, apresentamos duas novas versões para o referido método a partir de implementações computacionais de outros algoritmos. A primeira versão, denominada IED+NFDNA, recebeu a incorporação de uma implementação do algoritmo *Nonsmooth Feasible Direction Nonconvex Algorithm* (NFDNA). Esta versão, ao ser aplicada em experimentos numéricos com problemas teste da literatura, apresentou desempenho satisfatório quando comparada ao IED original e a outros solvers de otimização. Com o objetivo de aperfeiçoar mais o método, reduzindo sua dependência de parâmetros iniciais e também do cálculo de subgradientes, uma segunda versão, IED+GA, foi desenvolvida com a utilização de algoritmos genéticos. Além da resolução de problemas teste, o IED+GA obteve bons resultados quando aplicado a problemas de engenharia.

Palavras-chave: Dualidade Lagrangeana aumentada. Otimização não diferenciável. Otimização não convexa.

ABSTRACT

The classical theory of optimization assumes the existence of certain conditions, for example, that the functions involved in a problem of this nature are at least once continuously differentiable. However, in many practical applications that require the use of optimization methods, this characteristic is not present. Non-differentiable optimization problems are considered more difficult to deal with. In this class, those involving nonconvex functions are even more complex. Interior Epigraph Directions (IED) is an optimization method that is based on Lagrangean duality theory and applies to the resolution of non-differentiable, non-convex and constrained problems. In this study, we present two new versions for this method from computational implementations of other algorithms. The first version, called IED + NFDNA, received the incorporation of an implementation of the Nonsmooth Feasible Direction Nonconvex Algorithm (NFDNA) algorithm. This version, when applied in numerical experiments with problems in the literature, presented satisfactory performance when compared to the original IED and other optimization solvers. A second version, IED + GA, was developed with the use of genetic algorithms in order to further refine the method, reducing its dependence on initial parameters and also on the calculation of subgradients. In addition to solving test problems, IED + GA achieved good results when applied to engineering problems.

Keywords: Augmented Lagrangian duality. Nonsmooth optimization. Nonconvex optimization.

SUMÁRIO

1	INTRODUÇÃO.....	18
1.1	JUSTIFICATIVAS	19
1.2	OBJETIVOS	19
1.3	ESTRUTURA DO TRABALHO	20
2	NOÇÕES PRELIMINARES.....	21
2.1	DEFINIÇÕES GERAIS	21
2.1.1	<i>Direção viável.....</i>	21
2.1.2	<i>Direção de descida</i>	21
2.1.3	<i>Ponto de mínimo</i>	22
2.1.4	<i>Condição necessária de otimalidade de 1ª ordem</i>	22
2.2	CONVEXIDADE	22
2.2.1	<i>Conjunto convexo</i>	22
2.2.2	<i>Função convexa</i>	23
2.2.3	<i>Epígrafo</i>	23
2.2.4	<i>Subgradiente</i>	24
2.3	NÃO CONVEXIDADE	26
2.3.1	<i>Função Lipschitz contínua.....</i>	26
2.4	CONDIÇÕES DE KARUSH-KUHN-TUCKER	26
2.5	ELEMENTOS DA TEORIA DE DUALIDADE	27
2.5.1	<i>Supremo e ínfimo</i>	27
2.5.2	<i>Dualidade fraca</i>	30
3	MÉTODOS DE OTIMIZAÇÃO	31

3.1	MÉTODO DE PLANOS DE CORTE	31
3.2	MÉTODOS DE FEIXES	33
4	ALGORITMOS	34
4.1	A DIREÇÃO DE BUSCA DO FDIPA	34
4.2	NFDA E A APROXIMAÇÃO DO EPÍGRAFO	37
4.3	NFDNA E A EXTENSÃO PARA O CASO NÃO CONVEXO	39
4.3.1	<i>Algoritmo NFDNA</i>	41
4.4	O MÉTODO IED E O TRATAMENTO DAS RESTRIÇÕES VIA DUALIDADE LAGRANGIANA	43
4.4.1	<i>O cone de direções</i>	45
4.4.2	<i>Algoritmo IED</i>	48
4.4.2.1	<i>Algoritmo DSG</i>	50
4.4.2.2	<i>Função fminsearch: resolução do subproblema</i>	51
5	IED+NFDNA: UMA PRIMEIRA VERSÃO DO MÉTODO	58
5.1	EXPERIMENTOS NUMÉRICOS	60
5.1.1	<i>Experimento Numérico 1</i>	60
5.1.1.1	<i>Descrição dos problemas</i>	61
5.1.1.2	<i>Análise do experimento numérico 1</i>	63
5.1.2	<i>Experimento Numérico 2</i>	64
5.1.2.1	<i>Descrição dos problemas</i>	64
5.1.2.2	<i>Análise do experimento numérico 2</i>	65
5.1.3	<i>Experimento Numérico 3</i>	65
5.1.3.1	<i>Descrição dos problemas</i>	66
5.1.3.2	<i>Análise do experimento numérico 3</i>	67
6	IED+GA: UMA VERSÃO HÍBRIDA DO IED COM ALGORITMOS	

GENÉTICOS	68
6.1 ALGORITMOS GENÉTICOS	68
6.1.1 <i>Exemplo de passo evolutivo</i>	70
6.1.2 <i>Características dos operadores genéticos do IED+GA</i>	72
6.1.2.1 <i>Codificação</i>	72
6.1.2.2 <i>População inicial</i>	72
6.1.2.3 <i>Operador de seleção</i>	72
6.1.2.4 <i>Operador de recombinação</i>	73
6.1.2.5 <i>Operador de mutação</i>	74
6.2 O ALGORITMO IED+GA	75
6.2.1 <i>Experimentos Numéricos</i>	77
6.2.1.1 <i>Descrição dos problemas numéricos</i>	77
6.2.1.2 <i>Comparação dos resultados</i>	83
6.2.1.3 <i>Análise dos experimentos</i>	94
7 APLICAÇÃO DO MÉTODO IED NA RESOLUÇÃO DE PROBLEMAS DE ENGENHARIA	96
7.1 Comparação dos resultados dos experimentos numéricos	101
7.1.1 <i>Análise dos resultados</i>	104
8 CONSIDERAÇÕES FINAIS	106
8.1 CONTRIBUIÇÕES DESTE TRABALHO	107
8.2 TRABALHOS FUTUROS	107
REFERÊNCIAS	109

LISTA DE ILUSTRAÇÕES

2.1	Epígrafos: função convexa (a) e função não convexa (b) (Fonte: [1]).	24
2.2	y^1 e y^2 são subgradientes de f em x (Fonte: [2]).	24
2.3	Subgradientes no ponto $x = 0$	25
2.4	Gráfico função dual.	29
3.1	Três iterações do método de planos de corte (Fonte: [?]).	32
4.1	Direção de busca do FDIPA (Fonte: [3]).	36
4.2	Passo sério do NFDA (Fonte: o autor).	38
4.3	Cálculo de um plano viável (Fonte: [4]).	41
4.4	O cone de direções $K(v, z)$ (Fonte: o autor).	47
4.5	Passo sério: $(v^{k+1}, z^{k+1}) = (\hat{v}^k, \hat{z}^k)$ (Fonte: [5]).	49
4.6	Passo nulo: $(v^k, z^k) = (\tilde{v}^{k,i}, \tilde{z}^{k,i})$ (Fonte: [5]).	50
4.7	Método de Nelder-Mead após uma etapa de reflexão e expansão (Fonte: [6]).	52
4.8	Método de Nelder-Mead após uma contração externa, uma contração interna e uma redução (Fonte: [6]).	52
4.9	Reflexão e expansão.	55
4.10	Contração.	56
4.11	Sequência T_k dos triângulos [7, p. 409].	56
5.1	Diagrama de fluxo do método IED+NFDNA (Fonte: o autor).	59
5.2	Sequência histórica dos métodos (Fonte: o autor).	60
6.1	Função hipotética com um máximo local e outro global. (Fonte: [8]).	69
6.2	Ilustração da função (6.1).	70
6.3	Representação gráfica da Tabela 6.3.	73

6.4	Fluxograma IED+AG.	76
6.5	Fluxograma do AG utilizado pelo IED.	76
7.1	Mola sob tração/compressão.	97
7.2	Redutor de velocidade.	99
7.3	Viga soldada.	100
7.4	Vaso de pressão.	101

LISTA DE TABELAS

5.1	Valores iniciais das variáveis duais.	62
5.2	Resultados obtidos pelas duas versões do IED.	63
5.3	Resultados obtidos pelo IED+NFDNA.	65
5.4	Resultados comparativos entre IED+NFDNA e Fmincon.	67
5.5	Resultados comparativos entre IED+NFDNA e Fmincon.	67
6.1	População $P(0)$	71
6.2	População $P(1)$	72
6.3	Exemplo de seleção por roleta.	73
6.4	Parâmetros do IED.	84
6.5	Resultados comparativos do problema GLR-P1-1.	85
6.6	Resultados comparativos do problema GQR-P1-1.	85
6.7	Resultados comparativos do problema PPR-P1-2.	86
6.8	Resultados comparativos do problema PQR-T1-7.	86
6.9	Resultados comparativos do problema SQR-P1-1.	87
6.10	Resultados comparativos do problema LGR-P1-1.	87
6.11	Resultados comparativos do problema QBR-T1-1.	88
6.12	Resultados comparativos do problema PBR-T1-1.	88
6.13	Resultados comparativos do problema PBR-T1-2.	89
6.14	Resultados comparativos do problema QQR-P1-3.	89
6.15	Resultados comparativos do problema QQR-T1-6.	90
6.16	Resultados comparativos do problema PLR-T1-1.	90
6.17	Resultados comparativos do problema PBR-T1-3.	91
6.18	Resultados comparativos do problema QLR-T1-1.	91

6.19	Resultados comparativos do problema PQR-T1-1.	92
6.20	Resultados comparativos do problema QQR-T1-3.	92
6.21	Resultados comparativos do problema GBR-T1-1.	93
6.22	Resultados comparativos do problema PQR-T1-4.	93
6.23	Resultados comparativos do problema QQR-T1-2.	94
6.24	Resultados comparativos do problema QPR-T1-1.	94
7.1	Parâmetros do IED+GA.	102
7.2	Resultados dos problemas de engenharia.	102
7.3	Comparação das soluções do problema da mola sob tração.	102
7.4	Comparação das soluções do problema redutor de velocidade.	103
7.5	Comparação das soluções do problema da viga soldada.	104
7.6	Comparação das soluções do problema do vaso de pressão.	104

LISTA DE ABREVIATURAS E SIGLAS

DSG	Deflected Subgradient Method
EngOpt	Engineering Optimization
FDIPA	Feasible Direction Interior Point Algorithm
IED	Interior Epigraph Directions
NFDA	Nonsmooth Feasible Direction Algorithm
NFDNA	Nonsmooth Feasible Direction Nonconvex Algorithm

LISTA DE SÍMBOLOS

\mathbb{R}^n	espaço euclidiano de dimensão n
\mathbb{R}	conjunto dos números reais
$\nabla f(x)$	gradiente de f em x
$B(x, \epsilon)$	bola aberta de raio ϵ centrada no ponto x
$\langle x, y \rangle$	produto interno de dois vetores $x, y \in \mathbb{R}^n$
$\partial f(x)$	subdiferencial de f em x
$\ x\ $	norma euclidiana de $x \in \mathbb{R}^n$
$\nabla^2 f(x)$	matriz Hessiana da função f em x

1 INTRODUÇÃO

Otimização é uma área da Matemática que se refere ao estudo de problemas que consistem em minimizar ou maximizar funções cujas variáveis satisfaçam determinadas restrições na forma de igualdades ou desigualdades.

Os problemas reais de otimização estão presentes nos mais diversos campos científicos, por exemplo, em Economia, Mecânica, Gerenciamento, Engenharia e outros, em que se busca derivar soluções ótimas ou as melhores possíveis para estes problemas a partir dos modelos matemáticos que os representam [9].

Para além deste contexto mais prático, a Otimização (ou Programação Matemática) busca avançar no desenvolvimento e aperfeiçoamento de técnicas computacionais com o objetivo de resolver determinada classe de problemas.

Um problema de otimização, conforme as características das funções envolvidas, pode ser formulado como linear ou não linear, restrito ou irrestrito, diferenciável ou não diferenciável, convexo ou não convexo.

Do ponto de vista computacional, o problema de programação linear (PPL) se caracteriza pela existência de poucos métodos de resolução, por exemplo, o conhecido Simplex [10]. Para o caso não linear, diferentes algoritmos são necessários para os diversos tipos das funções objetivo e das restrições.

A Otimização Diferenciável trabalha com funções que possuem derivadas em todos os pontos de seu domínio, enquanto a Otimização Não Diferenciável (*Nonsmooth Optimization*), por tratar de problemas cujas funções não possuem derivadas em alguns pontos, utiliza técnicas da chamada Análise Convexa em substituição ao Cálculo Diferencial clássico [11].

Para lidar com a não diferenciabilidade das funções e consequente inviabilidade de aplicação dos métodos clássicos de Otimização Diferenciável nestes casos, outros métodos foram desenvolvidos, por exemplo, o método de Planos de Corte, proposto por Kelley [12] os métodos de Feixe de Lemaréchal [13].

A maioria dos algoritmos de Otimização Não Diferenciável utiliza o conceito de

subgradiente da função, o qual generaliza o gradiente de uma função convexa para pontos não diferenciáveis. No caso de funções não convexas, condições de otimalidade podem ser estendidas para a classe de funções localmente Lipschitz contínuas [14].

1.1 JUSTIFICATIVAS

1. Embora a teoria clássica de otimização pressuponha certa diferenciabilidade e regularidade das funções, em muitas aplicações, a própria natureza física do problema pode conduzir a um modelo em que tais características não estão presentes [15].
2. Nas últimas décadas a Otimização Não Diferenciável tem recebido avanços importantes. Os métodos considerados mais eficientes e confiáveis são aqueles aplicáveis a funções convexas, não diferenciáveis [1].
3. Problemas envolvendo funções não diferenciáveis e não convexas são considerados mais difíceis de lidar. Além disso, poucos métodos têm sido encontrados na literatura voltados à resolução de problemas com tais funções. ([16], [17]).

1.2 OBJETIVOS

O objetivo principal desta pesquisa é buscar o aperfeiçoamento do método *Interior Epigraph Directions* (IED), que se aplica à otimização não diferenciável, não convexa e com restrições.

Os objetivos específicos são:

1. Implementar computacionalmente outros algoritmos que possam substituir funções ou sub-rotinas do IED original, a fim de torná-lo mais robusto e aplicável a problemas nos quais esta versão não obtivera resolução;
2. Verificar a compatibilidade dessas alterações no desenvolvimento de outras versões do IED;
3. Comparar estas diferentes versões com outros solvers de otimização na resolução de problemas teste da literatura;

4. Avaliar o método IED na resolução de problemas de otimização de áreas afins.

1.3 ESTRUTURA DO TRABALHO

Para organizar o desenvolvimento desta pesquisa, estruturamos este trabalho conforme os tópicos a seguir:

A presente introdução, inicia-se com uma breve contextualização da área de pesquisa e, a partir das justificativas, é apresentado o objetivo principal e as possíveis etapas para cumpri-lo.

No Capítulo 2 são descritos conceitos e resultados matemáticos que fundamentam este estudo.

No Capítulo 3 descrevemos as características principais de alguns métodos clássicos de otimização que foram utilizados na construção dos algoritmos envolvidos na pesquisa.

O Capítulo 4 descreve o estado da arte dos algoritmos que historicamente antecederam o método IED. Foram destacados os principais aspectos de cada um desses métodos no que se refere à influência exercida no desenvolvimento do IED.

No Capítulo 5 é apresentada a primeira versão do método IED desenvolvida nesta pesquisa, denominada IED+NFDNA. Os resultados numéricos comparativos estão tabelados e analisados.

No capítulo 6 uma versão híbrida do IED é apresentada, o IED+GA, com a utilização de algoritmos genéticos na formulação do novo algoritmo.

O capítulo 7 mostra a aplicação do IED+GA na resolução de problemas de Engenharia.

E, finalmente, o Capítulo 8 apresenta as considerações finais desta pesquisa, suas contribuições e possíveis trabalhos futuros.

2 NOÇÕES PRELIMINARES

Com o propósito de contextualizar os problemas e métodos que serão aqui abordados, apresenta-se neste capítulo uma revisão de algumas definições e resultados necessários para o desenvolvimento dos capítulos subsequentes, objetivando a generalização das condições clássicas de otimalidade para o caso não diferenciável. Também serão revisados aspectos teóricos da Dualidade Lagrangeana, técnica utilizada pelo principal algoritmo estudado neste trabalho.

2.1 DEFINIÇÕES GERAIS

Dada uma função escalar $f : \mathbb{R}^n \rightarrow \mathbb{R}$, as definições seguintes são de uso comum em Otimização:

2.1.1 Direção viável

Uma direção se chama viável em um dado ponto de um conjunto X se um deslocamento suficientemente pequeno nesta direção, a partir deste ponto, não produz outro ponto fora de X . Formalmente, um vetor $d \in \mathbb{R}^n \setminus \{0\}$ se chama direção viável em $x \in X \subset \mathbb{R}^n$ se existe um $t_0 > 0$ tal que

$$x + td \in X, \quad \text{para } 0 < t \leq t_0.$$

2.1.2 Direção de descida

Um vetor $d \in \mathbb{R}^n$ se chama direção de descida de f no ponto x , se existe um $t_0 > 0$ tal que

$$f(x + td) < f(x), \quad \text{para } 0 < t \leq t_0.$$

Se f é diferenciável, a condição

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

é suficiente para que d seja uma direção de descida de f em x .

2.1.3 Ponto de mínimo

(i) Um ponto x é um mínimo **local** de f , se existe um $\epsilon > 0$, de modo que

$$f(x) \leq f(y), \quad \forall y \in B(x, \epsilon),$$

onde $B(x, \epsilon)$ representa a bola aberta de centro x e raio ϵ .

(ii) Um ponto x é um mínimo **global** de f , se satisfaz

$$f(x) \leq f(y), \quad \forall y \in \mathbb{R}^n.$$

Os conceitos anteriores são analogamente definidos para máximos, trocando-se apenas os sinais das desigualdades.

2.1.4 Condição necessária de otimalidade de 1ª ordem

Se x é um ponto de mínimo local de f ,

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

é satisfeita para todas as direções viáveis d em x .

Se o ponto x é interior ao domínio de f e também mínimo local, então

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

2.2 CONVEXIDADE

Na falta da diferenciabilidade das funções, a presença de outras características como a convexidade, por exemplo, permite a utilização de resultados importantes na resolução dos problemas de otimização. Introduz-se aqui alguns conceitos sobre convexidade de conjuntos e funções.

2.2.1 Conjunto convexo

Um conjunto convexo se caracteriza por conter todos os segmentos cujos extremos pertencem ao conjunto.

Formalmente, um subconjunto $C \subset \mathbb{R}^n$ é convexo se, para quaisquer dois pontos $x, y \in C$, o segmento

$$\lambda x + (1 - \lambda)y \in C, \quad \forall \lambda \in [0, 1].$$

O ponto $\lambda x + (1 - \lambda)y$, onde $\lambda \in [0, 1]$, se chama **combinação convexa** de x e y .

2.2.2 Função convexa

Seja f uma função definida em um conjunto convexo $C \subset \mathbb{R}^n$.

(i) f é convexa, se

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y), \quad \forall x, y \in C \quad \text{e} \quad \forall \lambda \in [0, 1].$$

(ii) f é estritamente convexa, se

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y), \quad \forall x, y \in C \quad \text{e} \quad \forall \lambda \in (0, 1).$$

(iii) f é (estritamente) côncava, se $-f$ é (estritamente) convexa.

Teorema 2.2.1. *Seja $f : C \rightarrow \mathbb{R}$ uma função convexa ($C \subset \mathbb{R}^n$ convexo). Cada ponto de mínimo local de f é global. ([1], pág.83)*

Para uma função convexa diferenciável, a condição de otimalidade necessária de 1ª ordem 2.1.4 é também suficiente para um mínimo global.

2.2.3 Epígrafo

O epígrafo de uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é o conjunto

$$\text{epi}(f) = \{(x, r) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq r\}.$$

Para $n = 1$ e $n = 2$, o epígrafo é o conjunto de todos os pontos acima do gráfico de f (figura 2.1).

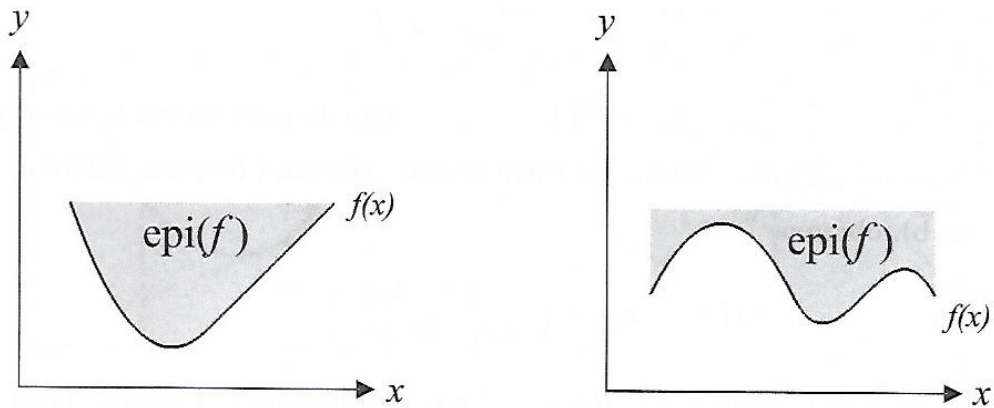


Figura 2.1: Epígrafos: função convexa (a) e função não convexa (b) (Fonte: [1]).

2.2.4 Subgradiente

Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa. O vetor $y \in \mathbb{R}^n$ se chama subgradiente de f no ponto x , se ele é o gradiente de uma função suporte de f em x , isto é, se y satisfaz

$$f(z) \geq f(x) + \langle y, z - x \rangle, \quad \forall z \in \mathbb{R}^n$$

onde $\langle u, v \rangle$ denota o produto escalar de u e v .

O conjunto de todos os subgradientes f em x , denotado por $\partial f(x)$, chama-se subdiferencial de f em x . Em outras palavras, o subgradiente define uma aproximação linear de f cujo gráfico fica abaixo daquele de f e cujo valor coincide com f no ponto x (Figura 2.2).

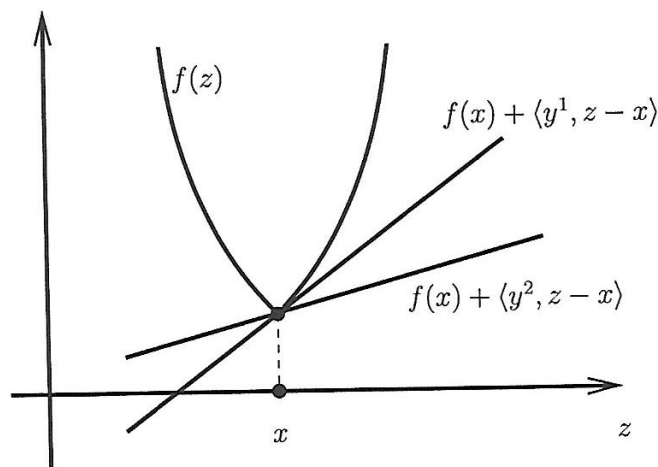


Figura 2.2: y^1 e y^2 são subgradientes de f em x (Fonte: [2]).

Considere o seguinte exemplo do cálculo do subdiferencial no ponto $x = 0$ da função convexa

$$f(x) = \begin{cases} -x, & \text{se } x < 0 \\ x^2, & \text{se } x \geq 0 \end{cases}$$

Em $x = 0$, cada reta $r = ax$ com $-1 \leq a \leq 0$ é uma função suporte de f , isto é, $a \in [-1, 0]$ é um subgradiente da função neste ponto (Figura 2.3).

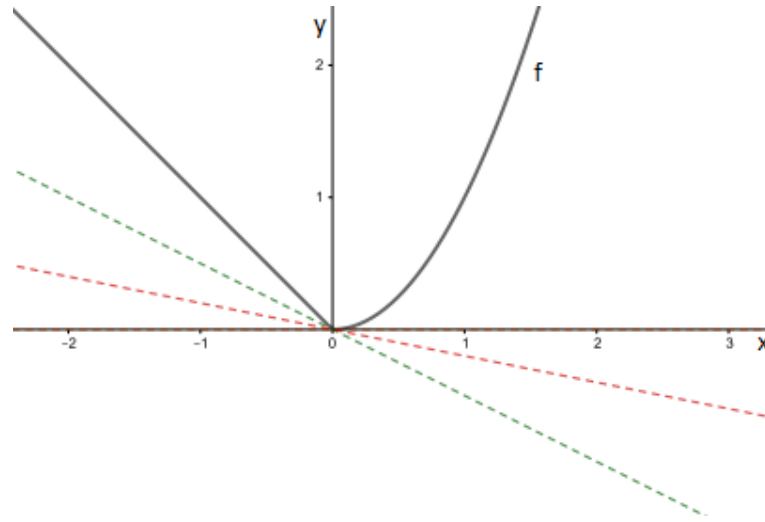


Figura 2.3: Subgradientes no ponto $x = 0$.

Generalizando as observações do exemplo anterior, existe um conjunto infinito de subgradientes se a função convexa f não for diferenciável no ponto. No caso de diferenciabilidade em x , o gradiente $\nabla f(x)$ é o único subgradiente.

O teorema a seguir fornece a condição de otimalidade para funções convexas não necessariamente diferenciáveis.

Teorema 2.2.2. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ convexa. As seguintes condições são equivalentes:*

- (i) f atinge um mínimo global em x ,
- (ii) $0 \in \partial f(x)$.

2.3 NÃO CONVEXIDADE

2.3.1 Função Lipschitz contínua

Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é localmente Lipschitz contínua com constante $L > 0$ em $a \in \mathbb{R}^n$, se existe $\epsilon > 0$ tal que

$$|f(x) - f(y)| \leq L\|x - y\|, \quad \forall x, y \in B(a, \epsilon).$$

O número L representa um limite superior para a “rapidez de variação” da função f .

Na ausência de convexidade da função objetivo f , mas sob hipótese de ser localmente Lipschitz contínua, tem-se a seguinte condição necessária de otimalidade:

“Se $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função localmente Lipschitz contínua e atinge um mínimo local em x^* , então $0 \in \partial f(x^*)$.”

2.4 CONDIÇÕES DE KARUSH-KUHN-TUCKER

Seja o problema geral de otimização na forma

$$\min f(x)$$

$$\text{sujeito a } g_i(x) \leq 0, \quad i = 1, \dots, m$$

com as funções diferenciáveis $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$.

O ponto x do conjunto de restrições do problema é chamado **regular** se os gradientes $\nabla g_i(x)$ são linearmente independentes. O teorema seguinte apresenta as condições de KKT necessárias para um ponto de mínimo local.

Teorema 2.4.1. *Dadas as funções objetivo f e as restrições g_i continuamente diferenciáveis, e seja x um ponto de mínimo local de f . Se x é regular, então existe um vetor $\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$ tal que*

$$\nabla f(x) + \sum_{i=1}^m \lambda_i \nabla g_i(x) = 0$$

$$\lambda_i g_i(x) = 0, \quad i = 1, \dots, m$$

$$g_i(x) \leq 0, \quad i = 1, \dots, m$$

$$\lambda_i \geq 0, \quad i = 1, \dots, m$$

onde os componentes $\lambda_1, \dots, \lambda_m$ são chamados multiplicadores de Lagrange.

2.5 ELEMENTOS DA TEORIA DE DUALIDADE

A teoria de dualidade se baseia em associar ao problema original (primal) um outro problema, denominado dual, que sob certas condições é equivalente ao primal. As relações de dualidade mais fortes são obtidas quando o problema primal é de minimização convexa, mas mesmo em casos bem gerais, a dualidade pode ser muito útil, tanto para as questões teóricas quanto para as técnicas computacionais [18].

Seja um problema em formato geral

$$(P) \quad \min f(x) \text{ sujeito a } h(x) = 0, \quad g(x) \leq 0, \quad x \in X \subset \mathbb{R}^n.$$

onde as funções $f : X \rightarrow \mathbb{R}$, $h : X \rightarrow \mathbb{R}^l$ e $g : X \rightarrow \mathbb{R}^m$ são contínuas. Chama-se (P) de problema primal (problema original). A Lagrangeana deste problema é $L : X \times \mathbb{R}^l \times \mathbb{R}^m \rightarrow \mathbb{R}$, dada por

$$L(x, u, c) = f(x) + \langle c, h(x) \rangle + \langle u, g(x) \rangle. \quad (2.1)$$

Antes de definir o problema dual de um problema não linear, apresenta-se aqui os conceitos do supremo e do ínfimo de uma função.

2.5.1 Supremo e ínfimo

Seja uma função $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$. O **supremo** de f (em X), denotado por $\sup_{x \in X} f(x)$, é definido como o menor valor $s \in \mathbb{R} \cup \{\infty\}$, tal que $s \geq f(x)$ para todo $x \in X$. Analogamente, o *ínfimo* de f (em X), denotado por $\inf_{x \in X} f(x)$, é definido como o maior valor $i \in \mathbb{R} \cup \{-\infty\}$, tal que $i \leq f(x)$ para todo $x \in X$.

Constata-se que o supremo e o ínfimo existem sempre e que estes valores são iguais, respectivamente, ao máximo ou ao mínimo, quando estes existem [c].

Exemplos:

1. Seja $f(x) = 1/x$ e $X = (0, \infty)$. Temos $\sup_{x \in X} f(x) = \infty$, $\inf_{x \in X} f(x) = 0$. Não existe o máximo nem o mínimo de f .
2. Seja $f(x) = x^2$ e $X = [-1, 1]$. Temos $\sup_{x \in X} f(x) = \max_{x \in X} = 1$ e $\inf_{x \in X} f(x) = \min_{x \in X} = 0$.

Pode-se reformular agora o problema (P). Para a função Lagrangeana (2.1), tem-se que

$$\sup_{(u,c) \in \mathbb{R}_+^m \times \mathbb{R}^l} L(x, u, c) = \begin{cases} f(x) & \text{para } h(x) = 0, g(x) \leq 0 \\ \infty & \text{caso contrário.} \end{cases}$$

O problema primal (P) pode ser formalmente escrito como

$$\min_{x \in X} \sup_{(u,c) \in \mathbb{R}_+^m \times \mathbb{R}^l} L(x, u, c). \quad (2.2)$$

O problema dual é obtido trocando a ordem da minimização e da maximização em 2.2, o que resulta no seguinte problema

$$\max_{(u,c) \in \mathbb{R}_+^m \times \mathbb{R}^l} \inf_{x \in X} L(x, u, c). \quad (2.3)$$

Definindo a função dual

$$H(u, c) = \inf_{x \in X} L(x, u, c),$$

o problema dual (2.3) pode ser escrito como

$$\max H(c, u) \text{ sujeito a } (u, c) \in \mathbb{R}_+^m \times \mathbb{R}^l. \quad (2.4)$$

O problema dual tem algumas propriedades interessantes, independentemente da estrutura do problema primal. Em particular, o problema dual sempre consiste na maximização de uma função côncava num conjunto convexo, o que equivale à minimização de uma função convexa neste conjunto. Portanto, toda solução local é global e o conjunto de soluções é convexo (Teorema 1.4.1 em [18], pág. 21).

A fim de exemplificar as definições (2.3) e (2.4), consideremos o seguinte problema primal [1]

$$\begin{aligned} \min \quad & f(x) = (x_1 + 3)^2 + x_2^2 \\ \text{sujeito a} \quad & x_1^2 - x_2 \leq 0 \end{aligned} \quad (2.5)$$

O ponto $x^* = (-1, 1)$ minimiza a função f do problema (2.5), pois, para cada ponto (x_1, x_2) do conjunto viável $M = \{(x_1, x_2) \mid x_1^2 - x_2 \leq 0\}$, $f(x_1, x_2)$ é o quadrado de sua distância ao ponto $(-3, 0)$.

A função Lagrangeana é $L(x_1, x_2, u) = (x_1 + 3)^2 + x_2^2 + u(x_1^2 - x_2)$.

Fixando $u \geq 0$, $L(x_1, x_2, u)$ é convexa e o único ponto de mínimo x^* satisfaz

$$\nabla_x L(x_1^*, x_2^*, u) = \begin{cases} 2(x_1^* + 3) + 2x_1^*u \\ 2x_2^* - u \end{cases} = 0.$$

Do gradiente anterior, temos $x_1^* = -3/(u + 1)$ e $x_2^* = u/2$, logo

$$\begin{aligned} \inf_{x \in \mathbb{R}^2} L(x_1, x_2, u) &= L(x_1^*, x_2^*, u) \\ &= \left(\frac{-3}{u+1} + 3 \right)^2 + \frac{u^2}{4} + u \left(\frac{9}{(u+1)^2} - \frac{u}{2} \right) \\ &= \frac{9u}{u+1} - \frac{u^2}{4} \end{aligned}$$

Assim, o problema dual é

$$\max_{u \geq 0} H(u) = \frac{9u}{u+1} - \frac{u^2}{4}$$

e a função dual $H(u)$ é côncava para valores não negativos de u . Na figura 2.4, observa-se o único ponto ótimo $u^* = 2$ com valor $H(u^*) = 5$.

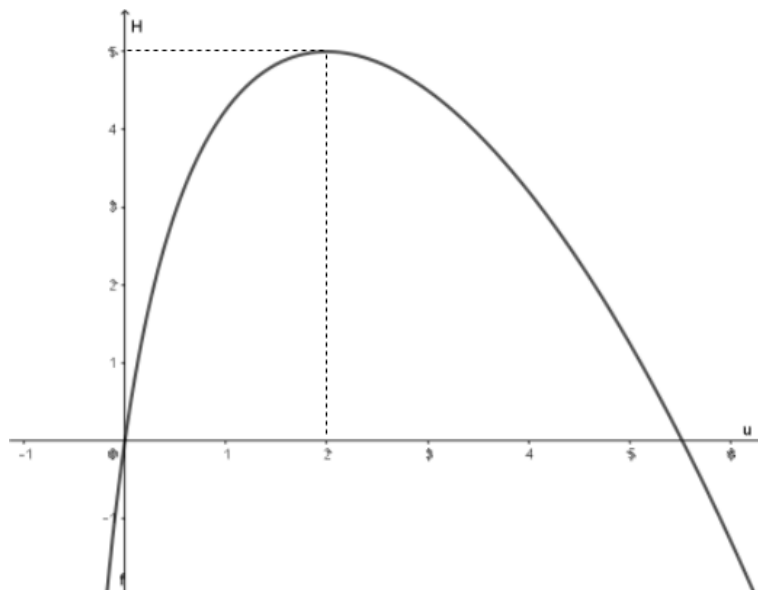


Figura 2.4: Gráfico função dual.

Mesmo nos casos em que não há equivalência entre os problemas primal e dual, a solução, possivelmente mais fácil de se obter do dual, fornece informação útil sobre o primal, conforme pode-se verificar na próxima definição.

2.5.2 Dualidade fraca

Sejam \bar{x} , (\bar{u}, \bar{c}) pontos viáveis dos problemas (2.2) e (2.4), respectivamente. Os valores das funções $f(\bar{x})$ e $H(\bar{u}, \bar{c})$ satisfazem a relação

$$H(\bar{u}, \bar{c}) \leq f(\bar{x}).$$

Demonstração: Veja Teorema 4.28 em [1], pág. 112.

O teorema (2.5.2) tem consequências importantes para procedimentos de resolução: se x^* é uma solução ótima do problema primal, tem-se

$$H(\bar{u}, \bar{c}) \leq f(x^*) \leq f(\bar{x}),$$

ou seja, um ponto viável do problema dual sempre fornece uma cota inferior para o valor ótimo do primal. De modo análogo, para (u^*, c^*) uma solução ótima do dual, $f(\bar{x})$ é uma cota superior para o valor ótimo do dual

$$H(\bar{u}, \bar{c}) \leq H(u^*, c^*) \leq f(\bar{x}).$$

A diferença $f(x^*) - H(u^*, c^*)$ é chamada “brecha de dualidade” (*duality gap*). Em geral, para garantir a igualdade deste valores, precisa-se de algumas hipóteses adicionais, tipicamente de convexidade. Mas, mesmo quando há brecha de dualidade, a relação entre os dois pode ser útil. Por exemplo, um algoritmo que produz iterativamente soluções viáveis dos problemas primal e dual pode determinar quando a diferença

$$f(\bar{x}) - H(\bar{u}, \bar{c})$$

é suficientemente pequena.

3 MÉTODOS DE OTIMIZAÇÃO

Apresenta-se aqui algumas noções sobre métodos (algoritmos) de Otimização, particularmente aqueles dedicados a problemas não diferenciáveis e convexos. Ressalta-se que essas técnicas são relevantes inclusive para a resolução de algumas classes de problemas que são, em princípio, diferenciáveis e/ou não convexos [?]. Como descrito no capítulo anterior, a resolução do problema dual pode fornecer informação muito útil sobre o problema primal e, sob certas condições, até resultar na resolução exata do primal [2].

A passagem dos métodos diferenciáveis para os não diferenciáveis acarreta dificuldades, tais como a direção de busca d não ser necessariamente de descida para a função f num dado ponto x , ou seja, não satisfazendo a condição para uma direção de descida

$$d^T s < 0,$$

em que s é um subgradiente pertencente ao subdiferencial $\partial f(x)$. Necessitaríamos conhecer todo o conjunto $\partial f(x)$ para cada x , i.e., todos os subgradientes, o que inviabiliza a maioria das aplicações práticas.

Outra questão importante consiste na escolha de critérios de parada. Na otimização diferenciável e irrestrita, uma regra de parada razoável é dada pela condição $\|\nabla f(x^k)\| \leq \epsilon$, em que $\epsilon > 0$ é suficientemente pequeno. Esta regra não pode ser aplicada a problemas não diferenciáveis como mostra o seguinte exemplo:

Considere a função valor absoluto $f(x) = |x|$, $x \in \mathbb{R}$. Para qualquer $x^k \neq 0$, não importa quão próximo x^k esteja do ponto de mínimo $x^* = 0$ da função, tem-se $\|\nabla f(x^k)\| = 1$.

Neste sentido, faremos a seguir uma breve exposição de alguns métodos de otimização não diferenciável, que são considerados clássicos e influenciaram o desenvolvimento dos algoritmos desta pesquisa.

3.1 MÉTODO DE PLANOS DE CORTE

Considere o problema inicial

$$(P) \quad \min f(x) \text{ sujeito a } x \in D,$$

onde $D \subset \mathbb{R}^n$ é um conjunto convexo e fechado, e $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função convexa.

Cada iteração deste método consiste na minimização do máximo de funções afins, acumuladas nas iterações anteriores, que aproximam a função objetivo inferiormente

$$f(x) \cong \max\{f(x^i) + s^i(x - x^i)\}, \quad s^i \in \partial f(x^i), \quad i = 1, 2, \dots, k$$

O problema (P) é substituído por

$$(P1) \begin{cases} \min \max\{f(x^i) + s^i(x - x^i)\} \\ x \in \mathbb{R}^n \end{cases}$$

que por sua vez é equivalente a

$$(P2) \begin{cases} \min z \\ \text{s.a. } f(x^i) + s^i(x - x^i) \leq z \end{cases}$$

O problema $(P2)$ é linear e sua solução (caso exista) é uma aproximação da solução do problema original e se torna mais precisa à medida que mais planos de corte são acrescentados (Figura 3.1).

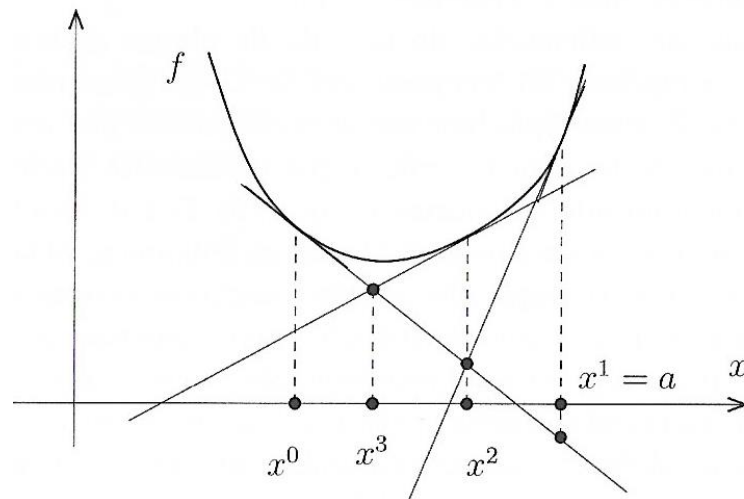


Figura 3.1: Três iterações do método de planos de corte (Fonte: [?]).

A vantagem mais importante deste método em relação aos métodos de subgradiente é a possibilidade de construir uma regra de parada com base na diferença entre os valores das funções do problema (P) e $(P2)$.

Entre as deficiências do método, está o número crescente de funções afins acumuladas que dificultam a resolução do problema, além da falta de garantia de decréscimo da função objetivo.

3.2 MÉTODOS DE FEIXES

Os métodos de feixes se baseiam na ideia de aproximação linear por partes, semelhante ao método de planos de corte, mas buscando-se evitar o acúmulo de funções afins que causam mal condicionamento.

A partir das informações obtidas nas iterações anteriores acumulando-os no chamado “feixe de informações”, esses métodos utilizam tanto as características dos métodos de descida quanto de planos de corte, garantindo assim o decréscimo da função objetivo e a estabilização.

Uma outra característica desses métodos é a utilização do conceito de passo sério e passo nulo. Calcula-se uma direção d^k e o ponto

$$x^{k+1} = x^k + td^k, \quad t > 0.$$

Se $f(x^{k+1}) \leq f(x^k) - \epsilon$, isto é, o passo resultar em um decréscimo suficiente da função objetivo, toma-se então x^{k+1} (passo sério). Senão, $x^{k+1} = x^k$ (passo nulo).

Para determinação da direção de busca, consideremos o seguinte método básico [1]:

- (1) Escolha $x^0 \in \mathbb{R}^n$ e faça $k = 0$;
- (2) Calcule $d^k = -a^k/|a^k|$, onde a^k é solução do problema $\min_{a \in \partial f(x^k)} |a|$;
- (3) Se $0 \in \partial f(x^k)$, pare. Senão, continue com o passo (4);
- (4) Determine uma solução t_k do problema $\min_{\lambda \in \mathbb{R}_+} f(x^k + \lambda d^k)$;
- (5) Faça $x^{k+1} = x^k + t_k d^k$, $k = k + 1$ e continue com o passo (2).

O critério de parada no passo (3) baseia-se no fato de que x^k é um ponto de mínimo global de f se, e somente se, $0 \in \partial f(x^k)$.

Para um estudo mais detalhado, consulte [19], [20] e [21].

4 ALGORITMOS

Numa linha histórica, o método IED possui características de outros algoritmos que o antecederam. A seguir serão apresentadas essas principais características que influenciaram o desenvolvimento do IED.

4.1 A DIREÇÃO DE BUSCA DO FDIPA

O método *Feasible Direction Interior Point Algorithm* (FDIPA), proposto em 1998 por Herkovits [22], aplica-se à otimização não linear diferenciável e converge globalmente para pontos de Karush-Khun-Tucker (KKT). Não é necessária a solução de problemas quadráticos, nem se trata de um método de penalidades ou de barreiras. O método requer somente a resolução de dois sistemas de equações lineares com a mesma matriz principal em cada iteração, seguida de uma busca linear inexata.

Considera-se o seguinte problema de otimização não linear diferenciável com restrições de desigualdade:

$$(P_{4.1}) \begin{cases} \min f(x) \\ \text{sujeito a } g(x) \leq 0, \quad i = 1, 2, \dots, m \end{cases}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ são funções diferenciáveis.

Se x^* é um ponto regular e mínimo local para o problema $(P_{4.1})$, então existe $\lambda^* \in \mathbb{R}^m$ tal que as condições de otimalidade de KKT podem ser escritas

$$\nabla f(x^*) + \nabla g(x^*)\lambda^* = 0 \tag{4.1}$$

$$G(x^*)\lambda^* = 0 \tag{4.2}$$

$$\lambda^* \geq 0 \tag{4.3}$$

$$g(x^*) \leq 0 \tag{4.4}$$

onde $G(x)$ é uma matriz diagonal com $G_{ii}(x) \equiv g_i(x)$.

O ponto x tal que $g(x) \leq 0$ é chamando “Ponto primal Viável”, e $\lambda \geq 0$, “Ponto Dual Viável”. Dado um par viável inicial (x^0, λ^0) , o algoritmo gera uma sequência de pontos

KKT a partir da resolução de um sistema formado pelas equações lineares (4.1) e (4.2) nas variáveis primal e dual (x, λ) .

Uma iteração de Newton para resolver o sistema linear das equações (4.1) and (4.2) pode ser escrito como

$$\begin{pmatrix} B^k & \nabla g(x^k) \\ \Lambda^k \nabla g(x^k)^t & G(x^k) \end{pmatrix} \begin{pmatrix} x - x^k \\ \lambda - \lambda^k \end{pmatrix} = - \begin{pmatrix} \nabla f(x^k) + \nabla g(x^k) \lambda^k \\ G(x^k) \lambda^k \end{pmatrix} \quad (4.5)$$

onde (x^k, λ^k) é o ponto corrente de iteração k , Λ é uma matriz diagonal com $\Lambda_{ii} \equiv \lambda_i$ e $B^k \equiv \nabla^2 f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla^2 g_i(x^k)$ é a hessiana da função Lagrangeana $L(x, \lambda) = f(x) + \lambda^T g(x)$ ou uma aproximação por alguma técnica quasi-Newton ou pela matriz identidade. Contudo, B^k deve ser definida positiva para garantir a convergência global do algoritmo.

Definindo $d = x - x^k$, pode-se reescrever o sistema (4.5) do seguinte modo:

$$B^k d + \nabla g(x^k) \lambda = -\nabla f(x^k) \quad (4.6)$$

$$\Lambda^k \nabla g(x^k)^T d + G(x^k) \lambda = 0. \quad (4.7)$$

A solução (d_1^k, λ_1^k) do sistema (4.6)-(4.7) fornece uma direção de descida d_1^k , demonstrada em [13]. Entretanto, d_1^k pode não ser uma direção viável pois, quando alguma restrição se aproxima de zero, d_1^k tende a uma direção tangente ao conjunto viável, fato decorrente de $\nabla g_i(x^k)^T d_1^k = 0$ na equação (4.7).

Para contornar essa questão, define-se um novo sistema em d e $\bar{\lambda}$ com adição de um vetor negativo $-\rho^k \Lambda^k$, $\rho^k > 0$, no lado direito de (4.7),

$$B^k d + \nabla g(x^k) \bar{\lambda} = -\nabla f(x^k) \quad (4.8)$$

$$\Lambda^k \nabla g(x^k)^T d + G(x^k) \bar{\lambda} = -\rho^k \Lambda^k. \quad (4.9)$$

Nesta nova situação, (4.9) é equivalente a $\lambda_i^k \nabla g_i(x^k)^T d^k + g_i(x^k) \bar{\lambda}^k = -\rho^k \lambda_i^k$, $i = 1, 2, \dots, m$. Consequentemente, se $g_i(x^k) = 0$, a última equação se reduz a $\nabla g_i(x^k)^T d^k = -\rho^k < 0$ e d^k é uma direção viável.

A adição do termo negativo $-\rho^k \lambda^k$ produz uma deflexão em d^k , proporcional à ρ , na direção do interior da região viável.

Para garantir agora que d seja uma direção de descida, ρ deve ser escolhido convenientemente. Observa-se que d e $\bar{\lambda}$ podem ser obtidos da resolução dos seguintes sistemas

$$B^k d_1 + \nabla g(x^k) \lambda_1 = -\nabla f(x^k) \quad (4.10)$$

$$\Lambda^k \nabla g(x^k)^T d_1 + G(x^k) \lambda_1 = 0 \quad (4.11)$$

e

$$B^k d_2 + \nabla g(x^k) \lambda_2 = 0 \quad (4.12)$$

$$\Lambda^k \nabla g(x^k)^T d_2 + G(x^k) \lambda_2 = -\Lambda^k \quad (4.13)$$

de modo que, definindo $d^k = d_1^k + \rho^k d_2^k$, tem-se $(d^k)^T \nabla f(x^k) = (d_1^k)^T \nabla f(x^k) + \rho^k (d_2^k)^T \nabla f(x^k)$. Uma vez que $(d_1^k)^T \nabla f(x^k) < 0$, se $(d_2^k)^T \nabla f(x^k) \leq 0$ então $(d^k)^T \nabla f(x^k) < 0$, $\forall \rho^k > 0$.

Portanto, ρ^k deve ser convenientemente escolhido no caso em que $(d_2^k)^T \nabla f(x^k) > 0$. Impõe-se $(d^k)^T \nabla f(x^k) \leq \xi (d_1^k)^T \nabla f(x^k)$ com $\xi \in (0, 1)$, obtendo-se $(d_1^k)^T \nabla f(x^k) + \rho^k (d_2^k)^T \nabla f(x^k) \leq \xi (d_1^k)^T \nabla f(x^k)$ ou $\rho^k \leq \frac{(\xi-1)(d_1^k)^T \nabla f(x^k)}{(d_2^k)^T \nabla f(x^k)}$.

Assim, se ρ^k for escolhido de modo a satisfazer a desigualdade anterior, $d^k = d_1^k + \rho^k d_2^k$ será uma direção viável e de descida para f (Figura 4.1).

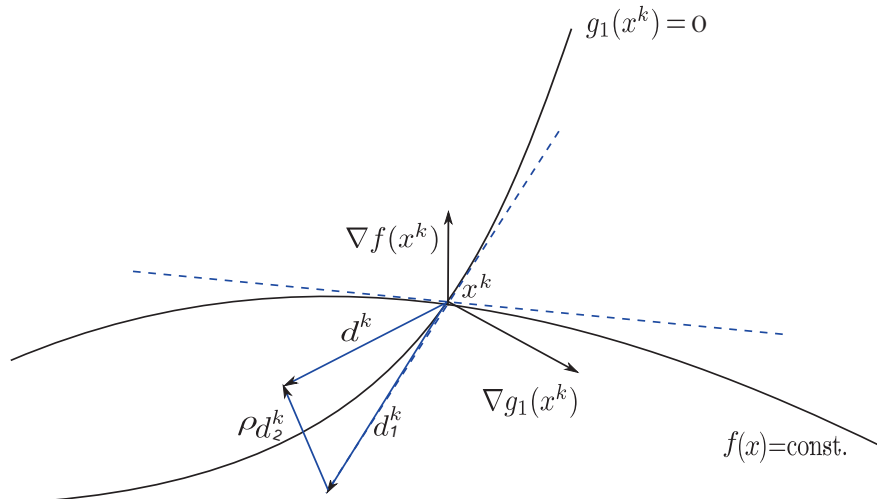


Figura 4.1: Direção de busca do FDIPA (Fonte: [3]).

Embora seja uma técnica para problemas diferenciáveis, os sistemas internos do

FDIPA são amplamente utilizados na determinação de direções viáveis nos métodos não diferenciáveis apresentados a seguir.

4.2 NFDA E A APROXIMAÇÃO DO EPÍGRAFO

A partir do FDIPA, Freire [23] desenvolveu o algoritmo *Nonsmooth Feasible Directions Algorithm* (NFDA) para lidar com problemas não diferenciáveis, convexos e irrestritos da forma:

$$(P_{4.2}) \begin{cases} \min f(x) \\ x \in \mathbb{R}^n, \end{cases}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função convexa não necessariamente diferenciável.

O problema original e irrestrito $(P_{4.2})$ é substituído por um outro equivalente onde se quer minimizar uma função linear $z \in \mathbb{R}$ sujeita a uma restrição não diferenciável de desigualdade, isto é,

$$(PE_{4.2}) \begin{cases} \min z \\ \text{sujeito a } f(x) \leq z \\ (x, z) \in \mathbb{R}^n \times \mathbb{R}. \end{cases}$$

O NFDA gera uma sequência de pontos (x^k, z^k) no interior do $\text{epi}(f)$, satisfazendo

$$z^{k+1} < z^k \quad \text{e} \quad f(x^k) < z^k, \quad \forall k.$$

Com a utilização das ideias do método de planos de corte de Kelley [4], o algoritmo constrói aproximações lineares do problema $(PE_{4.2})$ da seguinte forma:

$$g_i(x, z) = f(y_i) + s_i^T(x - y_i) - z, \quad i = 0, 1, \dots, l$$

onde $y_i \in \mathbb{R}^n$ são pontos auxiliares, $s_i \in \partial f(y_i)$ são subgradientes nesses pontos e l representa o número de planos corrente. Define-se então o vetor

$$\bar{g}_l(x, z) = [g_0(x, z), \dots, g_l(x, z)]^T,$$

e o problema auxiliar corrente

$$(PA_{4.2}) \begin{cases} \min z \\ \text{sujeito a } \bar{g}_l(x, z) \leq 0 \\ (x, z) \in \mathbb{R}^n \times \mathbb{R}. \end{cases}$$

O algoritmo NFDA não resolve o problema $(PA_{4.2})$, mas utiliza os sistemas internos do FDIPA para determinar uma direção de busca d_l^k para $(PA_{4.2})$, que pode ser calculada ainda que este não tenha mínimo finito.

O maior comprimento de passo é dado por

$$t = \max\{t | \bar{g}_l^k((x^k, z^k) + td_l^k) \leq 0\}.$$

Caso t não seja finito, considera-se

$$t^k = \min\{t, T\}, \quad T > 0 \text{ é um parâmetro pré-definido.}$$

Então, um novo iterado (x^{k+1}, z^{k+1}) é definido de acordo com o seguinte procedimento: o algoritmo produz pontos auxiliares

$$(y^k, \omega^k) = (x^k, z^k) + \mu t^k d_l^k, \quad \mu \in (0, 1)$$

e quando um ponto auxiliar estiver no interior do epígrafo de f , diz-se que foi realizado um passo **sério** e este será o novo iterado.

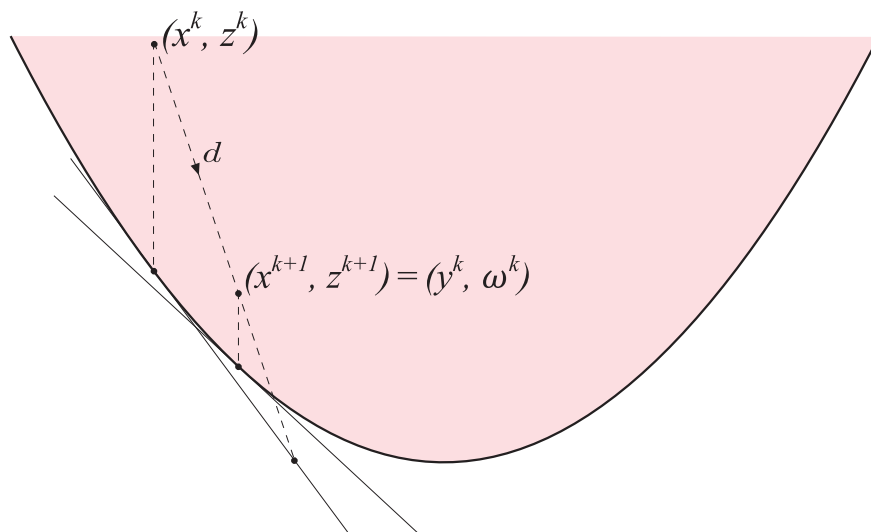


Figura 4.2: Passo sério do NFDA (Fonte: o autor).

Caso contrário, continua-se com o mesmo ponto corrente e diz-se que foi realizado um passo **nulo**. Em ambos os casos, um novo plano de corte é adicionado e o procedimento se repete até a obtenção de um passo sério.

O método termina quando se encontra uma direção d^k que satisfaça $\|d^k\| < \epsilon$, onde $\epsilon > 0$ é um parâmetro pré-estabelecido.

Nas duas seções anteriores apresentamos as formas de calcular a direção de busca do método FDIPA e a de aproximar o epígrafo da função objetivo por planos de corte do método NFDA. O próximo algoritmo é uma extensão do NFDA para o caso não convexo, inspirado nessas ideias até aqui vistas.

4.3 NFDNA E A EXTENSÃO PARA O CASO NÃO CONVEXO

O algoritmo *Nonsmooth Feasible Direction Nonconvex Algorithm* (NFDNA), proposto por Tanaka [3] para resolver problemas de otimização irrestrita com funções não necessariamente convexas ou diferenciáveis, combina características dos métodos apresentados anteriormente nas seções 4.1 e 4.2.

Considera-se o seguinte problema

$$(P_{4.3}) \begin{cases} \min f(x) \\ x \in \mathbb{R}^n, \end{cases}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função localmente Lipschitz contínua.

A teoria dos gradientes generalizados (Clarke [24]) mostrou-se um bom instrumento quando se lida com funções não diferenciáveis e não convexas.

Embora não seja uma tarefa simples generalizar o método de planos de corte para o caso não convexo, são uteis algumas ideias utilizadas no caso convexo. Por exemplo, nos métodos de feixe, a maneira mais comum de lidar com as dificuldades causadas pela não convexidade é usar o chamado subgradiente localmente mensurado ao invés do erro de linearização ([25], [20]).

A não convexidade traz também características adicionais para o problema, uma das

quais é que a função objetivo pode ter vários mínimos e máximos locais. Como em todo método não global, a convergência é provada para um ponto estacionário, isto é, um ponto que satisfaz a condição necessária de otimalidade $0 \in \partial f(x)$.

Como no NFDA, substitui-se o problema original ($P_{4.3}$) do NFDNA por um problema equivalente linear e sujeito a uma restrição não diferenciável. Constrói-se então uma sequência de problemas auxiliares lineares que são estabelecidos à medida em que se aproxima o epígrafo da função f por planos de corte. Para cada problema auxiliar, que também não será resolvido, o algoritmo gera uma direção de descida viável d^k de modo igual ao FDIPA.

Os próximos passos apresentam as novidades do NFDNA após o cálculo do tamanho do passo e de um novo ponto auxiliar (y^k, ω^k) de acordo com o que se segue:

- Quando o novo ponto é viável em relação ao problema auxiliar e de descida para a função objetivo f , a solução é atualizada, isto é, $(x^{k+1}, z^{k+1}) = (y^k, \omega^k)$ e um **passo sério de descida viável** foi realizado.
- Se o novo ponto é viável em relação ao problema auxiliar, mas não é de descida para f , considera-se que o ponto corrente (x^k, z^k) está longe da fronteira do epígrafo de f . Neste caso, troca-se a direção calculada pelo FDIPA pela direção de máxima descida (método do gradiente) da função linear z para se obter um ponto que esteja próximo da fronteira do epígrafo. Deste modo, tem-se $f(x^{k+1}) = f(x^k)$ e a próxima direção de busca gerada pelo FDIPA será de descida para f . Este é o **passo sério de máxima descida**.
- No caso de um passo sério, o algoritmo limpa todas as informações armazenadas até então dos planos de corte.
- Porém, se não ocorre nenhuma das alternativas anteriores, realiza-se o equivalente ao **passo nulo** dos métodos de feixe, ou seja, não se atualiza a solução fazendo-se $(x^{k+1}, z^{k+1}) = (x^k, z^k)$. Um novo plano de corte é calculado e uma nova direção de descida é obtida usando os sistemas do FDIPA. Em seguida, o algoritmo recomeça seu procedimento novamente.

Devido a não convexidade, pode ocorrer de o novo plano de corte tornar o ponto de iteração corrente (x^k, z^k) inviável (Figura 4.3). Nesta situação, ignora-se o plano de corte,

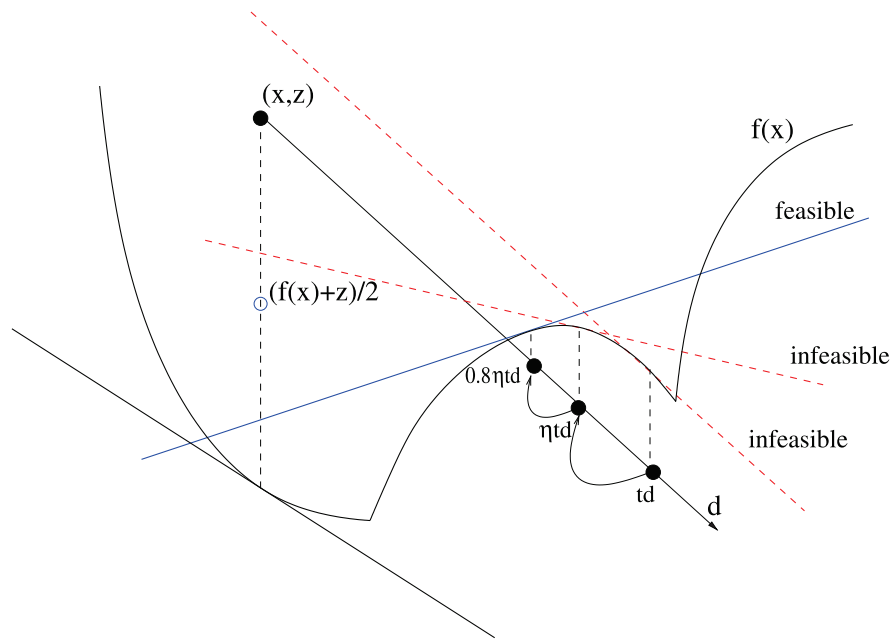


Figura 4.3: Cálculo de um plano viável (Fonte: [4]).

retorna-se ao longo da direção de busca e um novo plano é calculado. Este procedimento de retorno continua até que o ponto corrente seja viável com respeito ao plano de corte. De fato, para aumentar a eficiência do método, verifica-se a viabilidade do ponto $(x^k, \frac{f(x^k)+z^k}{2})$.

4.3.1 Algoritmo NFDNA

A seguir têm-se os passos do NFDNA, mantida a notação original do autor [4].

Parâmetros: Escolha uma tolerância $\varepsilon > 0$. Selecione os parâmetros de controle $\rho > 0$ e $v \in (0, 1)$ para o limite da deflexão. Selecione também os multiplicadores $\mu \in (0, 1)$ e $\eta \in (1/2, 1)$ para o tamanho do passo e o tamanho máximo $t_{\max} > 0$.

Dados iniciais: Seja o contador de iterações $k = 1$ e o contador do número de planos de corte $l = 1$. Escolha um ponto inicial estritamente viável $(x^1, z^1) \in \text{int}(\text{epi}f)$, um vetor positivo inicial $\lambda^1 \in \mathbb{R}^l$ e uma matriz simétrica positiva $B^1 \in \mathbb{R}^{(n+1) \times (n+1)}$. Seja $y_1^1 = x^1$. Calcule $f(x^1)$.

Passo 1: (Plano de corte inicial) Calcule $s_1^k \in \partial f(x^k)$ e o primeiro plano de corte $g_1^k(x^k, z^k) = f(x^k) - z^k \in \mathbb{R}$.

Faça $\nabla g_1^k(x^k, z^k) = (s_1^k, -1) \in \mathbb{R}^{n+1}$.

Defina $\bar{g}_1^k(x^k, z^k) = [g_1^k(x^k, z^k)] \in \mathbb{R}$ e $\nabla \bar{g}_1^k(x^k, z^k) = [\nabla g_1^k(x^k, z^k)]$.

Passo 2: (Cálculo da direção) Calcule $d_l^k = (d_x^k, d_z^k) \in \mathbb{R}^{n+1}$, uma direção de descida viável para (PA_l) :

1. (Direção de descida) Determine $d_{1,l}^k \in \mathbb{R}^{n+1}$ e $\lambda_{1,l}^k \in \mathbb{R}^l$ resolvendo o sistema linear

$$B^k d_{1,l}^k + \nabla \bar{g}_1^k(x^k, z^k) \lambda_{1,l}^k = -e_z \quad (4.14)$$

$$\Lambda_l^k [\nabla \bar{g}_1^k(x^k, z^k)]^t d_{1,l}^k + \bar{G}_l^k(x^k, z^k) \lambda_{1,l}^k = 0, \quad (4.15)$$

onde $e_z = [0, 0, \dots, 0, 1]^t \in \mathbb{R}^{n+1}$, $\Lambda_l^k = \text{diag}[\lambda_1^k, \dots, \lambda_l^k]$, e

$\bar{G}_l^k(x^k, z^k) = \text{diag}[g_1^k(x^k, z^k), \dots, g_l^k(x^k, z^k)]$.

2. (Direção viável) Determine os valores $d_{2,l}^k \in \mathbb{R}^{n+1}$ e $\lambda_{2,l}^k \in \mathbb{R}^l$ resolvendo o sistema linear

$$B^k d_{2,l}^k + \nabla \bar{g}_l^k(x^k, z^k) \lambda_{2,l}^k = 0 \quad (4.16)$$

$$\Lambda_l^k [\nabla \bar{g}_l^k(x^k, z^k)]^t d_{2,l}^k + \bar{G}_l^k(x^k, z^k) \lambda_{2,l}^k = -\lambda_l^k. \quad (4.17)$$

Se $e_z^t d_{2,l}^k > 0$, então

$$\rho = \min \left\{ \varrho \|d_{1,l}^k\|^2, \frac{(v-1)e_z^t d_{1,l}^k}{e_z^t d_{2,l}^k} \right\}. \quad (4.18)$$

Caso contrário,

$$\rho = \varrho \|d_{1,l}^k\|^2. \quad (4.19)$$

3. (Direção de descida viável) Calcule a direção de busca

$$d_l^k = d_{1,l}^k + \rho d_{2,l}^k. \quad (4.20)$$

Passo 3: (Tamanho de passo e atualização da solução) Calcule o tamanho de passo $t^k = \min \{t_{\max}, \max \{t | \bar{g}_l^k((x^k, z^k) + t d_l^k) \leq 0\}\}$.

Se $\|d_l^k\| \leq \varepsilon$ and $t^k < t_{\max}$, então pare com (x^k, z^k) sendo a solução final. Caso contrário, defina $(y_{l+1}^k, \omega_{l+1}^k) = (x^k, z^k) + \mu t^k d_l^k$ e calcule o valor de $f(y_{l+1}^k)$.

Se $\omega_{l+1}^k \leq f(y_{l+1}^k)$, o passo não é sério: siga para o passo 6. Caso contrário, faça $d^k = d_l^k$, $d_1^k = d_{1,l}^k$, $d_2^k = d_{2,l}^k$, $\lambda_1^k = \lambda_{1,l}^k$, e $\lambda_2^k = \lambda_{2,l}^k$. Se $f(x^k) \geq f(y_{l+1}^k)$ siga para o passo 4. Senão, siga para o passo 5.

Passo 4: (Passo sério de descida viável) Faça $(x^{k+1}, z^{k+1}) = (y_{l+1}^k, \omega_{l+1}^k)$ e $f(x^{k+1}) = f(y_{l+1}^k)$. Apague todos os planos de corte e atualize B^k para B^{k+1} e λ^k para λ^{k+1} . Faça $k = k + 1$, $l = 1$ e siga para o passo 1.

Passo 5: (Passo sério de máxima descida) Defina $(x^{k+1}, z^{k+1}) = (x^k, z^k) - \mu(z^k - F(x^k))e_z$ and $f(x^{k+1}) = f(x^k)$. Apague todos os planos de corte e atualize B^k para B^{k+1} e λ^k para λ^{k+1} . Faça $k = k + 1$, $l = 1$ e siga para o passo 1.

Passo 6: (Passo nulo)

1. (Erro de linearização) Calcule $s_{l+1}^k \in \partial f(y_{l+1}^k)$ e um erro de linearização $\alpha = f(x^k) - f(y_{l+1}^k) - (s_{l+1}^k)^t(x^k - y_{l+1}^k)$.
2. (Procedimento de retorno) Se $\alpha < g_1^k(x^k, z^k)/2$ retorne ao longo do vetor d^k até que um "ponto viável" seja encontrado, ou seja, faça $(y_{l+1}^k, \omega_{l+1}^k) = (x^k, z^k) + \eta \mu t^k d^k$, $\eta = 0.8\eta$ e siga para o passo 6(1).
3. (Planos de corte para passos nulos) Calcule o novo plano de corte e seus gradientes $g_{l+1}^k(x^k, z^k) = -\alpha + F(x^k) - z^k$ e $\nabla g_{l+1}^k(x^k, z^k) = (s_{l+1}^k, -1)$.
Defina $\bar{g}_{l+1}^k(x^k, z^k) = [g_1^k(x^k, z^k), \dots, g_l^k(x^k, z^k), g_{l+1}^k(x^k, z^k)]^t \in \mathbb{R}^{l+1}$ e
 $\nabla \bar{g}_{l+1}^k(x^k, z^k) = [\nabla g_1^k(x^k, z^k), \dots, \nabla g_l^k(x^k, z^k), \nabla g_{l+1}^k(x^k, z^k)] \in \mathbb{R}^{(n+1) \times (n+1)}$.
Faça $l = l + 1$ e siga para o passo 2.

4.4 O MÉTODO IED E O TRATAMENTO DAS RESTRIÇÕES VIA DUALIDADE LAGRANGIANA

Nesta seção apresentamos o método IED, desenvolvido por Burachik, Freire e Kaya [5], para resolver problemas de otimização não convexos, não diferenciáveis e com restrições. Em sua versão original, o IED combina características de outros dois algoritmos: NFDA, discutido na seção 4.2, e *Deflected Subgradient Method* (DSG), proposto por Burachik e Kaya [26].

Considera-se o seguinte problema

$$(P_{4.4}) \begin{cases} \min f(x) \\ \text{sujeito a } g(x) = 0 \\ x \in X \end{cases}$$

onde X é um subconjunto compacto do \mathbb{R}^n e as funções $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ são contínuas.

Embora o método tenha sido concebido para problemas com restrições de igualdade, restrições de desigualdade $g(x) \leq 0$ podem ser incorporadas ao problema $(P_{4.4})$ através do operador não diferenciável $\max\{g(x), 0\} = 0$.

A técnica utilizada para resolver $(P_{4.4})$ é a dualidade Lagrangeana, em que a função Lagrangeana aumentada $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_+ \rightarrow \mathbb{R}$ é a soma da função objetivo f com mais dois termos: uma função de penalidade σ mais um termo linear, conforme a seguir

$$L(x, (u, c)) = f(x) + c\sigma(g(x)) + \langle Au, g(x) \rangle, \quad (4.21)$$

onde $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $c \in \mathbb{R}_+ = [0, \infty)$, $A \in \mathbb{R}^{m \times m}$ é uma matriz real simétrica, $\sigma : \mathbb{R}^m \rightarrow \mathbb{R}$ é uma função contínua tal que $\sigma(x) > 0$, $\forall x \in \mathbb{R}^m \setminus \{0\}$ e $\sigma(0) = 0$.

Propriedades de ponto de sela e dualidade forte são válidas para funções Lagrangeanas do tipo (4.21) ([26],[27],[28]).

Para formular o problema dual, define-se a função dual (côncava) correspondente $H : \mathbb{R}^m \times \mathbb{R}_+ \rightarrow \mathbb{R}$

$$H(u, c) = \min_{x \in X} [f(x) + c\sigma(g(x)) + \langle Au, g(x) \rangle]. \quad (4.22)$$

A minimização da função Lagrangeana em (4.22) será aqui chamada de **subproblema**, sendo considerada uma etapa importante no método, e sobre o qual será apresentada, no próximo capítulo, uma forma diferente da original para resolvê-lo.

Portanto, o problema dual associado ao primal (4.4) é dado por

$$(D) \max_{(u,c) \in \mathbb{R}^m \times \mathbb{R}_+} H(u, c),$$

o qual é equivalente a

$$(D^*) \min_{(u,c) \in \mathbb{R}^m \times \mathbb{R}_+} q(u, c), \quad q = -H, \text{ por simplificação.}$$

Com esta equivalência, tem-se em (D^*) um problema irrestrito de otimização convexa.

Assumida a possibilidade de resolução do subproblema (4.22), dado um ponto dual (u, c) , tem-se agora um ponto primal x que pertence ao conjunto

$$X(u, c) = \underset{x \in X}{\text{Arg min}} [f(x) + c\sigma(g(x)) + \langle Au, g(x) \rangle]. \quad (4.23)$$

Em seguida, um subgradiente s pode ser calculado no ponto $((u, c), q(u, c))$ de acordo com o seguinte resultado:

Teorema 4.4.1. *Dados a função dual q e um ponto $(u, c) \in \mathbb{R}^m \times \mathbb{R}_+$. Se $x \in X(u, c)$, então $s = (-Ag(x), -\sigma(g(x))) \in \partial q(u, c)$. [26]*

Generalizando, a partir de um ponto corrente $(u^k, c^k) = v^k$ no espaço dual, um iterado primal x^k é obtido pela regra

$$x^k \in X(v^k), \quad (4.24)$$

Com este iterado x^k , encontra-se uma direção de busca que fornece um novo ponto dual (u^{k+1}, c^{k+1}) , que por sua vez pode estar ou não no interior do epígrafo da função dual. Se $(u^{k+1}, c^{k+1}) \in \text{int}(\text{epi}(q))$, então tem-se o chamado passo sério e este será o novo iterado. Caso (u^{k+1}, c^{k+1}) não esteja no interior do epígrafo, tem-se o passo nulo e emprega-se o passo do algoritmo DSG para gerar o novo iterado. Se o ponto x^k satisfizer a restrição $g(x^k) = 0$, então este é a solução do problema original (primal), tendo-se portanto um critério de parada para o algoritmo, que se baseia no teorema a seguir:

Teorema 4.4.2. *Seja $x \in X(v)$ para algum $v \in \mathbb{R}^m \times \mathbb{R}_+$. Então x é uma solução do problema primal e v é uma solução do problema dual se, e somente se, $g(x) = 0$. [26]*

Na próxima secção, apresentamos mais detalhes referentes ao cálculo da direção de busca do IED.

4.4.1 O cone de direções

Igualmente ao método NFDA, o problema dual (D^*) é substituído pelo problema de minimização de uma função linear z , sujeita a uma restrição de desigualdade não

necessariamente diferenciável

$$(ED^*) \begin{cases} \min \psi(u, c, z) = z \\ \text{sujeito a } q(u, c) \leq z. \end{cases}$$

A partir de um ponto corrente $(u^k, c^k) = v^k$ no espaço dual, um iterado primal x^k é calculado, tal que $g(x^k) \neq 0$. Utiliza-se x^k para encontrar um subgradiente s^k que define um plano de corte $h^k(v) = q(v^k) + \langle s^k, v - v^k \rangle$, suporte à restrição do problema (ED^*) no ponto $(v^k, q(v^k))$.

Então, uma aproximação do problema (ED^*) é dada pelo problema auxiliar

$$(AP_k) \begin{cases} \min \psi(v, z) = z \\ \text{sujeito a } g^k(v, z) \leq 0, \end{cases}$$

em que $g^k : \mathbb{R}^m \times \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$ é dada por $g^k(v, z) = h^k(v) - z$.

Com a mesma técnica do FDIPA, resolve-se dois sistemas lineares - mostrados no algoritmo à frente - e uma direção de busca $d^k = d_1^k + \rho d_2^k$ é calculada, na qual o parâmetro ρ deve ser escolhido de modo que d^k seja uma direção de descida para ψ . Impondo a condição adicional $(d^k)^T \nabla g(v^k, z^k) < 0$, garante-se também que d^k é de descida para a função dual q . Em outras palavras, o IED encontra uma direção que pertence ao interior do cone em cada iteração

$$K(v^k, z^k) = \{d \in \mathbb{R}^{m+1} \times \mathbb{R} \mid d^T \nabla g(v^k, z^k) < 0 \text{ e } d^T \nabla \psi(v^k, z^k) < 0\}.$$

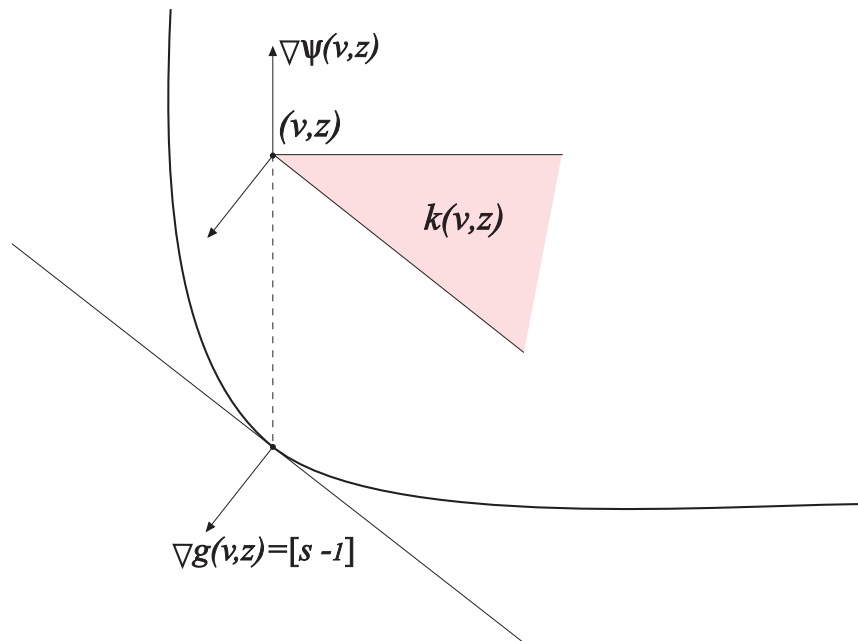


Figura 4.4: O cone de direções $K(v, z)$ (Fonte: o autor).

Em resumo, o método IED pode ser descrito:

- A partir de um ponto no interior do epígrafo da função dual, um subgradiente é usado para definir uma aproximação linear deste mesmo epígrafo;
- Essa aproximação linear é utilizada para encontrar uma direção de busca que fornece um ponto auxiliar;
- Se este ponto auxiliar estiver no interior do epígrafo, temos um passo sério e o auxiliar será o novo iterado, a partir do qual o processo se repete;
- Caso contrário, tem-se um passo nulo e uma atualização a partir do ponto corrente é realizada de acordo com o passo do algoritmo DSG;
- O novo ponto substitui o ponto corrente e o processo é repetido até um passo sério ocorrer;
- Assim, o IED gera uma sequência no interior do epígrafo da função dual, bem como uma sequência primal, o que motiva o nome dado ao método.

4.4.2 Algoritmo IED

Escolha $\eta_0 \in (0, 1)$, uma matriz simétrica A e uma função aumentada σ . Tome uma sequência $\{\alpha_k\} \in \mathbb{R}_+$ de números positivos e uma sequência $\{B^k\}$ de matrizes definidas positivas. Por conveniência, considere o conjunto $X(u, c) = \text{Arg min}_{x \in X} [f(x) + c\sigma(g(x)) + \langle Au, g(x) \rangle]$.

Passo 0: Escolha $u^0 \in \mathbb{R}^m$, $c^0 > 0$, $\lambda^0 > 0$ e $z^0 > q(v^0) = q(u^0, c^0)$. Tome $x^0 \in X(v^0)$ tal que $g(x^0) \neq 0$. Seja $k = 0$.

Passo k: Seja $(v^k, z^k) \in \text{int}(\text{epi}(q))$, $\lambda^k > 0$ e $x^k \in X(v^k)$ tal que $g(x^k) \neq 0$. Calcule $s^k = (-A(g(x^k)), -\sigma(g(x^k)))$.

Passo k.1: Seja $i = 0$. Dado (v^k, z^k, λ^k) , calcule $d_1^k, d_2^k, \lambda_1^k$ e λ_2^k resolvendo os sistemas

$$B^k d_1 + \nabla g(x^k) \lambda_1 = -\nabla \psi(u^k, c^k, z^k) \quad (4.25)$$

$$\Lambda^k \nabla g(x^k)^T d_1 + G(x^k) \lambda_1 = 0 \quad (4.26)$$

e

$$B^k d_2 + \nabla g(x^k) \lambda_2 = 0 \quad (4.27)$$

$$\Lambda^k \nabla g(x^k)^T d_2 + G(x^k) \lambda_2 = -\Lambda^k \quad (4.28)$$

Calcule ρ_1^k, ρ_2^k satisfazendo $\rho_1^k = -\frac{(d_1^k)^T \nabla g(v^k, z^k)}{(d_2^k)^T \nabla g(v^k, z^k)} < \rho^k \leq (1 - \xi^k) \frac{(d_1^k)^T \nabla \psi(v^k, z^k)}{(d_2^k)^T \nabla \psi(v^k, z^k)} < \rho_2^k$, com $\xi^k \in \left(\frac{1}{3} \left(1 - \frac{\rho_1^k}{\rho_2^k}\right), \frac{2}{3} \left(1 - \frac{\rho_1^k}{\rho_2^k}\right) \right) \subset (0, 1)$, e a direção de busca $d^k = d_1^k + \rho^k d_2^k$.

Calcule o tamanho de passo $t^k = \max\left\{\eta_0, \frac{z^k - \hat{q}}{-\nabla \psi(v^k, z^k)^T d^k}, \frac{|(v^k)^T s^k|}{-\nabla \psi(v^k, z^k)^T d^k}\right\}$.

Passo k.2: Calcule o ponto auxiliar $(\hat{v}^k, \hat{z}^k) = (v^k, z^k) + t^k d^k$. Calcule $\hat{x}^k \in X(\hat{v}^k, \hat{z}^k)$. Se $g(\hat{x}^k) = 0$, pare. Sejam $x^k = \hat{x}^k$ e $v^k = \hat{v}^k$.

Passo k.3.a: (Passo sério) Se o ponto auxiliar $(\hat{v}^k, \hat{z}^k) \in \text{int}(\text{epi}(q))$ então $(v^{k+1}, z^{k+1}) = (\hat{v}^k, \hat{z}^k)$, $x^{k+1} = \hat{x}^k$, $\lambda^{k+1} = \lambda^k + t^k(\lambda_1^k + \rho^k \lambda_2^k) = \hat{\lambda}^k$, $k = k + 1$ e siga para o passo k .

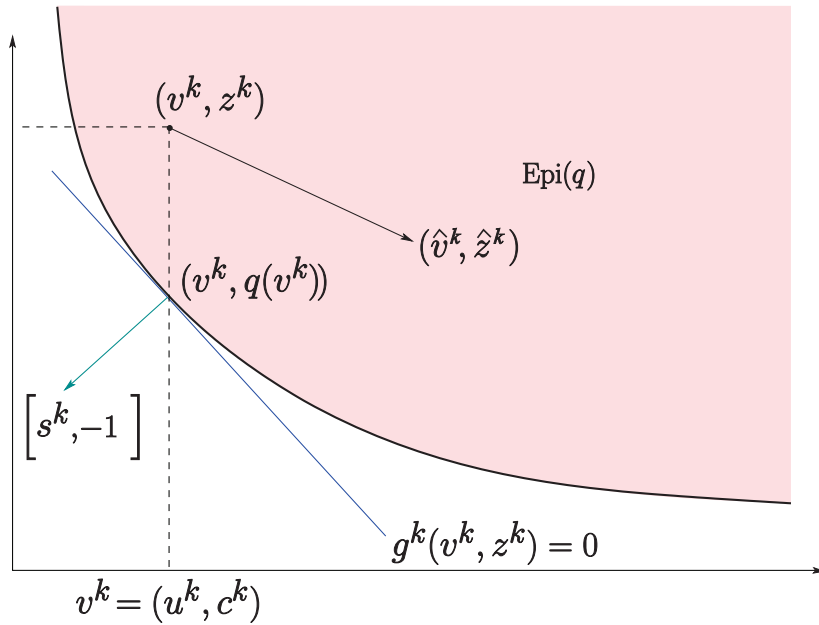


Figura 4.5: Passo s3rio: $(v^{k+1}, z^{k+1}) = (\hat{v}^k, \hat{z}^k)$ (Fonte: [5]).

Step k.3.b: (Passo nulo) Faça $i = i + 1$. Escolha $r_k > 0$ como no passo 2 do algoritmo DSG. Calcule um novo ponto no $\text{int}(\text{epi}(q))$ realizando o passo DSG a partir de v^k :

$$\tilde{u}^{k,i} = u^k - r_k A(g(x^k));$$

$$\tilde{c}^{k,i} = c^k - (1 + \alpha_k) r_k \sigma(g(x^k));$$

$$\tilde{v}^{k,i} = (\tilde{u}^{k,i}, \tilde{c}^{k,i}).$$

Calcule $\tilde{x}^{k,i} \in X(\tilde{v}^{k,i})$. Se $g(\tilde{x}^{k,i}) = 0$, pare. Caso contr3rio, calcule $q(\tilde{v}^{k,i})$ e tome $\tilde{z}^{k,i}$ tal que $q(\tilde{v}^{k,i}) < \tilde{z}^{k,i} < q(v^k)$. Seja $(v^k, z^k) = (\tilde{v}^{k,i}, \tilde{z}^{k,i})$; $x^k = \tilde{x}^{k,i}$, e siga para o passo k.

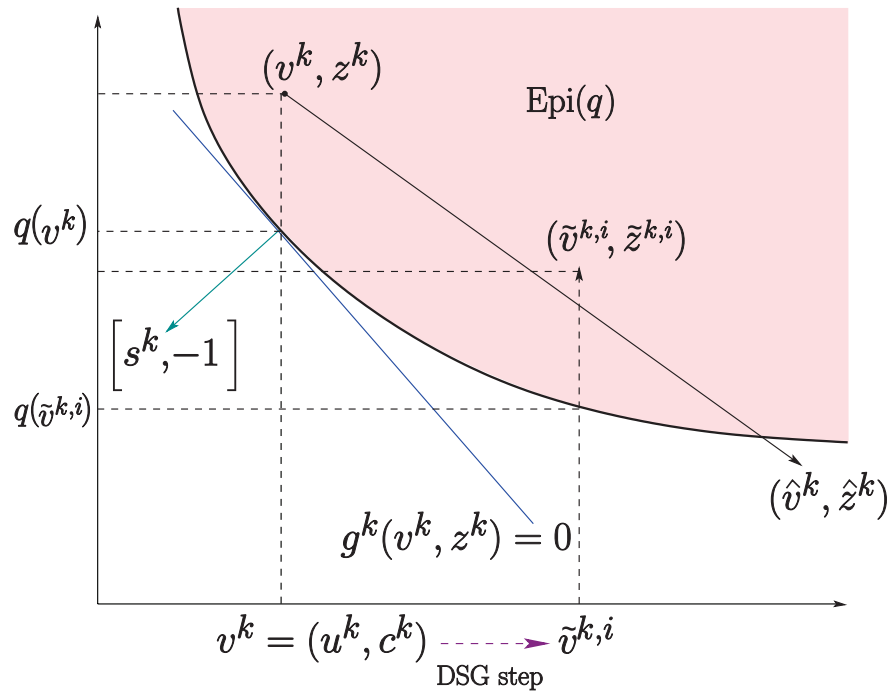


Figura 4.6: Passo nulo: $(v^k, z^k) = (\tilde{v}^{k,i}, \tilde{z}^{k,i})$ (Fonte: [5]).

4.4.2.1 Algoritmo DSG

Apresenta-se aqui os passos do algoritmo DSG [29]:

Step 0: Escolha $v^0 = (u^0, c^0) \in \mathbb{R}^m \times \mathbb{R}_+$. Fixe $\beta \geq \eta > 0$. Seja $k = 0$.

Step 1: (Critério de parada)

- Calcule $x^k \in X(v^k)$,
- Se $g(x^k) = 0$ pare. Caso contrário, siga para o passo 2.

Step 2: (Cálculo do passo e atualização das variáveis duais)

$\eta_k := \min\{\eta, \sigma(g(x^k))\}$, $\beta_k := \max\{\beta, \sigma(g(x^k))\}$, e escolha r_k em $[\eta_k, \beta_k]$,

$$u^{k+1} = u^k - r_k A(g(x^k)),$$

$$c^{k+1} = c^k - (1 + \alpha_k) r_k \sigma(g(x^k)).$$

Seja $k = k + 1$ e siga para o passo 1.

4.4.2.2 Função *fminsearch*: resolução do subproblema

Conforme dito na seção 4.4, a cada iteração do método IED é necessário resolver a minimização da função Lagrangeana 4.22, ou seja, a resolução do subproblema. Para tal fim, na versão original do IED [28], é utilizada a função de otimização *fminsearch* do Matlab, que se baseia no método de Nelder-Mead [6], um dos métodos de busca direta mais utilizados em problemas de minimização sem restrições de uma função de n variáveis.

A popularidade do método de Nelder-Mead deve-se a diferentes fatores, tais como [30]:

- não necessidade do cálculo de derivadas da função objetivo;
- acentuado decréscimo do valor da função objetivo nas primeiras iterações, além de serem feitas poucas avaliações da mesma por iteração;
- simplicidade de sua implementação computacional.

O método de Nelder-Mead, assim como a maior parte dos métodos de busca direta, utiliza os conceitos de um simplex no \mathbb{R}^n . No caso de duas variáveis, por exemplo, um simplex é um triângulo, e o método compara os valores da função nos três vértices do triângulo [6].

O algoritmo usa 4 operações possíveis, cujos coeficientes escalares são:

- coeficiente de reflexão: $\rho > 0$,
- coeficiente de expansão: $\chi > 1$ com $\chi > \rho$,
- coeficiente de contração: $0 < \gamma < 1$,
- coeficiente de redução (shrink): $0 < \sigma < 1$.

Uma escolha, praticamente universal, é definir $\rho = 1$, $\chi = 2$, $\gamma = 1/2$ e $\sigma = 1/2$. [6]

Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e um simplex com $n + 1$ vértices. Em uma dada iteração do método, os $n + 1$ vértices do simplex, x_1, x_2, \dots, x_{n+1} pertencentes a \mathbb{R}^n , são ordenados de acordo com o crescimento dos valores de f , isto é,

$$f(x_1) \leq f(x_2) \leq \dots \leq f(x_{n+1}),$$

onde x_1 é definido como o melhor vértice e x_{n+1} o pior, isto é, o vértice com o maior valor da função objetivo. O método busca, então, substituir este pior vértice do simplex

por outro com melhor valor. O novo vértice é obtido através da reflexão, expansão ou contração do pior vértice ao longo da reta que passa por este vértice e o centróide dos n melhores vértices (\bar{x}), isto é,

$$\bar{x} = \frac{1}{2} \sum_{i=1}^n x_i$$

A cada iteração o pior vértice é substituído por um novo vértice. Um novo simplex é formado e a pesquisa continua, gerando uma sequência de simplex na qual os valores da função nos vértices ficam cada vez menores [30].

Assim, as principais etapas do método podem ser observadas nas figuras 4.7 e 4.8 [31]:

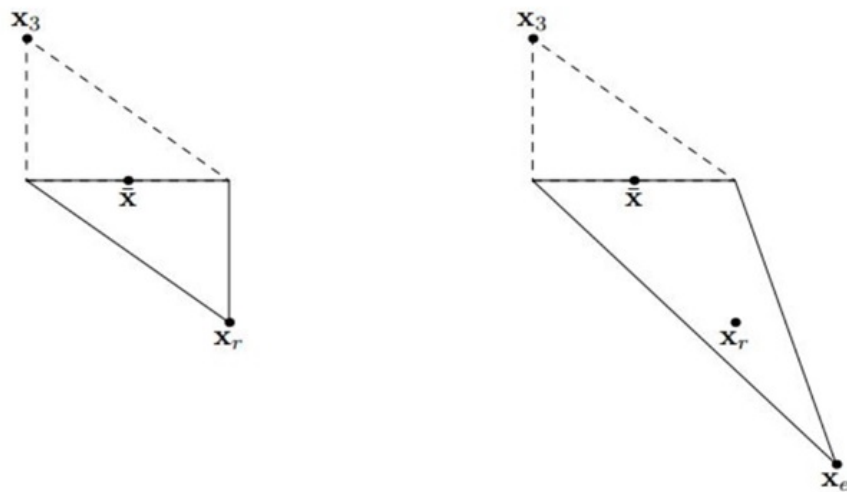


Figura 4.7: Método de Nelder-Mead após uma etapa de reflexão e expansão (Fonte: [6]).

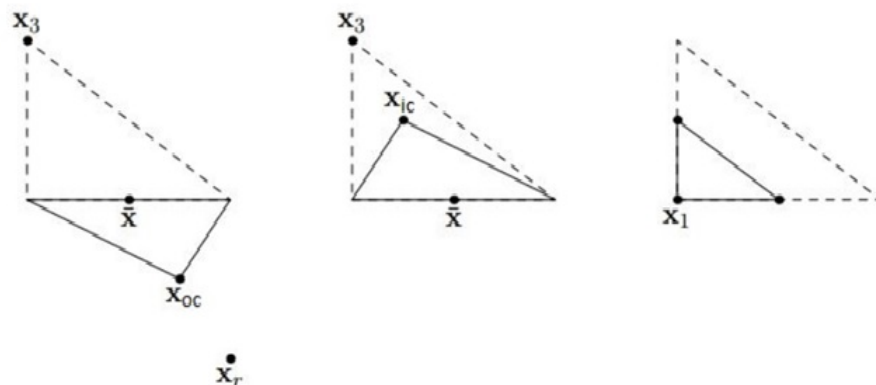


Figura 4.8: Método de Nelder-Mead após uma contração externa, uma contração interna e uma redução (Fonte: [6]).

Na figura 4.7 observa-se que, se o ponto refletido x_r é melhor que o vértice, então o simplex foi refletido numa direção que minimiza a função objetivo. É importante, neste caso, fazer a expansão x_e na mesma direção. Caso contrário, se o ponto refletido não possuir um valor melhor da função objetivo, então é provável que o simplex esteja próximo do ponto de mínimo da função, sendo testados, assim, os pontos os pontos de contração externa x_e e interna x_i . Se nenhuma das operações resultou em um novo vértice com valor da função objetivo pelo menos melhor do que aquele correspondente ao vértice a ser rejeitado, deve-se reduzir o simplex, pois o ponto de mínimo está em seu interior. Esta operação de redução é feita preservando o vértice x_1 e aproximando os demais vértices na direção de x_1 , conforme figura 4.8.

Uma iteração do método de Nelder-Mead é descrita a seguir:

Para exemplificarmos o método, consideremos a minimização irrestrita da seguinte função de duas variáveis [7, p. 408]:

$$f(x, y) = x^2 + y^2 - xy - 4x - y.$$

Sejam os três vértices iniciais

$$x_1 = (0, 0), \quad x_2 = (1.2, 0) \quad \text{e} \quad x_3 = (0, 0.8),$$

com a função assumindo os valores em ordem crescente

$$f(x_2) = -3.36, \quad f(x_3) = -0.16 \quad \text{e} \quad f(x_1) = 0.$$

Calcula-se o ponto médio \bar{x} do segmento $\overline{x_2x_3}$ (melhor lado)

$$\bar{x} = (0.6, 0.4).$$

Ao longo da linha que liga o pior vértice x_1 e o ponto médio \bar{x} , é calculado o ponto refletido x_r

$$x_r = 2\bar{x} - x_1 = (1.2, 0.8).$$

Como o valor da função $f(1.2, 0.8) = -4.48$ é menor que em x_2 , deve-se mover ao longo da mesma direção e calcula-se o ponto de expansão

$$x_e = 2x_r - \bar{x} = 2(1.2, 0.8) - (0.6, 0.4) = (1.8, 1.2).$$

Estes passos estão ilustrados na figura 4.9 a seguir:

Algoritmo 1: Método Simplex de Nelder-Mead

Dados:

- (1) Função $f(x)$ com $x \in \mathbb{R}^n$
- (2) Conjunto de vértices iniciais: $x_1, x_2, \dots, x_{n+1} \in \mathbb{R}^n$
- (3) Coeficiente de reflexão $\rho > 0$
- (4) Coeficiente de expansão $\chi > 1$
- (5) Coeficiente de contração $0 < \gamma < 1$
- (6) Coeficiente de redução $0 < \sigma < 1$

Passo 1: Calcule o valor da função objetivo para os $n + 1$ vértices e ordene-os de modo que $f(x_1) \leq f(x_2) \leq \dots \leq f(x_{n+1})$

Passo 2: Calcule o centróide \bar{x} dos n melhores vértices

Passo 3: Enquanto algum critério de parada não for atingido, faça

(3a) Calcule o ponto de reflexão $x_r = \bar{x} + \rho(\bar{x} - x_{n+1})$

(3b) Calcule $f(x_r)$

(3c) Se $f(x_1) \leq f(x_r) < f(x_n)$ então $x_{n+1} = x_r$

(3d) Se $f(x_r) \leq f(x_1)$ então calcule o ponto de expansão $x_e = \bar{x} + \chi(x_r - \bar{x})$

Calcule $f(x_e)$

Se $f(x_e) \leq f(x_r)$ então $x_{n+1} = x_e$

Senão $x_{n+1} = x_r$

(3e) Se $f(x_n) \leq f(x_r) < f(x_{n+1})$ então calcule o ponto de contração

externa $x_{oc} = \bar{x} + \gamma(x_r - \bar{x})$

Calcule $f(x_{oc})$

Se $f(x_{oc}) \leq f(x_r)$ então $x_{n+1} = x_{oc}$

Senão vá para o passo (3g)

(3f) Se $f(x_r) \geq f(x_{n+1})$ então calcule o ponto de contração

interna $x_{ic} = \bar{x} + \gamma(x_r - \bar{x})$

Calcule $f(x_{ic})$

Se $f(x_{ic}) \leq f(x_{n+1})$ então $x_{n+1} = x_{ic}$

Senão vá para o passo (3g)

(3g) Redução: para $2 \leq i \leq n + 1$ faça $x_i = x_1 + \sigma(x_i - x_1)$

(3h) Volte ao passo 2

Passo 4: Retorna x

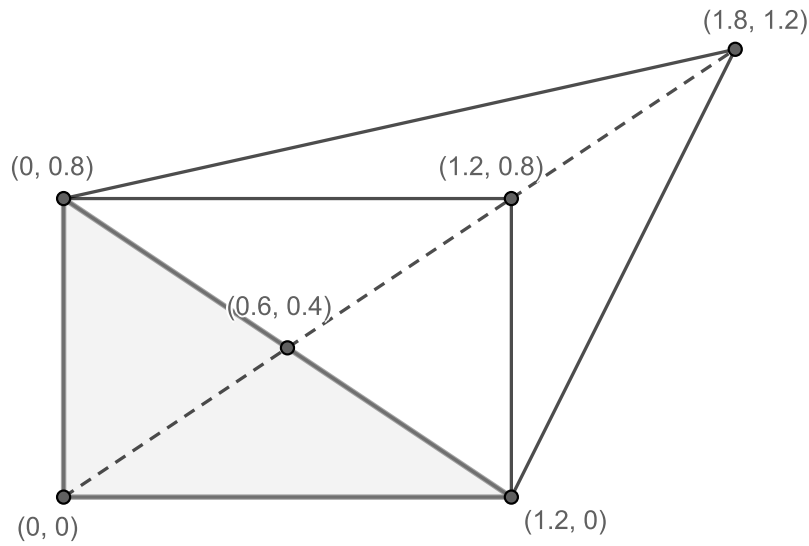


Figura 4.9: Reflexão e expansão.

A função em x_e é $f(1.8, 1.2) = -5.88$, logo o vértice x_1 é substituído pelo vértice x_e e o novo triângulo tem vértices

$$x_1 = (1.8, 1.2), \quad x_2 = (1.2, 0) \quad \text{e} \quad x_3 = (0, 0.8).$$

Na nona iteração os pontos dos vértices assumem os valores

$$f(3.0, 1.8) = -6.96, \quad f(3.15, 2.25) = -6.9525 \quad \text{e} \quad f(2.55, 2.05) = -6.7725.$$

O ponto médio do segmento que liga os dois melhores vértices é

$$\bar{x} = \left(\frac{3 + 3.15}{2}, \frac{1.8 + 2.25}{2} \right) = (3.075, 2.025).$$

O ponto refletido x_r é calculado

$$x_r = 2\bar{x} - x_1 = (3.6, 2),$$

com a função assumindo o valor $f(3.6, 2) = -0.24$, resultado pior que dos três vértices.

Neste caso, o algoritmo calcula o ponto de contração

$$x_{oi} = \left(\frac{3.075 + 2.55}{2}, \frac{2.025 + 2.05}{2} \right) = (2.8125, 2.0375),$$

no qual a função vale $f(2.8125, 2.0375) = -6.9564$. O décimo triângulo é formado com a substituição do pior vértice de T_9 pelo ponto de contração interna x_{oi} . A figura 4.10 mostra a passagem de T_9 para T_{10} .

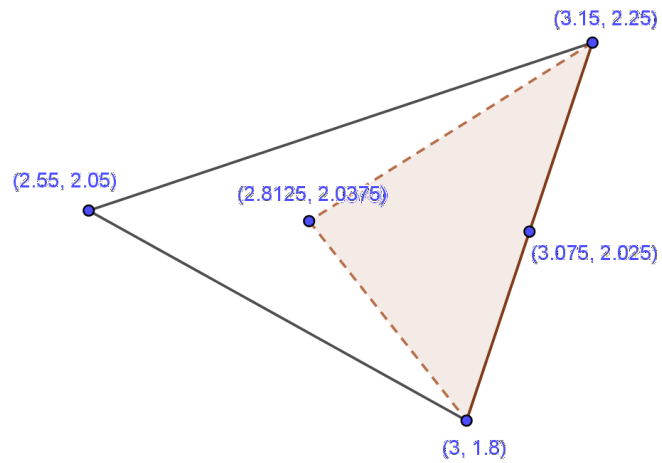


Figura 4.10: Contração.

O processo continua e gera uma sequência de triângulos que converge para a solução $(3, 2)$. A figura 4.11 mostra os triângulos das dez primeiras iterações.

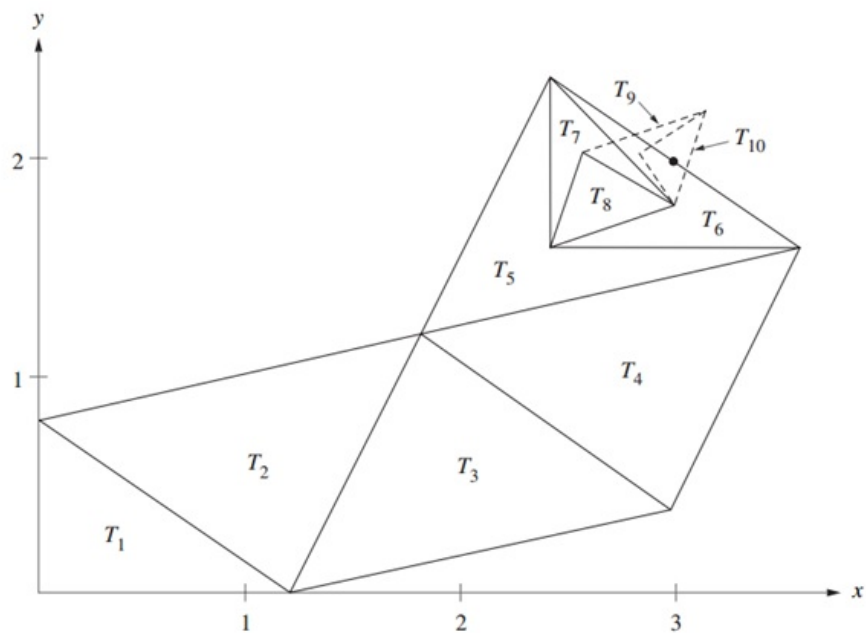


Figura 4.11: Sequência T_k dos triângulos [7, p. 409].

Dada uma tolerância Δ_{tol} , o critério de parada do algoritmo Nelder-Mead leva em conta o valor da função nos vértices do simplex [32]:

$$\sqrt{\sum_{i=1}^{n+1} \frac{(f(x_i) - f(\bar{x}))^2}{n}} < \Delta_{tol}$$

Embora o método funcione bem para problemas de baixa dimensão, pode-se tornar ineficiente para grandes dimensões [31].

5 IED+NFDNA: UMA PRIMEIRA VERSÃO DO MÉTODO

Descreve-se neste capítulo uma versão diferente para o IED, denominada IED+NFDNA.

As seguintes etapas foram realizadas:

1. Implementou-se o algoritmo NFDNA (subseção 4.3.1) para problemas irrestritos, não diferenciáveis e não convexos.
2. Eliminou-se da versão original do IED a função de otimização *fminsearch* do Matlab, substituindo-a pelo NFDNA, para resolver o subproblema (4.22), ou seja, minimizar a função Lagrangeana. Testes numéricos anteriores mostraram que, para alguns casos, o subproblema não havia sido resolvido com o IED original [5].
3. Realizou-se três experimentos numéricos com a nova versão IED+NFDNA, cujos resultados estão apresentados no capítulo 5.

A Figura 5.1 a seguir mostra o diagrama de fluxo da nova versão IED+NFDNA:

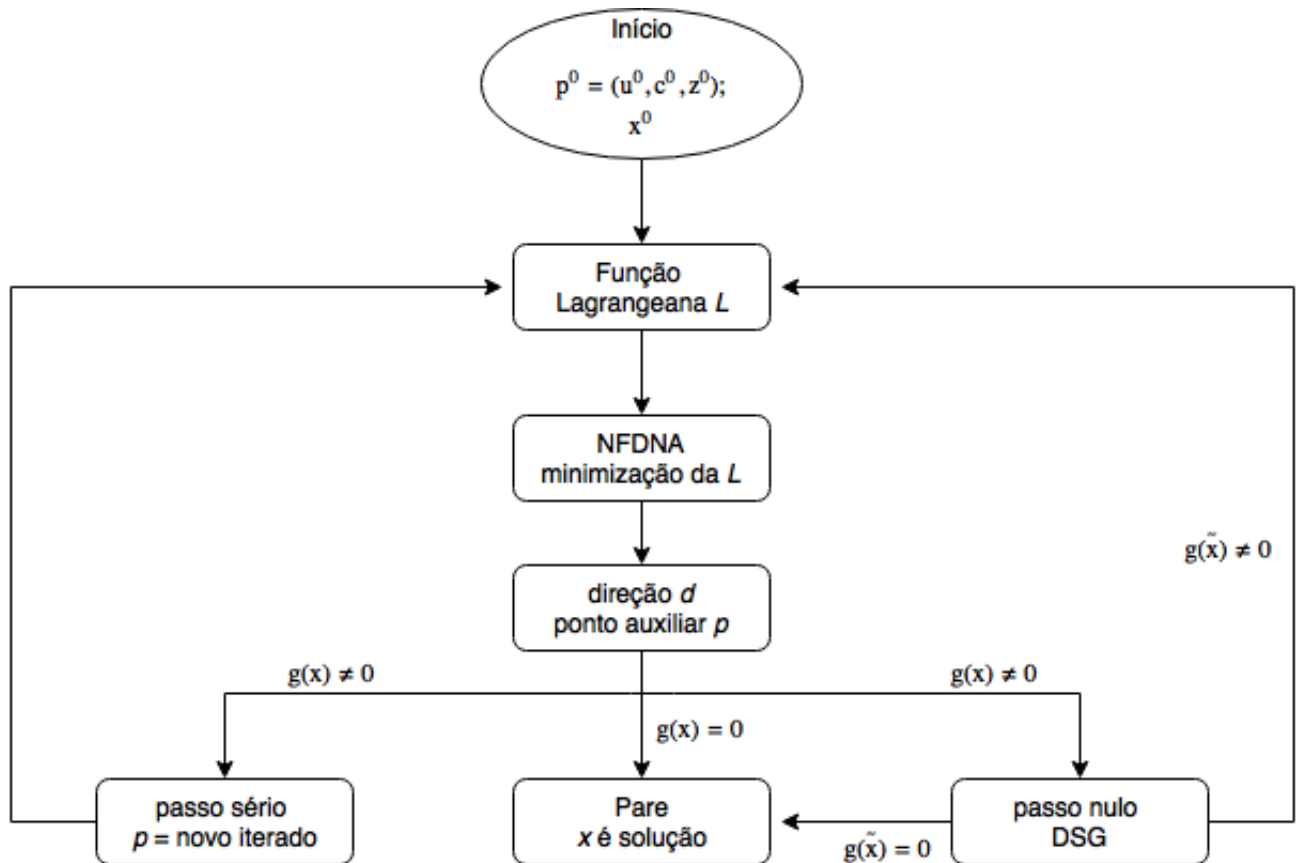


Figura 5.1: Diagrama de fluxo do método IED+NFDNA (Fonte: o autor).

Os métodos de otimização abordados, incluindo a nova versão do IED, estão organizados historicamente conforme o fluxograma apresentado na Figura 5.2:

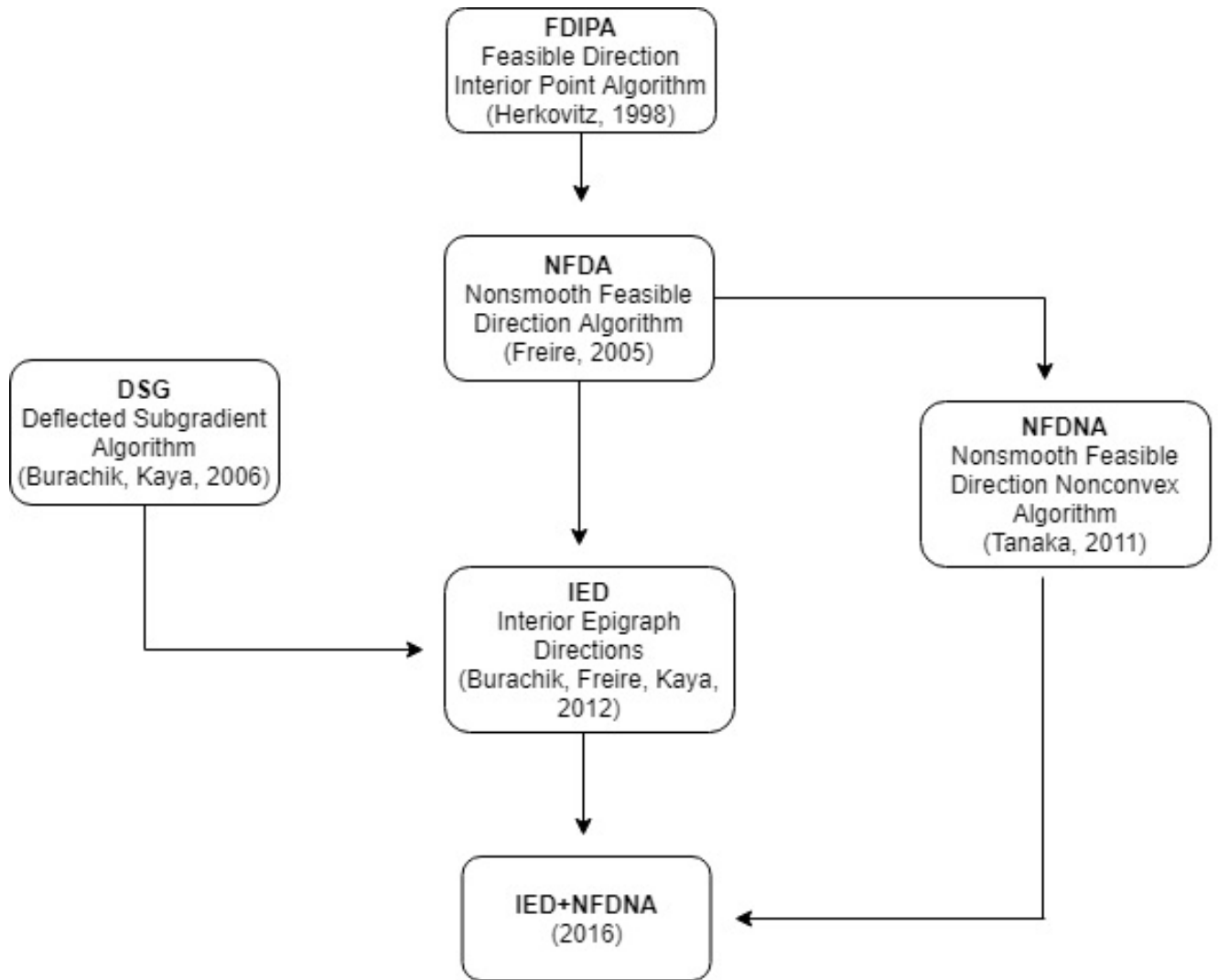


Figura 5.2: Sequência histórica dos métodos (Fonte: o autor).

5.1 EXPERIMENTOS NUMÉRICOS

Apresentamos nesta seção os resultados numéricos obtidos pelo algoritmo IED+NFDNA.

5.1.1 *Experimento Numérico 1*

Neste 1º experimento comparou-se as duas versões do método IED, a original IED+fmin e a nova IED+NFDNA. Foram utilizadas sete diferentes funções de penalidade $\sigma(\cdot)$ e, para

a matriz simétrica A do termo linear da Lagrangeana, escolheu-se as matrizes identidade $A=I$ e nula $A=0$. A comparação entre as versões foi feita com base no número de iterações do algoritmo n_{its} e no número de vezes em que a função Lagrangeana foi avaliada n_L no subproblema. Os cinco problemas teste deste experimento foram retirados de Hock e Schittkowski collection [33] e suas formulações e resultados estão apresentados a seguir:

5.1.1.1 Descrição dos problemas

PPR-P1-2

Função objetivo: $f(x) = 5x_1 + 50000/x_1 + 20x_2 + 72000/x_2 + 10x_3 + 144000/x_3$

Restrições: $1 - 4/x_1 - 32/x_2 - 120/x_3$; $1 \cdot 10^{-5} \leq x_i, \quad i = 1, 2, 3$

Valor ótimo: $f(x^*) = 6299,8424$

LGR-P1-1

Função objetivo: $f(x) = 0,2x_3 - 0,8x_1$

Restrições: $x_2 - \exp(x_1) \geq 0$; $x_3 - \exp(x_2) \geq 0$; $0 \leq x_1 \leq 100$; $0 \leq x_2 \leq 100$;

$0 \leq x_3 \leq 10$ Valor ótimo: $f(x^*) = 0,5181$

GPR-P1-2

Função objetivo: $f(x) = \exp(x_1x_2x_3x_4x_5) - 0,5(x_1^3 + x_2^3 + 1)^2$

Restrições: $x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 - 10 = 0$; $x_2x_3 - 5x_4x_5 = 0$; $x_1^3 + x_2^3 + 1 = 0$;

$-2,3 \leq x_i \leq 2,3, \quad i = 1, 2$; $-3,2 \leq x_i \leq 3,2, \quad i = 3, 4, 5$

Valor ótimo: $f(x^*) = 0,5394$

PPR-P1-3

Função objetivo: $f(x) = x_1x_4(x_1 + x_2 + x_3) + x_3$

Restrições: $x_1x_2x_3x_4 - 25 \geq 0$; $x_1^2 + x_2^2 + x_3^2 + x_4^2 - 40 = 0$; $1 \leq x_i \leq 5, \quad i = 1, \dots, 4$

Valor ótimo: $f(x^*) = 17,0114$

LPR-P1-1

Função objetivo: $f(x) = 1 + x_1 + x_2 + x_3 + x_4$

Restrições: $0,0401 - 4/x_1 - 2,25/x_2 - 1/x_3 - 0,25/x_4 \geq 0$; $0,010085 - 0.16/x_1 -$

$0,36/x_2 - 0,64/x_3 - 0,64/x_4 \geq 0; 0,001 \leq x_i \leq (5-1)10^5, i = 1, \dots, 4$

Valor ótimo: $f(x^*) = 727,6793$

A	$\sigma(\cdot)$	PPR-P1-2		LGR-P1-1		GPR-P1-2		PPR-P1-3		LPR-P1-1	
		c_0	u_0	c_0	u_0	c_0	u_0	c_0	u_0	c_0	u_0
I	$\ \cdot\ _2$	0	[1111]	1	[00000000]	3	[000000000000]	10	[1111111111]	2	[0000000000]
	$\ \cdot\ _1$	0	[1111]	0.1	[00000000]	3	[000000000000]	10	[1111111111]	2	[0000000000]
	$\sqrt{m}\ \cdot\ _\infty$	0	[1111]	1	[00000000]	3	[000000000000]	1	[0000000000]	2	[0000000000]
	$e^{\ \cdot\ _2} - 1$	0	[1111]	3	[00000000]	1	[55555555555555]	10	[1111111111]	0.1	[0000000000]
	$e^{\ \cdot\ _1} - 1$	0	[1111]	3	[11111111]	3	[000000000000]	10	[1111111111]	0.05	[0000000000]
	$\sqrt{m}(e^{\ \cdot\ _\infty} - 1)$	0	[1111]	2	[11111111]	3	[000000000000]	10	[1111111111]	0.1	[0000000000]
	$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	0	[1111]	0.1	[00000000]	3	[11111111111111]	1	[0000000000]	2	[0000000000]
0	$\ \cdot\ _2$	0	[1111]	1	[00000000]	3	[000000000000]	15	[1111111111]	2	[0000000000]
	$\ \cdot\ _1$	0	[1111]	0.1	[00000000]	3	[000000000000]	10	[1111111111]	2	[0000000000]
	$\sqrt{m}\ \cdot\ _\infty$	0	[1111]	1	[00000000]	3	[000000000000]	1	[0000000000]	2	[0000000000]
	$e^{\ \cdot\ _2} - 1$	0	[1111]	3	[00000000]	0.5	[55555555555555]	10	[1111111111]	0.05	[0000000000]
	$e^{\ \cdot\ _1} - 1$	0	[1111]	3	[11111111]	3	[000000000000]	10	[1111111111]	0.05	[0000000000]
	$\sqrt{m}(e^{\ \cdot\ _\infty} - 1)$	0	[1111]	3	[11111111]	3	[000000000000]	10	[1111111111]	0.1	[0000000000]
	$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	0	[1111]	1	[00000000]	1	[11111111111111]	1	[0000000000]	2	[0000000000]

Tabela 5.1: Valores iniciais das variáveis duais.

A	$\sigma(\cdot)$	PPR-P1-2				LGR-P1-1				GPR-P1-2				PPR-P1-3				LPR-P1-1			
		IED+fmin		IED+NFDNA		IED+fmin		IED+NFDNA		IED+fmin		IED+NFDNA		IED+fmin		IED+NFDNA		IED+fmin		IED+NFDNA	
		n_{its}	n_L	n_{its}	n_L	n_{its}	n_L	n_{its}	n_L	n_{its}	n_L	n_{its}	n_L	n_{its}	n_L	n_{its}	n_L	n_{its}	n_L	n_{its}	n_L
I	$\ \cdot\ _2$	3	1306	2	102	—	—	1	111	—	—	1	84	—	—	1	118	7	4166	7	115
	$\ \cdot\ _1$	3	1306	2	102	4	1877	1	111	—	—	—	—	—	—	—	7	5148	7	115	
	$\sqrt{m}\ \cdot\ _\infty$	3	993	2	102	2	1482	1	76	—	—	—	—	—	—	2	118	8	7764	3	115
	$e^{\ \cdot\ _2} - 1$	—	—	2	102	3	1039	1	76	2	1173	1	123	3	5105	—	—	7	3648	7	115
	$e^{\ \cdot\ _1} - 1$	—	—	2	102	4	2180	1	76	2	3941	—	—	2	5797	—	—	7	4266	5	115
	$\sqrt{m}(e^{\ \cdot\ _\infty} - 1)$	—	—	2	102	2	1196	1	111	4	5352	—	—	2	4813	—	—	8	4934	5	115
	$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	1	1145	2	102	2	1582	1	76	2	3060	2	123	2	2254	44	118	7	4420	9	115
0	$\ \cdot\ _2$	3	1345	2	102	—	—	1	111	—	—	1	84	—	—	1	118	8	4354	10	115
	$\ \cdot\ _1$	3	1345	2	102	4	1877	1	111	—	—	—	—	—	—	—	7	4984	10	115	
	$\sqrt{m}\ \cdot\ _\infty$	3	1236	2	102	2	1482	1	76	—	—	—	—	—	—	1	118	8	6237	5	115
	$e^{\ \cdot\ _2} - 1$	—	—	2	102	3	1039	1	76	1	3048	1	123	2	1663	—	—	7	4555	9	115
	$e^{\ \cdot\ _1} - 1$	—	—	2	102	4	2180	1	76	1	1790	—	—	2	3769	—	—	7	3926	5	115
	$\sqrt{m}(e^{\ \cdot\ _\infty} - 1)$	—	—	2	102	2	1196	1	76	1	3484	—	—	2	5637	—	—	8	5993	5	115
	$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	3	1309	2	102	3	1303	1	76	1	3060	1	123	2	2890	2	118	7	4252	9	115

Tabela 5.2: Resultados obtidos pelas duas versões do IED.

5.1.1.2 Análise do experimento numérico 1

- A tabela 5.1 mostra os valores iniciais das variáveis duais que foram mantidas para ambas as versões do IED para fins de comparação.
- Na tabela 5.2, tem-se que o número de avaliações da Lagrangeana n_L até a solução final é muito menor na nova versão IED+NFDNA quando comparado à versão original IED+fmin, em todos os problemas. Os valores em destaque são os menores encontrados pelos algoritmos em cada problema. Esta diferença é importante na medida em que a resolução do subproblema - a minimização da Lagrangeana - é uma das distinções entre as versões do método.
- Nos problemas PPR-P1-2, LGR-P1-1 e LPR-P1-1, o IED+NFDNA preencheu todas as funções σ nas quais a versão original não obteve solução.
- O número de iterações n_{its} , obtido por ambas as versões, é similar em todos os testes.
- Em algumas funções σ , assinaladas por traço (-), não foram encontradas soluções.

5.1.2 Experimento Numérico 2

Após a comparação entre as duas versões do método IED demonstrada na seção anterior, neste segundo experimento, foi testada a nova versão IED+NFDNA com outros problemas da literatura [33]. Comparou-se as soluções obtidas pelo algoritmo com as soluções de referência dos problemas, utilizando as mesmas sete funções σ e as duas matrizes A do termo linear na composição da Lagrangeana. Os problemas estão descritos a seguir, bem como os resultados obtidos com destaque para as melhores soluções.

5.1.2.1 Descrição dos problemas

QPR-T1-1

Função objetivo: $f(x) = (x_1 - 2)^2 + x_2^2$

Restrições: $(1 - x_1)^3 - x_2 \geq 0$; $0 \leq x_1$; $0 \leq x_2$

Valor ótimo: $f(x^*) = 1$

PQR-T1-7

Função objetivo: $f(x) = -x_1x_2x_3$

Restrições: $-x_1^2 - 2x_2^2 - 4x_3^2 + 48 \geq 0$

Valor ótimo: $f(x^*) = -22,6274$

QQR-T1-8

Função objetivo: $f(x) = x_1^2 + x_2^2 + x_3^2$

Restrições: $x_1^2 + x_2^2 - 1 \geq 0$; $1 \leq x_1 \leq 10$; $-10 \leq x_2 \leq 10$; $-10 \leq x_3 \leq 10$

Valor ótimo: $f(x^*) = 1$

QQR-P1-3

Função objetivo: $f(x) = (x_1 - x_2)^2 + (x_1 + x_2 - 10)^2/9 + (x_3 - 5)^2$

Restrições: $48 - x_1^2 - x_2^2 - x_3^2 \geq 0$; $-4,5 \leq x_i \leq 4,5$, $i = 1, 2$; $-5 \leq x_3 \leq 5$

Valor ótimo: $f(x^*) = 0,9535$

PBR-T1-1

Função objetivo: $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$

Restrições: $-x_2 - 1, 5 \leq 0$

Valor ótimo: $f(x^*) = 0$

A	$\sigma(\cdot)$	QPR-T1-1			PQR-T1-7			QQR-T1-8			QQR-P1-3			PBR-T1-1			
		n_{its}	n_L	Soluções	n_{its}	n_L	Soluções	n_{its}	n_L	Soluções	n_{its}	n_L	Soluções	n_{its}	n_L	Soluções	
I	$\ \cdot\ _2$	2	3362	0.9818	2	3013	-22.6274	1	532	1	1	381	0.9666	1	6368	0	
	$\ \cdot\ _1$	8	22844	0.9885	2	3013	-22.6274	1	21	1	1	478	0.9535	1	6368	0	
	$\sqrt{m}\ \cdot\ _\infty$	9	7375	1.0538	2	3013	-22.6274	1	2285	1	1	641	1.0613	1	6368	0	
	$e^{\ \cdot\ _2} - 1$	—	—	—	1	5665	-22.6274	—	—	—	1	1083	0.9541	1	6047	0	
	$e^{\ \cdot\ _1} - 1$	1	792	0.9811	1	7131	-22.6274	1	21	1	1	5439	0.9540	1	6047	0	
	$\sqrt{m}(e^{\ \cdot\ _\infty} - 1)$	10	18063	1.1801	1	7131	-22.6274	1	261	1.0003	1	1	4168	0.9535	1	6047	0
	$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	3	763	0.9893	5	2359	-22.6274	1	669	1	1	1	143	1.0518	1	6911	0
0	$\ \cdot\ _2$	—	—	—	1	4880	-22.6274	1	179	1.0015	1	1	381	0.9666	1	5391	0
	$\ \cdot\ _1$	1	325	0.9925	1	4880	-22.6274	1	4293	1	1	1	520	0.9578	1	5391	0
	$\sqrt{m}\ \cdot\ _\infty$	—	—	—	1	4880	-22.6274	1	1940	1	1	1	717	0.9544	1	5391	0
	$e^{\ \cdot\ _2} - 1$	8	8728	0.9897	1	2288	-22.6274	—	—	—	—	—	—	—	1	5391	0
	$e^{\ \cdot\ _1} - 1$	—	—	—	1	2288	-22.6274	1	290	1	1	1	6686	0.9564	1	5391	0
	$\sqrt{m}(e^{\ \cdot\ _\infty} - 1)$	11	57201	1.3005	1	2288	-22.6274	1	214	1	1	1	4168	0.9535	1	5391	0
	$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	2	353	1.0445	1	10167	-22.6274	1	216	1.0098	1	1	360	1.1377	1	5391	0
Valores de referência			1			-22.6274		1				0.9535			0		

Tabela 5.3: Resultados obtidos pelo IED+NFDNA.

5.1.2.2 Análise do experimento numérico 2

- As soluções encontradas pelo IED+NFDNA neste experimento (Tabela 5.3) mostram compatibilidade com os valores de referência dos problemas.
- Em destaque, no PQR-T1-7 e PBR-T1-1 foram obtidas as soluções de referência em todas as funções σ e matrizes A .

5.1.3 Experimento Numérico 3

Nesta terceira experimentação, o objetivo foi comparar o IED+NFDNA com um solver diferente do método. Para tal, foi escolhido o algoritmo de Pontos Interiores da função

Fmincon do MatLab, destinada à problemas não lineares com restrições de igualdade ou desigualdade. Seleccionamos a função $\sigma = \max(\|\cdot\|_2, \|\cdot\|_2^2)$, pois nos experimentos anteriores 1 e 2 foi a que obteve resultados em todos os problemas teste (Tabelas 5.2 e 5.3). Para o termo linear da Lagrangeana optou-se pela matriz $A = I$. Os problemas descritos, bem como o número de iterações, o número de vezes que a função objetivo é avaliada e as soluções obtidas estão apresentados a seguir:

5.1.3.1 Descrição dos problemas

QQR-T1-3

Função objetivo: $f(x) = 0,5x_1^2 + x_2^2 - x_1x_2 - 7x_1 - 7x_2$

Restrições: $-25 + 4x_1^2 + x_2^2 \leq 0$

Valor ótimo: $f(x^*) = -30$

QLR-T1-3

Função objetivo: $f(x) = 9 - 8x_1 - 6x_2 - 4x_3 + 2x_1^2 + 2x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3$

Restrições: $3 - x_1 - x_2 - 2x_3 \geq 0$; $0 \leq x_i, i = 1, 2, 3$;

Valor ótimo: $f(x^*) = 0,1111$

QLR-P1-1

Função objetivo: $f(x) = x_1^2 + 0,5x_2^2 + x_3^2 + 0,5x_4^2 - x_1x_3 + x_3x_4 - x_1 - 3x_2 + x_3 - x_4$

Restrições: $5 - x_1 - 2x_2 - x_3 - x_4 \geq 0$; $4 - 3x_1 - x_2 - 2x_3 + x_4 \geq 0$; $x_2 + 4x_3 - 1,5 \geq 0$;

$0 \leq x_i, i = 1, \dots, 4$

Valor ótimo: $f(x^*) = -4,6818$

PPR-P1-7

Função objetivo: $f(x) = (x_1 - 10)^2 + 5(x_2 - 12)^2 + x_3^2 + 3(x_4 - 11)^2 + 10x_5^6 + 7x_6^2 + x_7^4 - 4x_6x_7 - 10x_6 - 8x_7$

Restrições: $127 - 2x_1^2 - 3x_2^4 - x_3 - 4x_4^2 - 5x_5 \geq 0$; $282 - 7x_1 - 3x_2 - 10x_3^2 - x_4 + x_5 \geq 0$;

$196 - 23x_1 - x_2^2 - x_6^2 + 8x_7 \geq 0$; $-4x_1^2 - x_2^2 + 3x_1x_2 - 2x_3^2 - 5x_6 + 11x_7$

Valor ótimo: $f(x^*) = 680,6300$

QQR-T1-6

Função objetivo: $f(x) = (x_1 - 2)^2 + (x_2 - 1)^2$

Restrições: $x_1^2 - x_2 \leq 0$; $x_1 + x_2 - 2 \leq 0$

Valor ótimo: $f(x^*) = 1$

QQR-T1-3			PPR-P1-7			QQR-T1-6											
IED+NFDNA			Fmincon			IED+NFDNA			Fmincon			IED+NFDNA			Fmincon		
n_{its}	n_L	solução	n_{its}	n_L	solução	n_{its}	n_L	solução	n_{its}	n_L	solução	n_{its}	n_L	solução	n_{its}	n_L	solução
1	63	-30	8	30	-30	1	69	680.6301	15	147	680.6301	9	4016	1	9	31	1

Tabela 5.4: Resultados comparativos entre IED+NFDNA e Fmincon.

QLR-P1-1			QLR-T1-3								
IED+NFDNA			Fmincon			IED+NFDNA			Fmincon		
n_{its}	n_L	solução	n_{its}	n_L	solução	n_{its}	n_L	solução	n_{its}	n_L	solução
1	71	-4.6773	10	55	-4.6818	1	69	0.1118	11	48	0.1111

Tabela 5.5: Resultados comparativos entre IED+NFDNA e Fmincon.

5.1.3.2 Análise do experimento numérico 3

- As soluções ótimas obtidas pelos dois algoritmos são bastante similares, com diferenças mínimas nas casas decimais.
- O IED+NFDNA obteve menor número de iterações n_{its} em 4 dos 5 testes (Tabelas 5.4 e 5.5), entretanto, Fmincon realizou menos avaliações da Lagrangeana.
- Embora o n_{its} e o n_L tenham sido discrepantes, não foi observada variação relevante no tempo de processamento dos dois algoritmos.

6 IED+GA: UMA VERSÃO HÍBRIDA DO IED COM ALGORITMOS GENÉTICOS

Com o objetivo de permitir ao método IED uma redução na dependência dos valores dos parâmetros iniciais da função Lagrangiana, apresentamos, neste capítulo, uma versão híbrida denominada IED+GA. A inovação aqui está em associar uma abordagem heurística - algoritmos genéticos - com os métodos clássicos determinísticos que compõem o IED.

A seguir são apresentadas algumas características dos algoritmos genéticos que entraram na formulação do IED+GA, os resultados comparativos de problemas teste da literatura, além da resolução de problemas de Engenharia.

6.1 ALGORITMOS GENÉTICOS

Um algoritmo genético (AG) é um procedimento iterativo que simula o desenvolvimento de uma população baseado numa metáfora do processo biológico de evolução natural. Os indivíduos da população são codificados por uma cadeia de símbolos, chamada de *genoma*. Em otimização, os indivíduos correspondem a soluções candidatas dentro de um dado espaço de busca [1].

Como técnicas heurísticas de otimização, os AGs, diferentemente dos métodos que utilizam a derivada da função, não apresentam dependência tão forte dos valores iniciais [8]. A figura a seguir ilustra esta situação na qual, dependendo do ponto inicial e seguindo a direção de maior crescimento (gradiente), técnicas clássicas de otimização tenderão a ficar presas num máximo local.

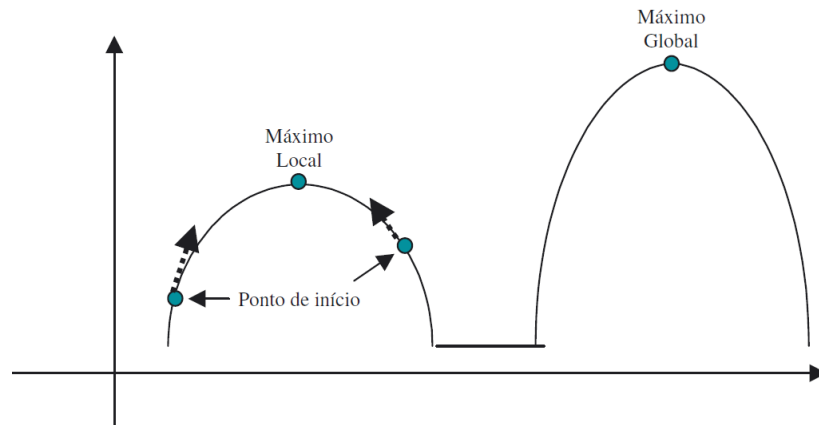


Figura 6.1: Função hipotética com um máximo local e outro global. (Fonte: [8]).

Basicamente, um AG funciona do seguinte modo [1]:

1. Escolhem-se pseudo-aleatoriamente os elementos de uma população inicial;
2. Em cada iteração (passo evolutivo ou geração), os indivíduos são avaliados segundo um determinado critério, chamado de aptidão (*fitness*). No caso da otimização, a aptidão é representada pelo valor da função objetivo da solução.
3. Para criar a próxima geração, indivíduos adequados são escolhidos para o processo reprodutório;
4. Para gerar novos pontos (indivíduos) no espaço de busca, aplicam-se operadores inspirados pela genética, como recombinação (*crossover*) e mutação. A recombinação visa levar o processo evolutivo para “regiões promissoras” do espaço de busca, enquanto o operador de mutação serve para evitar convergência prematura para ótimos locais, introduzindo diversidade;
5. Como critério de parada, pode-se especificar previamente um número máximo de iterações ou um nível aceitável de aptidão em que não se note mudança significativa desse valor durante várias iterações.

Pode-se resumir as considerações acima no seguinte pseudo-código de um algoritmo genético genérico [34]:

AG genérico

Inicialize a população

Avalie indivíduos na população

Repita

Selecione indivíduos para reprodução

Aplique operadores de recombinação e mutação

Avalie indivíduos na população

Selecione indivíduos para sobreviver

Até critério de parada satisfeito

Fim

6.1.1 Exemplo de passo evolutivo

Em Zörnig [1], tem-se uma exemplificação de um passo evolutivo do AG, o qual adaptamos para o seguinte problema de minimização

$$\begin{aligned} \min \quad & -x\text{sen}(x/2) + 25, \\ \text{sujeito a } & 0 \leq x \leq 31 \end{aligned} \tag{6.1}$$

A figura 6.2 ilustra o gráfico da função

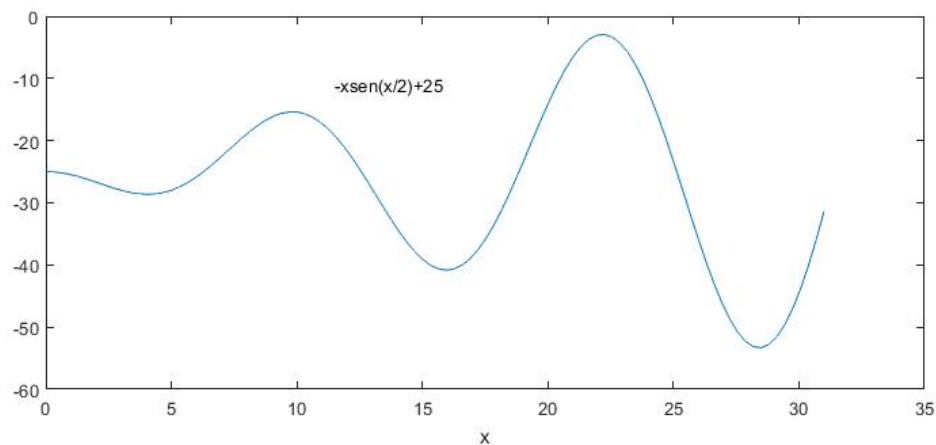


Figura 6.2: Ilustração da função (6.1).

Por simplicidade, considera-se somente valores inteiros de x , ou seja, escolhemos o conjunto espaço de busca $M = \{0, 1, \dots, 31\}$ como uma “aproximação discreta” do problema anterior, equação (6.1). Seja a população inicial $P(0)$ contendo quatro indivíduos x em codificação binária, conforme a tabela 6.1 a seguir:

x	genoma	aptidão f(x)	aptidão relativa
5	00101	-27,99	0,222
12	01100	-21,65	0,172
19	10011	-23,57	0,187
28	11100	-52,74	0,419
total		-125,95	1

Tabela 6.1: População $P(0)$

Para a reprodução com probabilidade proporcional à aptidão relativa, suponhamos que o indivíduo menos apto, 12, com 17,2% de aptidão relativa, não seja selecionado, e o indivíduo mais apto, 28, tenha sido escolhido duas vezes, formando os pares de “pais” $\{19, 28\}$ e $\{5, 28\}$. Um suposto cruzamento ocorre entre os pais 19 e 28 no segundo elemento do genoma, onde esta posição a ser trocada foi aleatória. Como resultado têm-se os genomas 11011 e 10100, correspondentes aos filhos 27 e 20. Supondo também que não ocorra cruzamento entre 5 e 28, cada um dos quatro filhos 5, 20, 27 e 28 pode sofrer mutação individual com probabilidade P_{mut} . Na hipótese de que somente os genomas de 5 e 20 sofram mutações nas posições 5 e 2, respectivamente, temos como resultado os genomas 00100 e 11100, correspondentes aos indivíduos 4 e 28.

A nova geração é apresentada na tabela 6.2, na qual observa-se que, neste passo evolutivo, a aptidão $f(x)$ é minimizada entre todos os os indivíduos. A aptidão média passou de -31,49 para -45,21.

x	genoma	aptidão f(x)	aptidão relativa
4	00100	-28,64	0,158
27	11011	-46,72	0,258
28	11100	-52,74	0,292
28	11100	-52,74	0,292
total		-125,95	1

Tabela 6.2: População P(1)

6.1.2 Características dos operadores genéticos do IED+GA

6.1.2.1 Codificação

Para representar os indivíduos que são as possíveis soluções do problema de otimização, os AG's utilizam, dentre outras, a codificação binária e a codificação real (decimal). A primeira, utilizada por Holland [35] em seus trabalhos, é a mais indicada para problemas com variáveis discretas, por outro lado, a codificação real é recomendada quando as variáveis são contínuas. Para o IED+GA foi escolhida a codificação real.

6.1.2.2 População inicial

Os elementos de uma população inicial, de modo mais geral, são escolhidos ao acaso do espaço de busca, através de funções pseudorrandômicas. Ao se definir o tamanho da população, alguns critérios devem ser verificados, pois, enquanto uma população grande pode obter uma solução global cobrindo melhor o espaço de busca, mas a um custo computacional elevado, uma população pequena, por sua vez, pode resultar em uma convergência prematura (solução local).

6.1.2.3 Operador de seleção

Para criar a próxima geração, elementos mais aptos da população são selecionados, de modo geral, proporcionalmente ao valor relativo da aptidão. Dessa forma, os melhores indivíduos da população têm mais chance de “reprodução” enquanto aqueles com aptidão

menor tendem a desaparecer [1].

Dentre as diferentes seleções possíveis, o IED+GA utiliza a seleção por roleta, na qual cada indivíduo ocupa uma porção da roleta de tamanho proporcional ao seu valor de aptidão. A tabela 6.3 e a figura 6.3 exemplificam este operador para problemas de maximização.

Individuo	aptidão	aptidão relativa (%)
1	42	12,3
2	90	26,4
3	23	6,7
4	136	39,9
5	50	14,7

Tabela 6.3: Exemplo de seleção por roleta.

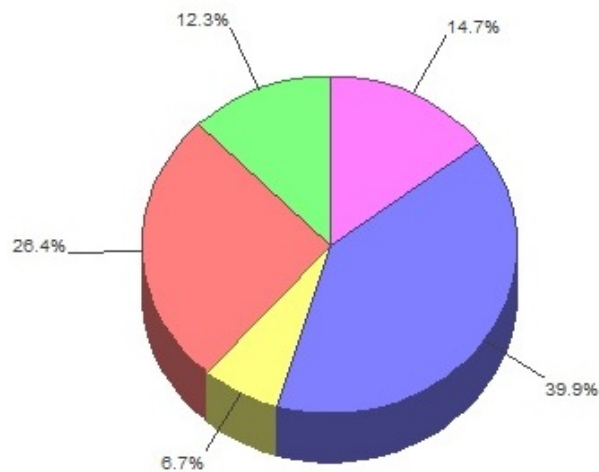


Figura 6.3: Representação gráfica da Tabela 6.3.

6.1.2.4 Operador de recombinação

Uma recombinação (*crossover*) entre dois indivíduos selecionados (pais) é realizado com probabilidade P_{cross} , trocando-se partes dos genomas. Este operador visa explorar o espaço de busca.

Para a formulação do IED+GA, foi escolhida a Recombinação Binária Simulada (*Simulated Binary Crossover - SBX*) cujas propriedades são [36]:

- A média dos pais é igual à média dos filhos após a operação de recombinação;
- Fator de propagação β : Na maioria dos eventos de recombinação os filhos tendem a ser parecidos com os pais, isto é, resulta em um fator de propagação

$$\beta = \left| \frac{c_1 - c_2}{p_1 - p_2} \right| \approx 1$$

onde p_1, p_2 são os pais e c_1, c_2 são os filhos.

Para manter a propriedade média, calculam-se os valores dos filhos do seguinte modo:

$$c_1 = \bar{x} - \frac{1}{2}\beta(p_2 - p_1)$$

$$c_2 = \bar{x} + \frac{1}{2}\beta(p_2 - p_1)$$

em que $\bar{x} = \frac{1}{2}\beta(p_1 + p_2)$, $p_2 > p_1$. Pode-se deduzir $\bar{c} = \bar{x}$.

Para se obter β no SBX, Deb e Agrawal (1995) propuseram uma função de distribuição probabilística:

$$c(\beta) = \begin{cases} 0.5(n+1)\beta^n, & \beta \leq 1 \\ \frac{0.5(n+1)}{\beta^{n+2}}, & \beta > 1 \end{cases}$$

onde valores menores de n geram filhos mais parecidos com seus pais.

Pode-se obter randomicamente o fator de propagação $\bar{\beta}$ a partir da geração de um número aleatório $u \in [0, 1]$. $\bar{\beta}$ é calculado no gráfico da distribuição de probabilidade em função de β , em que a área sob a curva é igual u . Neste caso, os valores dos filhos são calculados como:

$$c_1 = \bar{x} - \frac{1}{2}\bar{\beta}(p_2 - p_1)$$

$$c_2 = \bar{x} + \frac{1}{2}\bar{\beta}(p_2 - p_1)$$

Mais informações sobre o SBX em [36].

6.1.2.5 Operador de mutação

Este operador serve para evitar convergência “prematura” para ótimos locais, possibilitando a busca em regiões do espaço ainda não exploradas. Elementos do genoma com baixa probabilidade P_{mut} são trocados.

O IED+GA possui o operador de mutação polinomial (Deb e Goyal, 1996), que utiliza distribuição de probabilidade polinomial para perturbar uma solução de uma variável contínua na sua proximidade. Um fator de perturbação δ é definido como:

$$\delta = \frac{c - p}{\Delta_{max}},$$

onde c é o valor mutado, Δ_{max} é uma perturbação máxima permitida para o pai p .

O valor da mutação é calculado como uma distribuição de probabilidade que depende da perturbação δ :

$$P(\delta) = 0.5(n + 1)(1 - |\delta|)^n.$$

A distribuição acima é válida no intervalo $\delta \in (-1, 1)$. É gerado um número aleatório $u \in (0, 1)$ e calcula-se o fator de perturbação δ correspondente:

$$\bar{\delta} = \begin{cases} (2u)^{\frac{1}{n+1}} - 1, & u < 0.5 \\ 1 - [2(1 - u)]^{\frac{1}{n+1}}, & u \geq 0.5 \end{cases}$$

O novo indivíduo é então calculado

$$c = p + \bar{\delta}\Delta_{max}.$$

Mais detalhes sobre mutação polinomial em [37].

6.2 O ALGORITMO IED+GA

Para a resolução do subproblema, ou seja, minimizar a Lagrangiana em cada iteração do método, as versões anteriores do IED necessitam do cálculo de subgradientes, como no NFDNA, além de dependerem fortemente dos valores dos parâmetros iniciais. Esses aspectos motivaram a hibridização do IED com o GA.

Apresentamos, nesta seção, uma estratégia diferente que não utiliza derivadas e que não possui dependência significativa dos parâmetros iniciais u_0 e c_0 .

Esta versão, denominada IED+GA, é mostrada nos fluxogramas seguintes:

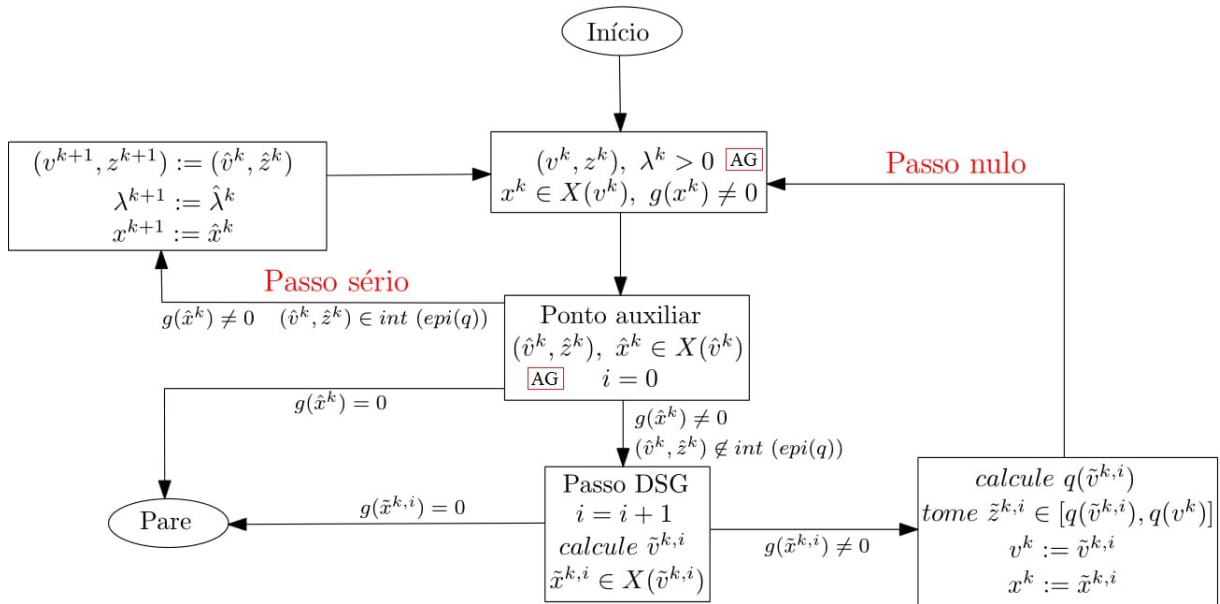


Figura 6.4: Fluxograma IED+AG.

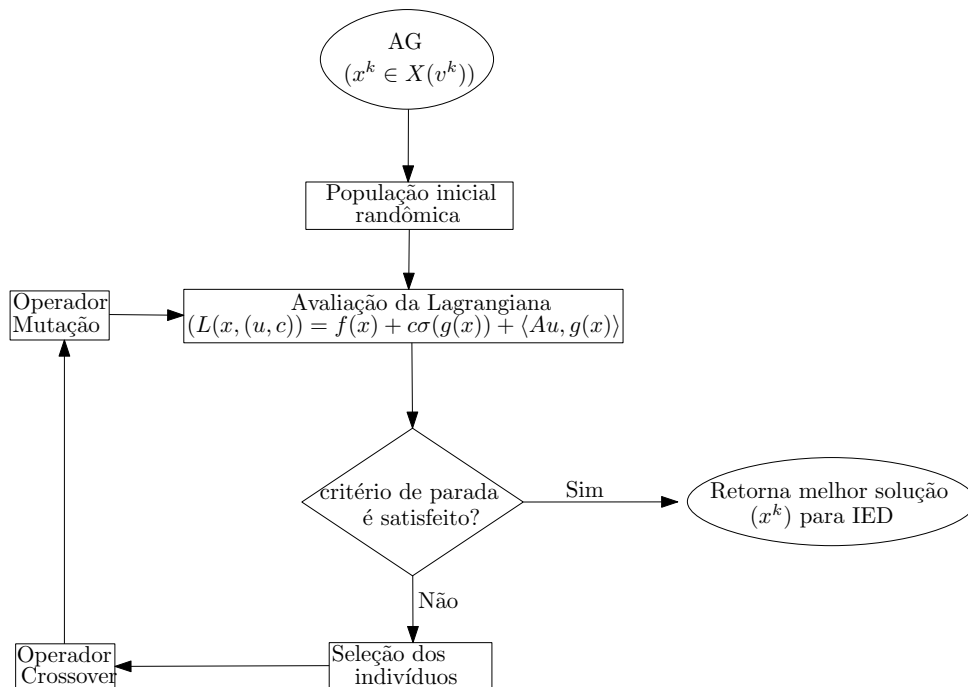


Figura 6.5: Fluxograma do AG utilizado pelo IED.

Algoritmo 2: O Algoritmo Genético para minimizar a função Lagrangiana.

Input: Parâmetros $popsiz$, $gensiz$, u , c ;
Output: A melhor solução encontrada em todas as gerações;
 $gen = 1$;
 $Pop[gen] = \text{InicializePopulação}()$;
Avalie ($L(x, (u, c)) = f(x) + c\sigma(g(x)) + \langle Au, g(x) \rangle$, $Pop[gen]$);
while $gen \leq gensiz$ **do**
 $selPop[gen] = \text{Seleção}(Pop[gen])$;
 $newPop[gen] = \text{Crossover}(selPop[gen])$;
 $\text{Muta\c{c}\~{a}o}(newPop[gen])$;
 Avalie ($L(x, (u, c)) = f(x) + c\sigma(g(x)) + \langle Au, g(x) \rangle$, $newPop[gen]$);
 $Pop[gen + 1] = \text{Novagera\c{c}\~{a}o}(newPop[gen], Pop[gen])$;
 $gen = gen + 1$;
end

Para os operadores de *crossover* e de mutação do AG foram utilizados respectivamente *Simulated Binary Crossover(SBX)* [38] e *Polynomial mutation* [37], além de codificação real.

6.2.1 Experimentos Numéricos

Nesta seção são apresentados os resultados numéricos da versão IED+GA, em comparação com a versão original do método IED.

Na construção da função Lagrangiana, utilizou-se as seguintes funções σ : $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$, $e^{\|\cdot\|_2} - 1$, $e^{\|\cdot\|_1} - 1$, $\sqrt{m} (e^{\|\cdot\|_\infty} - 1)$ and $\max(\|\cdot\|_2, \|\cdot\|_2^2)$. Optou-se pela matriz $A=I$ (Identidade) para o termo linear da Lagrangiana.

A tolerância do critério de parada do método foi de 10^{-6} , tendo sido realizadas 10 execuções do IED+GA para cada problema.

6.2.1.1 Descrição dos problemas numéricos

A função objetivo e as restrições dos 20 problemas-teste de otimização, retirados de Schittkowski collection [33], estão descritas a seguir:

GLR-P1-1:

$$\min f(x) = -32,174(255\ln((x_1 + x_2 + x_3 + 0,03)/(0,09x_1 + x_2 + x_3 + 0,03)) + 280\ln((x_2 + x_3 + 0,03)/(0,07x_2 + x_3 + 0,03)) + 290\ln((x_3 + 0,03)/(0,13x_3 + 0,03)))$$

sujeito a

$$x_1 + x_2 + x_3 = 0$$

$$0 \leq x_i \leq 1, \quad i = 1, 2, 3$$

GQR-P1-1:

$$\begin{aligned} \min f(x) = & -75,196 + 3,8112x_1 + 0,0020567x_1^3 - 1,0345(e - 5)x_1^4 + 6,8306x_2 - \\ & 0,030234x_1x_2 + 1,28134(e - 3)x_2x_1^2 + 2,266(e - 7)x_1^4x_2 - 0,25645x_2^2 + 0,0034604x_2^3 - \\ & 1,3514(e - 5)x_2^4 + 28,106/(x_2 + 1) + 5,2375e - 6x_1^2x_2^2 + 6,3(e - 8)x_1^3x_2^2 - 7(e - 10)x_1^3x_2^3 - \\ & 3,405(e - 4)x_1x_2^2 + 1,6638(e - 6)x_1x_2^3 + 2,8673\exp(0,0005x_1x_2) - 3,5256(e - 5)x_1^3x_2 \end{aligned}$$

sujeito a

$$x_1x_2 - 700 \geq 0$$

$$x_2 - x_1^2/125 \geq 0$$

$$(x_2 - 50)^2 - 5(x_1 - 55) \geq 0$$

$$0 \leq x_1 \leq 75$$

$$0 \leq x_2 \leq 65$$

PPR-P1-2:

$$\min f(x) = 5x_1 + 50000/x_1 + 20x_2 + 72000/x_2 + 10x_3 + 144000/x_3$$

sujeito a

$$1 - 4/x_1 - 32/x_2 - 120/x_3$$

$$1.10^{-5} \leq x_i, \quad i = 1, 2, 3$$

PQR-T1-7:

$$\min f(x) = -x_1x_2x_3$$

sujeito a

$$-x_1^2 - 2x_2^2 - 4x_3^2 + 48 \geq 0$$

SQR-P1-1:

$$\min f(x) = \sum_{i=1}^{44} f_i(x)^2, \quad f_i(x) = b_i - x_1 - (0,49 - x_1)\exp(-x_2(a_i - 8)), \quad i = 1, \dots, 44$$

sujeito a

$$0,49x_2 - x_1x_2 - 0,09 \geq 0$$

$$0,4 \leq x_1$$

$$-4 \leq x_2$$

LGR-P1-1:

$$\min f(x) = 0,2x_3 - 0,8x_1$$

sujeito a

$$x_2 - \exp(x_1) \geq 0$$

$$x_3 - \exp(x_2) \geq 0$$

$$0 \leq x_1 \leq 100$$

$$0 \leq x_2 \leq 100$$

$$0 \leq x_3 \leq 10$$

QBR-T1-1:

$$\min f(x) = x_2 + 10^{-5}(x_2 - x_1)^2$$

sujeito a

$$0 \leq x_2$$

PBR-T1-1:

$$\min f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

sujeito a

$$-x_2 - 1,5 \leq 0$$

PBR-T1-2:

$$\min f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

sujeito a

$$1,5 \leq x_2$$

QQR-P1-3:

$$\min f(x) = (x_1 - x_2)^2 + (x_1 + x_2 - 10)^2/9 + (x_3 - 5)^2$$

sujeito a

$$48 - x_1^2 - x_2^2 - x_3^2 \geq 0$$

$$-4,5 \leq x_i \leq 4,5, \quad i = 1, 2$$

$$-5 \leq x_3 \leq 5$$

QQR-T1-6:

$$\min f(x) = (x_1 - 2)^2 + (x_2 - 1)^2$$

sujeito a

$$x_1^2 - x_2 \leq 0$$

$$x_1 + x_2 - 2 \leq 0$$

PLR-T1-1:

$$\min f(x) = 1/(27\sqrt{3})((x_1 - 3)^2 - 9)x_2^3$$

sujeito a

$$x_1/\sqrt{3} - x_2 \geq 0$$

$$x_1 + \sqrt{3}x_2 \geq 0$$

$$-x_1 + \sqrt{3}x_2 - 6 \geq 0$$

$$0 \leq x_1$$

$$0 \leq x_2$$

PBR-T1-3:

$$\min f(x) = 1/3(x_1 + 1)^3 + x_2$$

sujeito a

$$1 \leq x_1$$

$$0 \leq x_2$$

QLR-T1-1:

$$\min f(x) = 0,01x_1^2 + x_2^2 - 100$$

sujeito a

$$10x_1 - x_2 - 10 \geq 0$$

$$2 \leq x_1 \leq 50$$

$$-50 \leq x_2 \leq 50$$

PQR-T1-1:

$$\min f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

sujeito a

$$x_1 x_2 - 1 \geq 0$$

$$x_1 + x_2^2 \geq 0$$

$$x_1 \leq 0,5$$

QQR-T1-3:

$$\min f(x) = 0,5x_1^2 + x_2^2 - x_1 x_2 - 7x_1 - 7x_2$$

sujeito a

$$-25 + 4x_1^2 + x_2^2 \leq 0$$

GBR-T1-1:

$$\min f(x) = \sin(x_1 + x_2) + (x_1 - x_2)^2 - 1,5x_1 + 2,5x_2 + 1$$

sujeito a

$$-1,5 \leq x_1 \leq 4$$

$$-3 \leq x_2 \leq 3$$

PQR-T1-4:

$$\min f(x) = (x_1 - 10)^3 + (x_2 - 20)^3$$

sujeito a

$$-100 + (x_1 - 5)^2 + (x_2 - 5)^2 \geq 0$$

$$-(x_2 - 5)^2 - (x_1 - 6)^2 + 82.81 \geq 0$$

$$13 \leq x_1 \leq 100$$

$$0 \leq x_2 \leq 100$$

QQR-T1-2:

$$\min f(x) = (x_1 - 5)^2 + x_2^2 - 25$$

sujeito a

$$-x_1^2 + x_2^2 \geq 0$$

QPR-T1-1:

$$\min f(x) = (x_1 - 2)^2 + x_2^2$$

sujeito a

$$(1 - x_1)^3 - x_2 \geq 0$$

$$0 \leq x_1$$

$$0 \leq x_2$$

6.2.1.2 Comparação dos resultados

A tabela 6.4 apresenta os valores iniciais dos parâmetros c e u , o tamanho da população, o número de gerações, os limites superior e inferior do algoritmo genético e a solução ótima de referência dos problemas. Para os primeiros seis problemas, a fim de comparação, utilizou-se os mesmos valores de c_0 e u_0 descritos em [5]; para os demais, c_0 e u_0 são aleatórios.

Problemas	c_0	u_0	População	Gerações	Limite inferior	Limite superior	Valor ótimo
GLR-P1-1	50	[0 0 0 0]	50	300	[0.01, 0.01, 0.01]	[1, 1, 1]	-2.6273×10^4
GQR-P1-1	1	[1 1 1 1 1 1 1]	50	100	[10, 10]	[60, 60]	-7.8028
PPR-P1-2	0.1	[1 1 1 1]	50	200	[0, 0, 0]	[300, 300, 300]	6.2998×10^3
PQR-T1-7	1.5	[1 1 1 1 1 1 1]	50	100	[-5, -5, -5]	[5, 5, 5]	-22.6274
SQR-P1-1	0.2	[0 0 0]	50	100	[-10, -10]	[10, 10]	0.0285
LGR-P1-1	0.1	[0 0 0 0 0 0 0 0]	50	100	[0, 0, 0]	[5, 5, 5]	0.5181
QBR-T1-1	1	[1 1 1 1 1 1 1 1]	50	100	[0, 0, 0]	[5, 5, 5]	0
PBR-T1-1	1	[1]	50	100	[-2, -2]	[2, 2]	0
PBR-T1-2	1	[1]	50	100	[-2, -2]	[2, 2]	0.0504
QQR-P1-3	1	[1 1 1 1 1 1 1]	50	100	[0, 0, 0]	[5, 5, 5]	0.9535
QQR-T1-6	1	[1 1]	50	100	[-2, -2]	[2, 2]	1
PLR-T1-1	1	[1 1 1 1 1]	50	100	[1, 1]	[4, 4]	-1
PBR-T1-3	1	[1 1]	50	100	[-1, -1]	[2, 2]	2.6666
QLR-T1-1	1	[1 1 1 1 1]	50	100	[-1, -1]	[3, 3]	-99.96
PQR-T1-1	5	[1 1 1]	50	100	[0, 0]	[3, 3]	306.5
QQR-T1-3	0.1	[1]	20	50	[1, 1]	[4, 4]	-30
GBR-T1-1	1	[1 1 1 1]	50	100	[-2, -2]	[0, 0]	-1.9132
PQR-T1-4	1	[1 1 1 1 1 1]	50	100	[0, 0]	[15, 15]	-6961.8138
QQR-T1-2	1	[1]	20	50	[0, 0]	[2, 2]	-8.4984
QPR-T1-1	0.1	[1 1 1]	50	100	[-1, -1]	[2, 2]	1

Tabela 6.4: Parâmetros do IED.

Os resultados numéricos obtidos por ambas as versões, IED+fmin e IED+GA, estão apresentados a seguir nas tabelas 6.5 até 6.24. Para a versão IED+GA, o melhor e o pior valores, a média e o desvio padrão foram calculados. As soluções do IED+fmin estão em uma coluna única. Soluções não encontradas estão indicadas por uma linha horizontal na tabela.

$\sigma(\cdot)$	GLR-P1-1				
	IED+fmin	IED+GA			
	solução	melhor	pior	média	desvio
$\ \cdot\ _1$	-2.6273×10^4	-2.6272×10^4	-2.6267×10^4	-2.6271×10^4	2.0661
$\ \cdot\ _2$	-2.6273×10^4	-2.6272×10^4	-2.6267×10^4	-2.6271×10^4	2.0661
$\sqrt{m} \ \cdot\ _\infty$	—	-2.6272×10^4	-2.6267×10^4	-2.6271×10^4	2.0661
$e^{\ \cdot\ _2} - 1$	-2.6273×10^4	-2.6273×10^4	-2.6270×10^4	-2.6272×10^4	0.8555
$e^{\ \cdot\ _1} - 1$	-2.6273×10^4	-2.6273×10^4	-2.6270×10^4	-2.6272×10^4	0.8555
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	—	-2.6272×10^4	-2.6225×10^4	-2.6261×10^4	14.5763
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	—	-2.6273×10^4	-2.4727×10^4	-2.6049×10^4	477.8463

Tabela 6.5: Resultados comparativos do problema GLR-P1-1.

$\sigma(\cdot)$	GQR-P1-1				
	IED+fmin	IED+GA			
	solução	melhor	pior	média	desvio
$\ \cdot\ _1$	—	-7.7978	-6.7495	-7.5575	0.4268
$\ \cdot\ _2$	—	-7.7978	-6.7495	-7.5572	0.4276
$\sqrt{m} \ \cdot\ _\infty$	-7.8028	-7.7984	-6.7495	-7.0554	0.4927
$e^{\ \cdot\ _2} - 1$	-7.8028	-7.7925	-6.7495	-7.4544	0.4872
$e^{\ \cdot\ _1} - 1$	-7.8028	-7.8012	-7.7443	-7.7801	0.0171
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	-7.8028	-7.7985	-6.7495	-7.4601	0.4914
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	-7.8028	-7.8024	-6.7495	-7.3560	0.5228

Tabela 6.6: Resultados comparativos do problema GQR-P1-1.

$\sigma(\cdot)$	PPR-P1-2				
	IED+fmin		IED+GA		
	solução	melhor	pior	média	desvio
$\ \cdot\ _1$	6.2998×10^3	6.3002×10^3	6.3018×10^3	6.3008×10^3	0.4452
$\ \cdot\ _2$	6.2998×10^3	6.3002×10^3	6.3018×10^3	6.3008×10^3	0.4452
$\sqrt{m} \ \cdot\ _\infty$	6.2998×10^3	6.3001×10^3	6.3035×10^3	6.3015×10^3	1.0803
$e^{\ \cdot\ _2} - 1$	—	6.3000×10^3	6.3023×10^3	6.3007×10^3	0.7107
$e^{\ \cdot\ _1} - 1$	—	6.3000×10^3	6.3023×10^3	6.3007×10^3	0.7107
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	—	6.3000×10^3	6.3053×10^3	6.3012×10^3	1.6442
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	6.2999×10^3	6.2999×10^3	6.3045×10^3	6.3013×10^3	1.5406

Tabela 6.7: Resultados comparativos do problema PPR-P1-2.

$\sigma(\cdot)$	PQR-T1-7				
	IED+fmin		IED+GA		
	solução	melhor	pior	média	desvio
$\ \cdot\ _1$	—	-22.5467	-22.1224	-22.4160	0.1450
$\ \cdot\ _2$	—	-22.6033	-22.1466	-22.4274	0.1383
$\sqrt{m} \ \cdot\ _\infty$	-22.6274	-22.6233	-22.3691	-22.5153	0.1151
$e^{\ \cdot\ _2} - 1$	-22.6262	-22.5424	-22.0668	-22.3474	0.1526
$e^{\ \cdot\ _1} - 1$	-22.6270	-22.5468	-21.9238	-22.2471	0.1891
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	-22.6274	-22.5215	21.6514	-17.8111	13.8676
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	-22.6156	-22.5704	-22.0331	-22.3381	0.1683

Tabela 6.8: Resultados comparativos do problema PQR-T1-7.

$\sigma(\cdot)$	SQR-P1-1				
	IED+fmin	IED+GA			
	solução	melhor	pior	média	desvio
$\ \cdot\ _1$	0.0285	0.0285	0.0287	0.0286	8.4993×10^{-5}
$\ \cdot\ _2$	0.0285	0.0285	0.0287	0.0286	8.2863×10^{-5}
$\sqrt{m} \ \cdot\ _\infty$	0.0285	0.0285	0.0293	0.0287	2.3577×10^{-4}
$e^{\ \cdot\ _2} - 1$	0.0285	0.0285	0.0293	0.0287	5.7703×10^{-4}
$e^{\ \cdot\ _1} - 1$	0.0285	0.0285	0.0306	0.0289	6.5958×10^{-4}
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	0.0285	0.0285	0.0290	0.0286	1.8364×10^{-4}
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	0.0285	0.0285	0.0289	0.0286	1.2819×10^{-4}

Tabela 6.9: Resultados comparativos do problema SQR-P1-1.

$\sigma(\cdot)$	LGR-P1-1				
	IED+fmin	IED+GA			
	solução	melhor	pior	média	desvio
$\ \cdot\ _1$	0.5187	0.5234	0.5407	0.5314	0.0061
$\ \cdot\ _2$	—	0.5216	0.5427	0.5301	0.0083
$\sqrt{m} \ \cdot\ _\infty$	0.5228	0.5213	0.5726	0.5364	0.0157
$e^{\ \cdot\ _2} - 1$	0.5199	0.5213	0.5726	0.5364	0.0157
$e^{\ \cdot\ _1} - 1$	0.5195	0.5232	0.5622	0.5384	0.0125
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	0.5183	0.5239	0.5677	0.5390	0.0161
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	0.5182	0.5201	0.5488	0.5313	0.0090

Tabela 6.10: Resultados comparativos do problema LGR-P1-1.

$\sigma(\cdot)$	QBR-T1-1				
	IED+fmin		IED+GA		
	solução	melhor	pior	média	desvio
$\ \cdot\ _1$	—	5.4134×10^{-7}	1.5704×10^{-4}	6.7404×10^{-5}	5.3498×10^{-5}
$\ \cdot\ _2$	—	5.4134×10^{-7}	1.5704×10^{-4}	6.7404×10^{-5}	5.3498×10^{-5}
$\sqrt{m} \ \cdot\ _\infty$	—	5.4134×10^{-7}	1.5704×10^{-4}	6.7404×10^{-5}	5.3498×10^{-5}
$e^{\ \cdot\ _2} - 1$	3.668×10^{-17}	7.2718×10^{-6}	2.0872×10^{-4}	9.4788×10^{-5}	8.2756×10^{-5}
$e^{\ \cdot\ _1} - 1$	3.668×10^{-17}	7.2718×10^{-6}	2.0872×10^{-4}	9.4788×10^{-5}	8.2756×10^{-5}
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	3.668×10^{-17}	7.2718×10^{-6}	2.0872×10^{-4}	9.4788×10^{-5}	8.2756×10^{-5}
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	2.4563×10^{-14}	5.4134×10^{-7}	1.5704×10^{-4}	6.7404×10^{-5}	5.3498×10^{-5}

Tabela 6.11: Resultados comparativos do problema QBR-T1-1.

$\sigma(\cdot)$	PBR-T1-1				
	IED+fmin		IED+GA		
	solução	melhor	pior	média	desvio
$\ \cdot\ _1$	-9.0273×10^{-14}	0	0.0134	0.0027	0.0052
$\ \cdot\ _2$	-9.0273×10^{-14}	0	0.0134	0.0027	0.0052
$\sqrt{m} \ \cdot\ _\infty$	-9.0273×10^{-14}	0	0.0134	0.0027	0.0052
$e^{\ \cdot\ _2} - 1$	-9.0273×10^{-14}	0	0.0134	0.0027	0.0052
$e^{\ \cdot\ _1} - 1$	-9.0273×10^{-14}	0	0.0134	0.0027	0.0052
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	-9.0273×10^{-14}	0	0.0134	0.0027	0.0052
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	-9.0273×10^{-14}	0	0.0134	0.0027	0.0052

Tabela 6.12: Resultados comparativos do problema PBR-T1-1.

$\sigma(\cdot)$	PBR-T1-2				
	IED+fmin	IED+GA			
	solução	melhor	pior	média	desvio
$\ \cdot\ _1$	0.0504	0.0506	0.0677	0.0536	0.0052
$\ \cdot\ _2$	0.0504	0.0506	0.0677	0.0536	0.0052
$\sqrt{m} \ \cdot\ _\infty$	0.0504	0.0506	0.0677	0.0536	0.0052
$e^{\ \cdot\ _2} - 1$	0.0504	0.0508	0.0839	0.0593	0.0123
$e^{\ \cdot\ _1} - 1$	0.0504	0.0508	0.0839	0.0593	0.0123
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	0.0504	0.0508	0.0839	0.0593	0.0123
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	0.0504	0.0505	0.0790	0.0545	0.0088

Tabela 6.13: Resultados comparativos do problema PBR-T1-2.

$\sigma(\cdot)$	QQR-P1-3				
	IED+fmin	IED+GA			
	solução	melhor	pior	média	desvio
$\ \cdot\ _1$	0.9535	0.9576	0.9837	0.9676	0.0099
$\ \cdot\ _2$	0.9535	0.9565	0.9853	0.9637	0.0890
$\sqrt{m} \ \cdot\ _\infty$	0.9535	0.9550	0.9881	0.9665	0.0115
$e^{\ \cdot\ _2} - 1$	0.9535	0.9550	0.9684	0.9603	0.0039
$e^{\ \cdot\ _1} - 1$	0.9535	0.9541	0.9860	0.9652	0.0096
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	-25.0694	0.9558	0.9683	0.9628	0.0040
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	0.9535	0.9568	1.0124	0.9708	0.0185

Tabela 6.14: Resultados comparativos do problema QQR-P1-3.

$\sigma(\cdot)$	QQR-T1-6				
	IED+fmin	IED+GA			
	solução	melhor	pior	média	desvio
$\ \cdot\ _1$	1	1.0002	1.0185	1.0052	0.0057
$\ \cdot\ _2$	1	1.0023	1.0179	1.0101	0.0044
$\sqrt{m} \ \cdot\ _\infty$	1	1.0003	1.0404	1.0112	0.0125
$e^{\ \cdot\ _2} - 1$	1	1.0005	1.0287	1.0126	0.0093
$e^{\ \cdot\ _1} - 1$	1	1.0005	1.0400	1.0122	0.0137
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	1	1.0002	1.0366	1.0133	0.0106
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	1	1.0002	1.0097	1.0046	0.0037

Tabela 6.15: Resultados comparativos do problema QQR-T1-6.

$\sigma(\cdot)$	PLR-T1-1				
	IED+fmin	IED+GA			
	solução	melhor	pior	média	desvio
$\ \cdot\ _1$	—	-0.9976	-0.9729	-0.9915	0.0071
$\ \cdot\ _2$	—	-0.9966	-0.9817	-0.9892	0.0052
$\sqrt{m} \ \cdot\ _\infty$	—	-0.9984	-0.9663	-0.9899	0.0102
$e^{\ \cdot\ _2} - 1$	-1	-0.9974	-0.9806	-0.9916	0.0057
$e^{\ \cdot\ _1} - 1$	-1	-0.9966	-0.9805	-0.9925	0.0048
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	-1	-0.9993	-0.9793	-0.9926	0.0061
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	-1	-0.9996	-0.9758	-0.9920	0.0073

Tabela 6.16: Resultados comparativos do problema PLR-T1-1.

$\sigma(\cdot)$	PBR-T1-3				
	IED+fmin	IED+GA			
	solução	melhor	pior	média	desvio
$\ \cdot\ _1$	—	2.6686	2.7135	2.6827	0.0134
$\ \cdot\ _2$	—	2.6698	2.6855	2.6743	0.0050
$\sqrt{m} \ \cdot\ _\infty$	—	2.6697	2.6909	2.6784	0.0079
$e^{\ \cdot\ _2} - 1$	2.6667	2.6692	2.6947	2.6784	0.0082
$e^{\ \cdot\ _1} - 1$	2.6667	2.6685	2.7022	2.6809	0.0110
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	2.6667	2.6687	2.7059	2.6909	0.0125
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	2.6667	2.6684	2.7125	2.6847	0.0150

Tabela 6.17: Resultados comparativos do problema PBR-T1-3.

$\sigma(\cdot)$	QLR-T1-1				
	IED+fmin	IED+GA			
	solução	melhor	pior	média	desvio
$\ \cdot\ _1$	-99.9600	-99.9600	-99.9597	-99.9599	8.9350×10^{-5}
$\ \cdot\ _2$	-99.9600	-99.9600	-99.9595	-99.9598	1.4311×10^{-4}
$\sqrt{m} \ \cdot\ _\infty$	-99.9600	-99.9600	-99.9599	-99.9599	5.1418×10^{-5}
$e^{\ \cdot\ _2} - 1$	-99.9600	-99.9600	-99.9594	-99.9599	1.5289×10^{-4}
$e^{\ \cdot\ _1} - 1$	-99.9600	-99.9600	-99.9595	-99.9599	1.3253×10^{-4}
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	-99.9600	-99.9600	-99.9594	-99.9598	1.8280×10^{-4}
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	-99.9600	-99.9599	-99.9598	-99.9599	3.7864×10^{-5}

Tabela 6.18: Resultados comparativos do problema QLR-T1-1.

$\sigma(\cdot)$	PQR-T1-1				
	IED+fmin	IED+GA			
	solução	melhor	pior	média	desvio
$\ \cdot\ _1$	306.5	306.5146	326.6933	310.9483	6.2614
$\ \cdot\ _2$	306.5	308.7375	329.3517	312.3063	6.1269
$\sqrt{m} \ \cdot\ _\infty$	306.5	307.5067	327.2424	311.6510	5.8228
$e^{\ \cdot\ _2} - 1$	306.5	308.9435	317.7634	312.9006	2.7421
$e^{\ \cdot\ _1} - 1$	306.5	306.5408	315.6229	310.7872	3.0505
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	306.5	307.8027	316.7068	312.1473	2.7962
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	306.5	307.9935	315.4329	311.4265	2.4128

Tabela 6.19: Resultados comparativos do problema PQR-T1-1.

$\sigma(\cdot)$	QQR-T1-3				
	IED+fmin	IED+GA			
	solução	melhor	pior	média	desvio
$\ \cdot\ _1$	—	-29.9998	-29.8286	-29.9714	0.0532
$\ \cdot\ _2$	—	-29.9998	-29.8286	-29.9714	0.532
$\sqrt{m} \ \cdot\ _\infty$	—	-29.9998	-29.8286	-29.9714	0.0532
$e^{\ \cdot\ _2} - 1$	-30	-29.9972	-29.3533	-29.8566	0.2162
$e^{\ \cdot\ _1} - 1$	-30	-29.9972	-29.3533	-29.8566	0.2162
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	-30	-29.9972	-29.3533	-29.8566	0.2162
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	-30	-29.9960	-29.2340	-29.8780	0.2316

Tabela 6.20: Resultados comparativos do problema QQR-T1-3.

$\sigma(\cdot)$	GBR-T1-1				
	IED+fmin	IED+GA			
	solução	melhor	pior	média	desvio
$\ \cdot\ _1$	-1.9132	-1.9132	-1.9132	-1.9132	5.1999×10^{-7}
$\ \cdot\ _2$	-1.9132	-1.9132	-1.9132	-1.9132	5.1999×10^{-7}
$\sqrt{m} \ \cdot\ _\infty$	-1.9132	-1.9132	-1.9132	-1.9132	1.4981×10^{-6}
$e^{\ \cdot\ _2} - 1$	-1.9132	-1.9132	-1.9132	-1.9132	9.3127×10^{-7}
$e^{\ \cdot\ _1} - 1$	-1.9132	-1.9132	-1.9132	-1.9132	9.3127×10^{-7}
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	-1.9132	-1.9132	-1.9132	-1.9132	1.2166×10^{-6}
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	-1.9132	-1.9132	-1.9132	-1.9132	5.1999×10^{-7}

Tabela 6.21: Resultados comparativos do problema GBR-T1-1.

$\sigma(\cdot)$	PQR-T1-4				
	IED+fmin	IED+GA			
	solução	melhor	pior	média	desvio padrão
$\ \cdot\ _1$	—	-6791.5	173850	1237.6	5.6747×10^4
$\ \cdot\ _2$	—	-6936.9	1792.1	-5153.5	2.8982×10^3
$\sqrt{m} \ \cdot\ _\infty$	—	-6800.2	-2548.9	-5962.7	1.3694×10^3
$e^{\ \cdot\ _2} - 1$	-6961.8	-6874.3	-6258.0	-6737.2	188.5621
$e^{\ \cdot\ _1} - 1$	-6961.8	-6915.8	-1642.2	-6165.1	1606.1
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	-6961.8	-6777.9	-6159.4	-6590.4	222.8402
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	-6961.8	-6958.0	-1735.3	-6274.8	1609.0

Tabela 6.22: Resultados comparativos do problema PQR-T1-4.

$\sigma(\cdot)$	QQR-T1-2				
	IED+fmin		IED+GA		
	solução	melhor	pior	média	desvio padrão
$\ \cdot\ _1$	-8.4985	-8.4958	-8.3391	-8.4729	0.0481
$\ \cdot\ _2$	-8.4985	-8.4958	-8.3391	-8.4729	0.0481
$\sqrt{m} \ \cdot\ _\infty$	-8.4985	-8.4958	-8.3391	-8.4729	0.0481
$e^{\ \cdot\ _2} - 1$	-8.4985	-8.4958	-7.5694	-8.2675	0.3201
$e^{\ \cdot\ _1} - 1$	-8.4985	-8.4958	-7.5694	-8.2675	0.3201
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	-8.4985	-8.4958	-7.5694	-8.2675	0.3201
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	-8.4985	-8.4980	-8.0434	-8.3737	0.1564

Tabela 6.23: Resultados comparativos do problema QQR-T1-2.

$\sigma(\cdot)$	QPR-T1-1				
	IED+fmin		IED+GA		
	solução	melhor	pior	média	desvio padrão
$\ \cdot\ _1$	—	1.1349	1.4578	1.2706	0.0846
$\ \cdot\ _2$	—	1.0632	1.4943	1.2132	0.1356
$\sqrt{m} \ \cdot\ _\infty$	—	1.1317	1.6091	1.2670	0.1421
$e^{\ \cdot\ _2} - 1$	0.9978	1.1316	1.4839	1.2715	0.1220
$e^{\ \cdot\ _1} - 1$	—	1.0729	1.5440	1.2488	0.1375
$\sqrt{m} (e^{\ \cdot\ _\infty} - 1)$	0.9979	1.0818	1.4799	1.2300	0.1555
$\max(\ \cdot\ _2, \ \cdot\ _2^2)$	0.9987	1.0967	1.6843	1.2762	0.1853

Tabela 6.24: Resultados comparativos do problema QPR-T1-1.

6.2.1.3 Análise dos experimentos

- A análise dos resultados numéricos apresentados dos 20 experimentos anteriores mostra que os mesmos foram compatíveis com as soluções ótimas de referência. Entretanto, ressalta-se o fato de que, em mais da metade destes, a versão IED+fmin falhou em algumas funções σ .

- O IED+GA preencheu todas as funções de penalização σ , apresentando em geral um baixo desvio padrão. Isto denota um comportamento robusto neste conjunto de testes.

7 APLICAÇÃO DO MÉTODO IED NA RESOLUÇÃO DE PROBLEMAS DE ENGENHARIA

Estudamos o comportamento do método IED, a partir de sua versão IED+GA, na resolução de problemas reais de engenharia mecânica.

As outras versões não estão contempladas neste capítulo, uma vez que, em testes preliminares, não obtiveram soluções para estes problemas.

A seguir são mostradas as formulações dos quatro problemas, suas funções objetivo e o conjunto de restrições.

Mola sob tração/compressão

O objetivo deste problema é minimizar o volume V de uma mola sob tração/compressão, conforme figura seguinte:

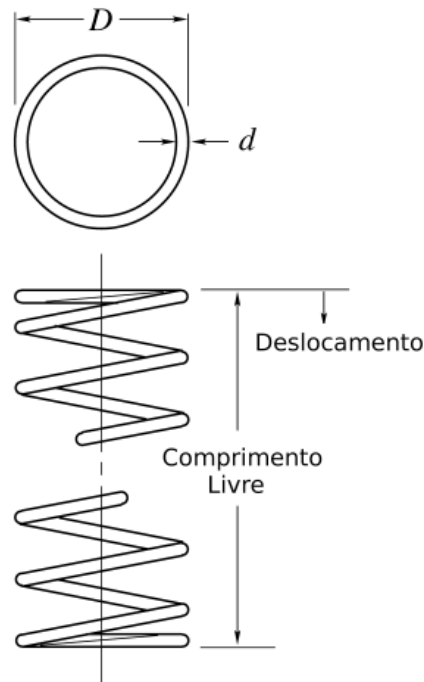


Figura 7.1: Mola sob tração/compressão.

As variáveis de projeto são o número de espirais ativas da mola ($N \in [2, 15]$), o diâmetro de cada volta ($D \in [0.25, 1.3]$) e o diâmetro do arame ($d \in [0.05, 2]$). O volume V (função objetivo) e as restrições mecânicas g_i , $i = 1, \dots, 4$ estão descritas a seguir, de acordo com [39]:

$$\begin{aligned}
 V &= (N + 2)Dd^2 \\
 g_1 &= 1 - \frac{D^3N}{71785d^4} \leq 0 \\
 g_2 &= \frac{4D^2 - dD}{12566(Dd^3 - d^4)} + \frac{1}{5108d^2} - 1 \leq 0 \\
 g_3 &= 1 - \frac{140.45d}{D^2N} \leq 0 \\
 g_4 &= \frac{D + d}{1.5} - 1 \leq 0
 \end{aligned}$$

onde

$$0.05 \leq d \leq 2, \quad 0.25 \leq D \leq 1.3, \quad 2 \leq N \leq 15.$$

Redutor de velocidade

Neste problema o objetivo é minimizar o peso W de um redutor de velocidade, mostrado na figura 7.2. As variáveis de projeto são a largura da face ($b = x_1 \in [2.6, 3.6]$), o módulo dos dentes ($m = x_2 \in [0.7, 0.8]$), o número de dentes ($n = x_3 \in [17, 18]$), o tamanho da haste 1 entre os suportes ($l_1 = x_4$), o tamanho da haste 2 entre os suportes ($l_2 = x_5$), o diâmetro da haste 1 ($d_1 = x_6$) e o diâmetro da haste 2 ($d_2 = x_7$). A variável x_3 é inteira e as demais são contínuas. As restrições incluem limitações da tensão de flexão e de superfície da engrenagem de dentes, deslocamento transversal das hastes 1 e 2 gerado pela força transmitida e as tensões nas hastes 1 e 2. O peso e as restrições mecânicas são

$$\begin{aligned} W(x) = & 0.7854x_1x_2^2 (3.3333x_3^2 + 14.9334x_3 - 43.0934) - \\ & 1.508x_1 (x_6^2 + x_7^2) + 7.477 (x_6^3 + x_7^3) + \\ & 0.7854 (x_4x_6^2 + x_5x_7^2) \end{aligned}$$

sujeito a

$$g_1(x) = 27x_1^{-1}x_2^{-2}x_3^{-1} \leq 1$$

$$g_2(x) = 397.5x_1^{-1}x_2^{-2}x_3^{-2} \leq 1$$

$$g_3(x) = 1.93x_2^{-1}x_3^{-1}x_4^3x_6^{-4} \leq 1$$

$$g_4(x) = 1.93x_2^{-1}x_3^{-1}x_5^3x_7^{-4} \leq 1$$

$$g_5(x) = \left[\left(\frac{745x_4}{x_2x_3} \right)^2 + 16.9x_6 \right]^{0.5} / 0.1x_6^3 \leq 1100$$

$$g_6(x) = \left[\left(\frac{745x_5}{x_2x_3} \right)^2 + 157.5x_7 \right]^{0.5} / 0.1x_7^3 \leq 850$$

$$g_7(x) = x_2x_3 \leq 40$$

$$g_8(x) = x_1/x_2 \geq 5$$

$$g_9(x) = x_1/x_2 \leq 12$$

$$g_{10}(x) = (1.5x_6 + 1.9)x_4^{-1} \leq 1$$

$$g_{11}(x) = (1.1x_7 + 1.9)x_5^{-1} \leq 1$$

$$2.6 \leq x_1 \leq 3.6$$

$$0.7 \leq x_2 \leq 0.8$$

$$17 \leq x_3 \leq 28$$

$$7.3 \leq x_4 \leq 8.3$$

$$7.8 \leq x_5 \leq 8.3$$

$$2.9 \leq x_6 \leq 3.9$$

$$5.0 \leq x_7 \leq 5.5$$

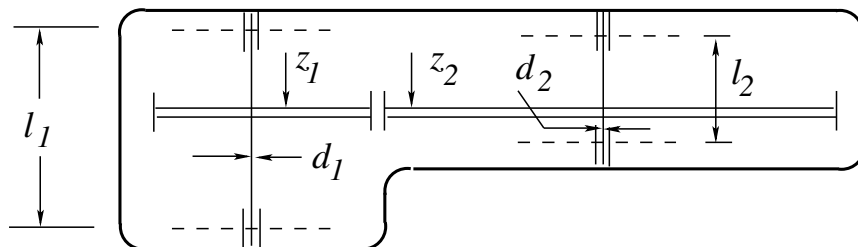


Figura 7.2: Redutor de velocidade.

Viga soldada

O objetivo é minimizar o custo $C(h, l, t, b)$ de uma viga soldada (fig. 7.3), onde $h \in [0.125, 10]$ e $0.1 \leq l, t, b \leq 10$. A função objetivo e as restrições são as seguintes

$$C(h, l, t, b) = 1.10471h^2l + 0.04811tb(14.0 + l)$$

sujeito a

$$g_1(\tau) = 13600 - \tau \geq 0$$

$$g_2(\sigma) = 30000 - \sigma \geq 0$$

$$g_3(b, h) = b - h \geq 0$$

$$g_4(P_c) = P_c - 6000 \geq 0$$

$$g_5(\delta) = 0.25 - \delta \geq 0$$

As expressões para τ , σ , P_c e δ são dadas por

$$\begin{aligned} \tau &= \sqrt{\tau'^2 + \tau''^2 + l\tau'\tau''/\alpha} \\ \alpha &= \sqrt{0.25(l^2 + (h+t)^2)}, \quad \sigma = \frac{504000}{t^2b} \\ P_c &= 64746.022(1 - 0.0282346t)tb^3 \\ \delta &= \frac{2.1952}{t^3b}, \quad \tau' = \frac{6000}{\sqrt{2hl}} \\ \tau'' &= \frac{6000(14 + 0.5l)\alpha}{2(0.707hl(l^2/12 + 0.25(h+t)^2))} \end{aligned}$$

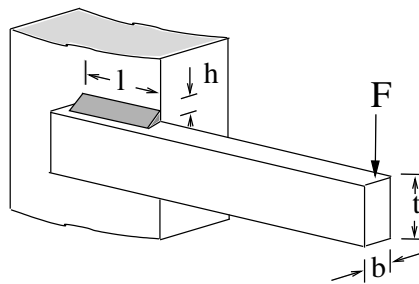


Figura 7.3: Viga soldada.

Vaso de pressão

Este problema corresponde à minimização de um vaso de pressão cilíndrico com duas tampas esféricas (fig. 7.4). São quatro variáveis de projeto: a espessura do vaso de pressão (T_s), a espessura da tampa (T_h), o raio interno do vaso (R) e a altura do componente

cilíndrico (L). Dentre essas, (T_s e T_h) são variáveis discretas e as outras duas (R e L) são contínuas. O peso a ser minimizado e as restrições são

$$w(T_s, T_h, R, L) = 0,6224T_sT_hRL + 1.7781T_hR^2 + 3.1661T_s^2L + 19.84T_s^2R$$

sujeito a

$$g_1(T_s, R) = T_s - 0.0193R \geq 0$$

$$g_2(T_h, R) = T_h - 0.00954R \geq 0$$

$$g_3(R, L) = \pi R^2L + 4/3\pi R^3 - 1296000 \geq 0$$

$$g_4(L) = -L + 240 \geq 0$$

onde

$$0.0625 \leq T_s, T_h \leq 5 \text{ e } 10 \leq R, L \leq 200$$

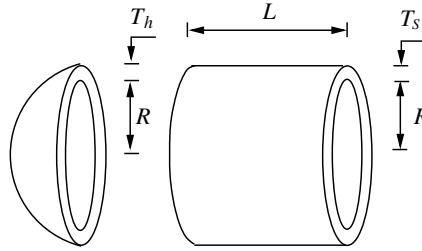


Figura 7.4: Vaso de pressão.

7.1 Comparação dos resultados dos experimentos numéricos

Para estes experimentos utilizou-se as mesmas 7 funções σ : $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$, $e^{\|\cdot\|_2} - 1$, $e^{\|\cdot\|_1} - 1$, $\sqrt{m} (e^{\|\cdot\|_\infty} - 1)$, $\max(\|\cdot\|_2, \|\cdot\|_2^2)$, e também a matriz $A = I$. Foram realizadas 30 execuções do IED+GA para cada problema, com critério de parada $\|g(x)\| \leq 10^{-6}$.

A tabela 7.1 mostra os valores iniciais dos parâmetros c_0 e u_0 do IED+GA, o tamanho da população e o número de gerações utilizados pelo GA.

Problemas	c_0	u_0	População	Gerações
Mola sob tração	10	[1 1 1 1]	100	300
Redutor de velocidade	1	[1 1 1 1 1 1 1 1 1 1 1]	100	300
Viga soldada	10	[1 1 1 1 1]	100	300
Vaso de pressão	1	[1 1 1 1]	100	300

Tabela 7.1: Parâmetros do IED+GA.

Na Tabela 7.2 são apresentados o melhor e o pior resultados, a média e o desvio padrão encontrados de cada um dos problemas.

As melhores soluções obtidas pelo IED+GA - variáveis, restrições e função objetivo - estão comparadas com outras soluções de referência nas tabelas 7.3 até 7.6.

Problemas de Engenharia	IED+GA			
	melhor	pior	média	desvio
Mola sob tração	0.0128	0.0164	0.0144	0.0009
Redutor de velocidade	2996.7	3039.7	3015.0	13.7485
Viga soldada	2.4151	5.4177	3.1995	0.7005
Vaso de pressão	6477.2	8787.4	7454.1	636.1933

Tabela 7.2: Resultados dos problemas de engenharia.

Var.	Alg-1 Ref. [40]	Alg-2 Ref. [40]	IED+GA
x_1	11.852177	11.329555	11.184056
x_2	0.34747463	0.35603234	0.36004951
x_3	0.051301897	0.051660806	0.051889349
g_1	-0.00000012	-0.000006437	-0.003088921
g_2	-0.00000047	-0.000013709	-0.003358194
g_3	-4.03513200	-4.052324300	-4.026621039
g_4	-0.73414900	-0.728204600	-0.725374090
V	0.012668	0.012666	0.012781

Tabela 7.3: Comparação das soluções do problema da mola sob tração.

Var.	Ref. [39]	Ref. [41]	Alg-1 Ref. [40]	Alg-2 Ref. [40]	IED+GA
x_1	3.506163	3.500000	3.500001	3.5	3.499684
x_2	0.700831	0.700000	0.700000	0.7	0.700011
x_3	17.0	17.0	17.0	17.0	17.290709
x_4	7.460181	7.300008	7.300017	7.3000035	7.302923
x_5	7.962143	7.715322	7.715326	7.7153225	7.806459
x_6	3.362900	3.350215	3.350216	3.3502147	3.350185
x_7	5.308949	5.286655	5.286654	5.2866545	5.286355
g_1	-0.077734	-0.07391524	-0.07391554	-0.07391524	-0.07386154
g_2	-0.201305	-0.19799852	-0.19799876	-0.19799852	-0.19795199
g_3	-0.474119	-0.49917084	-0.49916983	-0.49917156	-0.49856110
g_4	-0.897068	-0.90464383	-0.90464365	-0.90464383	-0.90120377
g_5	-0.011021	-0.00000023	-0.00000155	-0.000000119	0.000030538
g_6	-0.012500	-0.00000029	0.00000000	0.00000000	0.000187373
g_7	-0.702147	-0.70250000	-0.70250000	-0.702500000	-0.702495228
g_8	-0.000573	0.00000000	-0.00000029	0.00000000	0.00010613
g_9	-0.583095	-0.58333333	-0.58333320	-0.583333300	-0.583377553
g_{10}	-0.069144	-0.051326692	-0.05132753	-0.051326156	-0.051711423
g_{11}	-0.027920	-0.00000018	-0.00000077	-0.000000357	-0.011717075
V	3025.0051	2994.4717	2994.4720	2994.4712	2996.7085

Tabela 7.4: Comparação das soluções do problema redutor de velocidade.

Var.	Ref. [42]	Alg-1 Ref. [40]	Alg-2 Ref. [40]	IED+GA
x_1	0.2442949	0.24432427	0.24438575	0.240091515
x_2	6.2116738	6.2201996	6.2183037	6.390631032
x_3	8.3015486	8.291464	8.291165	8.314142911
x_4	0.2443003	0.24436942	0.24438748	0.246214266
g_1	0.0004447	0.000000000	0.001953125	0.004494538
g_2	64.378068	0.001953125	0.056640625	0.012900125
g_3	0.0000054	0.000045150	0.000001728	0.006122751
g_4	0.0002553	0.029785156	1.210937500	0.024765021
g_5	0.2342937	0.234240830	0.234240280	0.234486548
V	2.38159	2.381246	2.3812175	2.4151069

Tabela 7.5: Comparação das soluções do problema da viga soldada.

Var.	Ref. [41]	Ref. [42]	Alg-1 Ref. [40]	Alg-2 Ref. [40]	IED+GA
T_s	0.8125	0.8125	0.8125	0.8125	0.90625
T_h	0.4375	0.4375	0.4375	0.4375	0.46250
R	42.086994	42.0946558	42.093082	42.094967	46.189759
L	176.779128	176.684062	176.70308	176.67972	138.552311
g_1	0.000221	0.000073	0.0001035	0.000007	0.014787
g_2	0.035990	0.035917	0.0359320	0.035914	0.021849
g_3	3.219817	2.929000	0.1562500	0.0625	0.035064
g_4	63.220872	63.315938	63.296920	63.320282	101.447688
V	6061.1229	6060.187934	6060.3677	6060.138	6477.182

Tabela 7.6: Comparação das soluções do problema do vaso de pressão.

7.1.1 Análise dos resultados

- A análise da resolução dos problemas de engenharia mecânica mostra que o IED+GA pode ser aplicado em problemas reais de otimização com um certo grau de complexidade.
- A inclusão do algoritmo genético GA no método IED e a consequente criação da versão

híbrida IED+GA trouxe vantagens significativas no que tange as especificidades e características de cada um dos algoritmo, a saber:

- A resolução do subproblema (minimização da Lagrangiana), que nas versões anteriores IED+fmin e IED+NFDNA dependia fortemente dos valores iniciais dos parâmetros da função Lagrangiana, ganha agora liberdade de escolha com o GA.
- Por sua vez, o GA não precisa tratar diretamente as restrições do problema de otimização, haja vista que o IED já as incorpora na função Lagrangiana.
- A hibridização permitiu ao IED e ao GA um melhor ajuste nas suas limitações individuais, trazendo maior robustez ao método.

8 CONSIDERAÇÕES FINAIS

Apresentamos neste trabalho duas novas versões algorítmicas para o método de otimização IED, que se aplica à resolução de problemas não diferenciáveis, não convexos e com restrições. A base conceitual matemática - a teoria da dualidade Lagrangeana - presente na versão original do IED [5] foi preservada, não afetando sua prova de convergência.

O principal objetivo foi o de buscar novas estratégias para a solução do chamado **subproblema**, isto é, a cada iteração, realizar a minimização da função Lagrangeana. Nesse sentido, buscamos alternativas à função *fminsearch* (MatLab), responsável por esta tarefa no IED original.

Na primeira versão desenvolvida neste estudo, denominada IED+NFDNA, realizou-se inicialmente uma implementação do NFDNA [3], algoritmo voltado à otimização não diferenciável, não convexa e irrestrita, e que foi em seguida incorporado ao IED, em substituição ao *fminsearch*. Os experimentos numéricos realizados com esta versão, com a resolução de 15 problemas de otimização, foram divididos do seguinte modo:

1. Comparação entre o IED+NFDNA e a versão original IED+fmin, em relação aos números de iterações e avaliações da função Lagrangeana;
2. Comparação com as soluções de referência;
3. Comparação com outros solvers de otimização.

Os resultados destes testes numéricos comparativos, tabelados e apresentados no capítulo 5, demonstraram a eficiência do IED+NFDNA.

Com o intuito de continuar o aperfeiçoamento do método IED e de aplicá-lo a problemas mais complexos, foi desenvolvida e apresentada no capítulo 6 uma segunda versão, o IED+GA, com a entrada agora de algoritmos genéticos para a resolução do subproblema. Tal escolha baseou-se, num primeiro momento, na busca de uma programação que diminuísse a dependência do método no que se refere aos parâmetros iniciais c_0 e u_0 da função Lagrangeana. Essa hibridização mostrou-se robusta, associando características da parte determinística do IED com a probabilística dos genéticos, dentre as quais destacamos:

1. A incorporação das restrições à função Lagrangeana pelo IED que permitiu ao GA lidar no método somente com otimização sem restrições;
2. O GA trouxe contribuições importantes como o fato de não depender dos valores iniciais dos parâmetros u e c da Lagrangeana, e também a não necessidade do cálculos de subgradientes.

Os resultados dos experimentos numéricos obtidos pelo IED+GA mostram o preenchimento completo das tabelas dos 20 problemas resolvidos, bem como as melhores soluções compatíveis com as referências dos problemas, além de baixos valores dos desvios padrões.

O capítulo 7 apresenta a aplicação, pela primeira vez do método, na resolução de problemas de otimização que não são da literatura clássica matemática. O IED+GA encontrou resultados compatíveis com às soluções de outros algoritmos ao resolver problemas de engenharia.

8.1 CONTRIBUIÇÕES DESTE TRABALHO

A versão IED+NFDNA foi apresentada no EngOpt 2016 - 5th International Conference on Engineering Optimization, entre 19 a 23 de junho de 2016, com o trabalho completo publicado em Anais e disponível em (<http://engopt.org/proceedings.php?by=authorcod=615desc>).

O algoritmo IED+GA com os resultados numéricos dos problemas teste e dos problemas de engenharia resultou no artigo intitulado “*A hybrid Epigraph Directions Method for Nonsmooth and Nonconvex Constrained Optimization via Generalized Augmented Lagrangian Duality and a Genetic Algorithm*”, submetido ao journal *Mathematical Problems in Engineering*. O referido artigo foi aceito e será publicado no segundo semestre de 2018.

8.2 TRABALHOS FUTUROS

Ao longo do desenvolvimento desta pesquisa, diferentes observações realizadas nos permitem apontar para desdobramentos futuros, entre os quais, destacamos:

- Um estudo sobre a influência da matriz A que compõe o termo linear da função Lagrangeana. Até aqui, foram utilizadas apenas as matrizes nula e identidade.
- A possibilidade de se aplicar outras funções de penalização σ na formulação da Lagrangeana estendida.
- Outras estratégias que possam eventualmente substituir o algoritmo DSG para tratar os passos nulos.
- A aplicação da versão IED+GA em problemas de otimização de maior porte.

REFERÊNCIAS

- [1] ZÖRNIG, P., “Introdução à programação não linear”, *Brasília: UNB*, 2011.
- [2] IZMAILOV, A., SOLODOV, M., “Otimização: condições de otimalidade, elementos de análise convexa e de dualidade, Vol. 1”, *IMPA, Rio de Janeiro, Brasil*, 2005.
- [3] TANAKA FILHO, M., *ALGORITMOS DE DIREÇÕES VIÁVEIS PARA OTIMIZAÇÃO NÃO DIFERENCIÁVEL*, Tese (Doutorado), Universidade Federal do Rio de Janeiro, 2011.
- [4] KARMITSA, N., TANAKA FILHO, M., HERSKOVITS, J., “Globally convergent cutting plane method for nonconvex nonsmooth minimization”, *Journal of Optimization Theory and Applications*, v. 148, n. 3, pp. 528–549, 2011.
- [5] BURACHIK, R. S., FREIRE, W. P., KAYA, C. Y., “Interior Epigraph Directions method for nonsmooth and nonconvex optimization via generalized augmented Lagrangian duality”, *Journal of Global Optimization*, v. 60, n. 3, pp. 501–529, 2014.
- [6] LAGARIAS, J. C., REEDS, J. A., WRIGHT, M. H., WRIGHT, P. E., “Convergence properties of the Nelder–Mead simplex method in low dimensions”, *SIAM Journal on optimization*, v. 9, n. 1, pp. 112–147, 1998.
- [7] MATHEWS, J. H., FINK, K. D., OTHERS, *Numerical methods using MATLAB*. v. 4. Pearson Prentice Hall Upper Saddle River, NJ, 2004.
- [8] LINDEN, R., *Algoritmos genéticos (2a edição)*. Brasport, 2008.
- [9] HILLIER, F. S., LIEBERMAN, G. J., *Introdução à pesquisa operacional*. McGraw Hill Brasil, 2013.
- [10] RIBEIRO, A. A., KARAS, E. W., “Otimização Contínua: aspectos teóricos e computacionais”, *Cengage Learning, São Paulo, Brazil, in press (in Portuguese)*, 2013.
- [11] ROCKAFELLAR, R. T., *Convex analysis*. Princeton university press, 2015.

- [12] KELLEY, JR, J. E., “The cutting-plane method for solving convex programs”, *Journal of the society for Industrial and Applied Mathematics*, v. 8, n. 4, pp. 703–712, 1960.
- [13] LEMARECHAL, C., MIFFLIN, R., *Nonsmooth optimization: proceedings of a IIASA workshop, March 28-April 8, 1977*. Elsevier, 2014.
- [14] FUDULI, A., GAUDIOSO, M., GIALLOMBARDO, G., “Minimizing nonconvex nonsmooth functions via cutting planes and proximity control”, *SIAM Journal on Optimization*, v. 14, n. 3, pp. 743–756, 2004.
- [15] FLETCHER, R., “Practical methods of optimization john wiley & sons”, *New York*, v. 80, 1987.
- [16] FRANGIONI, A., “Generalized bundle methods”, *SIAM Journal on Optimization*, v. 13, n. 1, pp. 117–156, 2002.
- [17] KIWIEL, K. C., “A constraint linearization method for nondifferentiable convex minimization”, *Numerische Mathematik*, v. 51, n. 4, pp. 395–414, 1987.
- [18] IZMAILOV, A., SOLODOV, M., “Otimização Vol. 2: Condições de Otimalidade, Métodos Computacionais”, 2014.
- [19] LEMARÉCHAL, C., “Chapter vii nondifferentiable optimization”, *Handbooks in operations research and management science*, v. 1, pp. 529–572, 1989.
- [20] MÄKELÄ, M. M., NEITTAANMÄKI, P., *Nonsmooth optimization: analysis and algorithms with applications to optimal control*. World Scientific, 1992.
- [21] MIFFLIN, R., “Semismooth and semiconvex functions in constrained optimization”, *SIAM Journal on Control and Optimization*, v. 15, n. 6, pp. 959–972, 1977.
- [22] HERSKOVITS, J., “Feasible direction interior-point technique for nonlinear optimization”, *Journal of optimization theory and applications*, v. 99, n. 1, pp. 121–146, 1998.
- [23] FREIRE, W., *Um Algoritmo de Direções Viáveis para Otimização Convexa Não Diferenciável*, Tese (Doutorado), Tese (Doutorado), COPPE/UFRJ, Rio de Janeiro, 2005.

- [24] CLARKE, F. H., *Optimization and nonsmooth analysis*. SIAM, 1990.
- [25] SCHRAMM, H., ZOWE, J., “A version of the bundle idea for minimizing a nonsmooth function: Conceptual idea, convergence analysis, numerical results”, *SIAM journal on optimization*, v. 2, n. 1, pp. 121–152, 1992.
- [26] BURACHIK, R. S., KAYA, C. Y., “A deflected subgradient method using a general augmented Lagrangian duality with implications on penalty methods”, In: *Variational Analysis and Generalized Differentiation in Optimization and Control*, pp. 109–132, Springer, 2010.
- [27] BURACHIK, R. S., IUSEM, A. N., MELO, J. G., “The exact penalty map for nonsmooth and nonconvex optimization”, *Optimization*, v. 64, n. 4, pp. 717–738, 2015.
- [28] BURACHIK, R. S., IUSEM, A. N., MELO, J. G., “Strong duality and exact penalization for general augmented Lagrangians”, *Journal of optimization theory and applications*, v. 147, n. 1, pp. 125–140, 2010.
- [29] BURACHIK, R. S., IUSEM, A. N., MELO, J. G., “A primal dual modified subgradient algorithm with sharp Lagrangian”, *Journal of Global Optimization*, v. 46, n. 3, pp. 347–361, 2010.
- [30] GONÇALVES, A. M. S., *O Problema de Min-Max-Min com restrições pelo Método de Nelder-Mead*, Tese (Doutorado), Universidade Federal do Rio de Janeiro - COPPE, 2013.
- [31] GOUVÊA, E. J. C., *Métodos convergentes de otimização global baseados no vetor q -Gradiente*, Tese (Doutorado), Instituto Nacional de Pesquisas Espaciais - INPE, São José dos Campos, 2016.
- [32] NELDER, J. A., MEAD, R., “A simplex method for function minimization”, *The computer journal*, v. 7, n. 4, pp. 308–313, 1965.
- [33] HOCK, W., SCHITTKOWSKI, K., “Test examples for nonlinear programming codes”, *Journal of Optimization Theory and Applications*, v. 30, n. 1, pp. 127–129, 1980.

- [34] LEMONGE, A. C. C., “Aplicação de algoritmos genéticos em otimização estrutural”, *English: Genetic algorithms applications on structural optimization*). PhD Thesis, COPPE, Universidade Federal do Rio de Janeiro, 1999.
- [35] HOLLAND, J. H., *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT press, 1992.
- [36] AGRAWAL, R. B., DEB, K., AGRAWAL, R. B., “Simulated binary crossover for continuous search space”, *Complex systems*, v. 9, n. 2, pp. 115–148, 1995.
- [37] DEB, K., GOYAL, M., “A Combined Genetic Adaptive Search (GeneAS) for Engineering Design”, *Computer Science and Informatics*, v. 26, pp. 30–45, 1996.
- [38] DEB, K., AGRAWAL, R. B., “Simulated Binary Crossover for Continuous Search Space”, *Complex Systems*, pp. 115–148, 1995.
- [39] MEZURA-MONTES, E., COELLO, C. C., LANDA-BECERRA, R., “Engineering optimization using simple evolutionary algorithm”. In: *Tools with Artificial Intelligence, 2003. Proceedings. 15th IEEE International Conference on*, pp. 149–156, 2003.
- [40] BERNARDINO, H. S., BARBOSA, H. J., LEMONGE, A. C., “Constraint handling in genetic algorithms via artificial immune systems”. In: *Late-breaking paper at Genetic and Evolutionary Computation Conference (GECCO 2006)*, 2006.
- [41] COELLO, C. A. C., CORTÉS, N. C., “Hybridizing a genetic algorithm with an artificial immune system for global optimization”, *Engineering Optimization*, v. 36, n. 5, pp. 607–634, 2004.
- [42] LEMONGE, A. C., BARBOSA, H. J., “An adaptive penalty scheme for genetic algorithms in structural optimization”, *International Journal for Numerical Methods in Engineering*, v. 59, n. 5, pp. 703–736, 2004.