



ALGORITMOS NUMÉRICOS PARA PROGRAMAÇÃO SEMIDEFINIDA E  
APLICAÇÕES EM OTIMIZAÇÃO ESTRUTURAL

Elmer Giovanni Bazán Córdova

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia Mecânica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia Mecânica.

Orientadores: José Herskovits Norman  
Jean Rodolphe Roche

Rio de Janeiro  
Dezembro de 2016

ALGORITMOS NUMÉRICOS PARA PROGRAMAÇÃO SEMIDEFINIDA E  
APLICAÇÕES EM OTIMIZAÇÃO ESTRUTURAL

Elmer Giovanni Bazán Córdova

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ  
COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE)  
DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS  
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR  
EM CIÊNCIAS EM ENGENHARIA MECÂNICA.

Examinada por:

---

Prof. José Herskovits Norman, Dr.Ing.

---

Prof. Antonio André Novotny, D.Sc.

---

Prof. Thiago Gamboa Ritto, D.Sc.

---

Prof. Fernando Pereira Duda, D.Sc.

---

Prof. Susana Scheimberg de Makler, D.Sc.

RIO DE JANEIRO, RJ – BRASIL  
DEZEMBRO DE 2016

Córdova, Elmer Giovanni Bazán

Algoritmos Numéricos para Programação Semidefinida e Aplicações em Otimização Estrutural/Elmer Giovanni Bazán Córdova. – Rio de Janeiro: UFRJ/COPPE, 2016.

XV, 116 p.: il.; 29, 7cm.

Orientadores: José Herskovits Norman

Jean Rodolphe Roche

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia Mecânica, 2016.

Referências Bibliográficas: p. 97 – 102.

1. Programação semidefinida. 2. Programação não linear. 3. Otimização estrutural. I. Norman, José Herskovits *et al.* II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Mecânica. III. Título.

*A meus pais, Elmer e María,  
pelo carinho e constante apoio*

# Agradecimentos

Ao Prof. José Herskovits e ao Prof. Jean Roche, pela orientação, apoio e conhecimentos transmitidos ao longo da realização deste trabalho.

Ao Prof. Nelson Maculan, pelo apoio e motivação para fazer o Curso de Pós-Graduação na UFRJ.

Aos docentes e pessoal administrativo do Programa de Engenharia Mecânica PEM-COPPE/UFRJ.

A todos os colegas do Laboratório OptimizE.

Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado de Rio de Janeiro (FAPERJ) pelo suporte financeiro.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

## ALGORITMOS NUMÉRICOS PARA PROGRAMAÇÃO SEMIDEFINIDA E APLICAÇÕES EM OTIMIZAÇÃO ESTRUTURAL

Elmer Giovanni Bazán Córdova

Dezembro/2016

Orientadores: José Herskovits Norman  
Jean Rodolphe Roche

Programa: Engenharia Mecânica

As técnicas de programação semidefinida permitem lidar com problemas de otimização sujeitos a restrições matriciais semidefinidas e são muito eficientes em aplicações tais como otimização combinatória, teoria da aproximação, teoria de controle e otimização estrutural. Neste trabalho apresentam-se algoritmos numéricos para programação semidefinida não linear incluindo também restrições clássicas de igualdade e desigualdade. Estes algoritmos requerem um ponto inicial no interior de uma região viável definida pelas restrições e geram uma sequência de pontos no interior desta região que converge para a solução local do problema. Para isto, em cada iteração é definida uma direção de busca viável seguida de uma busca linear ao longo desta direção para definir o novo ponto da sequência para garantir convergência global. Finalmente, apresentam-se resultados numéricos de aplicações das técnicas desenvolvidas em problemas de otimização estrutural sujeitos a restrições locais de tensão, flambagem e deslocamento nodal e também restrições globais de frequência natural para mostrar o desempenho dos algoritmos.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

NUMERICAL ALGORITHMS FOR SEMIDEFINITE PROGRAMMING AND  
APPLICATIONS IN STRUCTURAL OPTIMIZATION

Elmer Giovanni Bazán Córdova

December/2016

Advisors: José Herskovits Norman  
Jean Rodolphe Roche

Department: Mechanical Engineering

The Semidefinite Programming Techniques allows us to deal with optimization problems subject to semidefinite matrix constraints and are very efficient in applications such as combinatorial optimization, approximation theory, control theory and structural optimization. In this work numerical algorithms for nonlinear semidefinite programming including also classical equality and inequality constraints are presented. These algorithms require an initial point at the interior of the feasible region defined by the constraints and generate a sequence of points within this region that converges to the local solution of the problem. For this, at each iteration a feasible direction search is defined followed by the linear search along this direction for to define a new point of the sequence and to ensure global converge. Finally, numerical examples of applications of developed techniques in structural optimization problems subject to local stress, buckling and nodal displacement constraints and also to global natural frequency constraints are presented to show the performance of the algorithms.

# Sumário

<b>Lista de Figuras</b>	<b>xi</b>
<b>Lista de Tabelas</b>	<b>xii</b>
<b>Lista de Símbolos</b>	<b>xiv</b>
<b>Introdução</b>	<b>1</b>
<b>1 Preliminares</b>	<b>6</b>
1.1 Algoritmo de pontos interiores e direções viáveis (FDIPA)	6
1.1.1 Notações	6
1.1.2 Definições	7
1.1.3 Condições de otimalidade de primeira ordem (KKT)	9
1.1.4 Algoritmo FDIPA	11
1.2 Algoritmo de pontos interiores e direções viáveis para programação semidefinida (FDIPA-SDP)	12
1.2.1 Algoritmo FDIPA-SDP	13
1.3 Algoritmo PENLAB	14
1.3.1 Observações	16
<b>2 Algoritmo de pontos interiores e direções viáveis para programação   semidefinida (FDIPA-SDP-NS)</b>	<b>18</b>
2.1 Introdução	18
2.2 Conceitos básicos	19
2.2.1 Notação e terminologia	19
2.2.2 Definições	21
2.2.3 Resultados auxiliares	22
2.3 Condições de otimalidade de primeira ordem	23
2.4 Hipóteses sobre o problema	24
2.5 Descrição do algoritmo	24
2.5.1 Matriz do sistema linear	25



2.5.2	Cálculo da direção de busca . . . . .	26
2.5.3	Direção viável . . . . .	28
2.5.4	Busca linear . . . . .	30
2.6	Algoritmo FDIPA-SDP-NS . . . . .	31
2.6.1	Hipóteses do algoritmo . . . . .	32
2.6.2	Atualizações . . . . .	32
2.7	Análise de convergência global . . . . .	33
2.8	Detalhes da implementação do algoritmo . . . . .	39
2.8.1	Resolução dos sistemas lineares internos do algoritmo FDIPA-SDP-NS . . . . .	39
2.8.2	Critérios de parada . . . . .	40
2.8.3	Busca de um ponto inicial viável . . . . .	41
<b>3</b>	<b>Algoritmo de pontos interiores e direções viáveis para programação semidefinida generalizada (FDIPA-GSDP)</b>	<b>42</b>
3.1	Introdução . . . . .	42
3.2	Problema de otimização com restrições de desigualdade . . . . .	43
3.2.1	Notações . . . . .	43
3.2.2	Definições . . . . .	44
3.2.3	Condições de otimalidade de primeira ordem . . . . .	44
3.2.4	Hipóteses sobre o problema . . . . .	45
3.2.5	Descrição do algoritmo FDIPA-GSDP . . . . .	46
3.2.6	Análise de convergência global . . . . .	51
3.3	Problema de otimização com restrições de igualdade . . . . .	55
3.3.1	Notações . . . . .	55
3.3.2	Condições de otimalidade de primeira ordem . . . . .	56
3.3.3	Hipóteses sobre o problema . . . . .	57
3.3.4	Descrição do algoritmo . . . . .	57
3.3.5	Análise de convergência global . . . . .	61
3.4	Algoritmo FDIPA-GSDP . . . . .	63
<b>4</b>	<b>Resultados Numéricos</b>	<b>65</b>
4.1	Introdução . . . . .	65
4.2	Dados de controle do algoritmo . . . . .	67
4.3	Problemas algébricos . . . . .	68
4.3.1	Exemplo 1 - Problema 43 aumentado . . . . .	68
4.3.2	Exemplo 2 - Problema 71 aumentado . . . . .	70
4.4	Aplicações em problemas de otimização estrutural dimensional . . . . .	72
4.4.1	Normalização das restrições . . . . .	73
4.4.2	Análise de sensibilidade . . . . .	74

4.4.3	Exemplo 3 - Treliça bidimensional de 10 barras . . . . .	75
4.4.4	Exemplo 4 - Treliça espacial de 25 barras . . . . .	78
4.4.5	Exemplo 5 - Treliça espacial de 72 barras . . . . .	81
4.5	Aplicações em problemas de otimização estrutural dimensional e geométrica . . . . .	85
4.5.1	Exemplo 6 - Treliça espacial de 25 barras . . . . .	85
4.6	Aplicações em problemas de otimização estrutural topológica . . . . .	88
4.6.1	Exemplo 7 - Treliça bidimensional de 26 barras . . . . .	89
4.7	Resumo dos resultados . . . . .	91
4.7.1	Esforço computacional . . . . .	92
4.7.2	Valores finais das condições de complementariedade . . . . .	92
4.7.3	Número de restrições ativas . . . . .	93
<b>5</b>	<b>Conclusões e sugestões</b>	<b>94</b>
5.1	Conclusões . . . . .	94
5.2	Sugestões para trabalhos futuros . . . . .	95
	<b>Referências Bibliográficas</b>	<b>97</b>
<b>A</b>	<b>Definições matemáticas</b>	<b>103</b>
A.1	Matrizes . . . . .	103
A.2	Problema de autovalores . . . . .	105
A.3	Transformação linear . . . . .	105
A.4	Complacência estrutural . . . . .	105
A.5	Complemento de Schur . . . . .	106
<b>B</b>	<b>Histórico das condições de complementariedade dos problemas teste resolvidos com o algoritmo FDIPA-GSDP</b>	<b>107</b>
B.1	Problema 43 aumentado . . . . .	108
B.2	Problema 71 aumentado . . . . .	108
B.3	Treliça de 10 barras . . . . .	109
B.4	Treliça de 25 barras . . . . .	110
B.5	Treliça de 72 barras . . . . .	111
B.6	Treliça de 25 barras . . . . .	112
B.7	Treliça de 26 barras . . . . .	113
<b>C</b>	<b>Análise estrutural de treliças com CALFEM</b>	<b>114</b>
C.1	Coordenadas locais e globais . . . . .	114
C.2	Vetores de deslocamentos e de carregamentos . . . . .	115
C.3	Matrizes de rigidez e de massa . . . . .	115
C.4	Tensão axial . . . . .	116

# Lista de Figuras

1.1	Direção de descida [1]. . . . .	8
1.2	Direção viável [1]. . . . .	9
1.3	Condições de KKT do problema 1.1. . . . .	10
3.1	Direção de busca não viável no algoritmo FDIPA-GSDP. . . . .	48
4.1	Evolução da redução da função objetivo do problema 43 aumentado. . . . .	69
4.2	Evolução da redução da função objetivo do problema 71 aumentado. . . . .	71
4.3	Geometria da treliça de 10 barras: $L = 360$ in (914.4 cm). . . . .	76
4.4	Configurações inicial e final da treliça de 10 barras. . . . .	77
4.5	Evolução da redução do peso da treliça de 10 barras. . . . .	78
4.6	Geometria da treliça de 25 barras: $a = 25$ in (63.5 cm). . . . .	78
4.7	Configurações inicial e final da treliça de 25 barras. . . . .	80
4.8	Evolução da redução do peso da treliça de 25 barras. . . . .	81
4.9	Geometria da treliça de 72 barras: $b = 60$ in (152.4 cm). . . . .	81
4.10	Configurações inicial e final da treliça de 72 barras. . . . .	83
4.11	Evolução da redução do peso da treliça de 72 barras. . . . .	84
4.12	Configurações inicial e final da treliça de 25 barras. . . . .	87
4.13	Evolução da redução do peso da treliça de 25 barras. . . . .	88
4.14	Geometria da treliça de 26 barras: $L = 360$ in (914.4 cm). . . . .	89
4.15	Configurações inicial e final da treliça de 26 barras. . . . .	90
4.16	Evolução da redução do peso da treliça de 26 barras. . . . .	91
C.1	Definição do elemento barra bidimensional [2, 3]. . . . .	114

# Lista de Tabelas

4.1	Notação empregada para definir a dimensão do problema . . . . .	66
4.2	Notação empregada para definir o esforço computacional . . . . .	66
4.3	Número de variáveis e de restrições do problema 43 aumentado. . . . .	68
4.4	Solução ótima do problema 43 aumentado. . . . .	69
4.5	Comparação dos resultados obtidos. . . . .	69
4.6	Número de variáveis e de restrições do problema 71 aumentado. . . . .	70
4.7	Solução do problema 71 aumentado. . . . .	70
4.8	Comparação dos resultados obtidos. . . . .	71
4.9	Casos de carregamento. Componentes da força em Kips (KN). . . . .	76
4.10	Definição dos elementos e conectividade dos nós. . . . .	76
4.11	Número de variáveis e de restrições da treliça de 10 barras. . . . .	77
4.12	Estrutura ótima obtida para a treliça de 10 barras. . . . .	77
4.13	Casos de carregamento. Componentes da força em Kips (KN). . . . .	79
4.14	Definição dos elementos e conectividade dos nós. . . . .	79
4.15	Número de variáveis e de restrições da treliça de 25 barras. . . . .	80
4.16	Estrutura ótima da treliça de 25 barras. . . . .	80
4.17	Casos de carregamento. Componentes da força em Kips (KN). . . . .	82
4.18	Definição dos elementos e conectividade dos nós. . . . .	82
4.19	Número de variáveis e de restrições da treliça de 72 barras. . . . .	83
4.20	Estrutura ótima da treliça de 72 barras. . . . .	84
4.21	Coordenadas nodais iniciais em in (cm). . . . .	86
4.22	Limites das variáveis geométricas da treliça de 25 barras em in (cm). . . . .	86
4.23	Número de variáveis e de restrições da treliça de 25 barras. . . . .	86
4.24	Estrutura ótima obtida para a treliça de 25 barras. . . . .	87
4.25	Casos de carregamento. Componentes da força em Kips (KN). . . . .	89
4.26	Número de variáveis e de restrições da treliça de 26 barras. . . . .	89
4.27	Definição dos elementos e conectividade dos nós. . . . .	90
4.28	Estrutura ótima obtida para a treliça de 26 barras. . . . .	91
4.29	Esforço computacional dos resultados numéricos obtidos com o algoritmo FDIPA-GSDP. . . . .	92

4.30	Valores finais de $\ \mathbf{d}_0\ $ , $\ \nabla L\ $ e das condições de complementariedade $\ \mathcal{G}\gamma_0\ $ , $ e_h $ e $ tr(\Lambda_0 \mathcal{A}) $ para o algoritmo FDIPA-GSDP. . . . .	92
4.31	Número de restrições ativas na solução dos problemas teste com o algoritmo FDIPA-GSDP. . . . .	93
B.1	Notação empregada nas tabelas dos resultados numéricos . . . . .	107
B.2	Evolução da redução da função objetivo para o problema 43 aumentado.108	
B.3	Evolução da redução da função objetivo para o problema 71 aumentado.108	
B.4	Evolução da redução do peso da treliça de 10 barras. . . . .	109
B.5	Evolução da redução do peso da treliça de 25 barras. . . . .	110
B.6	Evolução da redução do peso da treliça de 72 barras. . . . .	111
B.7	Evolução da redução do peso da treliça de 25 barras. . . . .	112
B.8	Evolução da redução do peso da treliça de 26 barras. . . . .	113

# Lista de Símbolos

$L_i$	Comprimento da barra $i$ , p. 71
$NF$	Número de avaliações da função objetivo, p. 65
$NI$	Número de iterações do algoritmo, p. 65
$\kappa^+$	Limite superior da complacência estrutural, p. 87
$\lambda^-$	Limite inferior da primeira frequência natural, p. 71
$\rho_i$	peso específico da barra $i$ , p. 71
$\sigma^+$	Tensão admissível de tração da barra $i$ , p. 71
$\sigma^-$	Tensão admissível de compressão da barra $i$ , p. 71
$\sigma_i$	Tensão da barra $i$ , p. 71
$\gamma$	Vetor dos multiplicadores de Lagrange das restrições de desigualdade, p. 9, 44
$\lambda_0$	Nova estimativa dos multiplicadores de Lagrange das restrições matriciais, p. 26
$\mathbf{K}$	Matriz de rigidez global reduzida, p. 71
$\mathbf{M}$	Matriz de massa global reduzida, p. 71
$b_i$	Área da seção transversal da barra $i$ , p. 71
$b_{\max}$	Limite superior da área da seção transversal $b_i$ , p. 71
$b_{\min}$	Limite inferior da área da seção transversal $b_i$ , p. 71
$iter$	Iteração $k$ , p. 106
$itlin$	Número de iterações na busca linear de Armijo, p. 106
$m$	Número de restrições de desigualdade, p. 65

$n$	Número de variáveis de projeto, p. 65
$nc$	Número de restrições de caixa, p. 65
$ne$	Número de elementos, p. 71
$ng$	Número de graus de liberdade reduzidos, p. 71
$nt$	Número total de restrições de igualdade e desigualdade, p. 65
$p$	Número de restrições de igualdade, p. 65
$q$	Número de filas ou colunas da restrição matricial, p. 65
$s$	Número de casos de carregamento, p. 72
$t$	Comprimento do passo $t^k$ da busca linear de Armijo, p. 106
$u^+$	Limite superior do deslocamento nodal $u_i$ , p. 71
$u^-$	Limite inferior do deslocamento nodal $u_i$ , p. 71
$u_i$	Deslocamento nodal ao longo do grau de liberdade $i$ , p. 71
$G$	Norma da condição de complementariedade das restrições de desigualdade, p. 91
$\ \nabla L\ $	Norma do gradiente da função lagrangiana, p. 91
$ tr(\Lambda_0 \mathcal{A}) $	Valor absoluto do traço da condição de complementariedade das restrições matriciais, p. 91
$\ \mathbf{d}_0\ $	Norma da direção de busca, p. 91
$ e_h $	Máximo erro das restrições de igualdade, p. 91

# Introdução

Um problema de programação matemática (ou simplesmente problema de otimização) pode ser definido de acordo com as características da função objetivo e do tipo de restrições impostas as quais podem ser funções vetoriais ou matriciais. Desta maneira, um problema é dito Problema de Programação Não Linear (PPNL) se pelo menos alguma das funções for não linear e Problema de Programação Semidefinida (PPSD) se as restrições são funções matriciais semidefinidas (positivas ou negativas).

Matematicamente, um problema clássico de programação não linear com restrições de desigualdade é formulado da seguinte maneira

$$\begin{cases} \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimizar}} & f(\mathbf{x}) \\ \text{sujeito a:} & \mathbf{g}(\mathbf{x}) \leq \mathbf{0} \end{cases} \quad (1)$$

onde  $\mathbf{x} \in \mathbb{R}^n$  é o vetor das variáveis de projeto,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é a função objetivo e  $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  é o vetor das  $m$  restrições de desigualdade do problema (1). Neste problema, as funções  $f$  e  $\mathbf{g}$  são consideradas suaves e continuamente diferenciáveis.

Existem muitas algoritmos na literatura que permitem resolver este tipo de problema tais como: método de direções viáveis, método do lagrangiano aumentado, método do gradiente conjugado, métodos de penalidade exata, algoritmos genéticos, entre outros.

Um algoritmo de pontos interiores muito eficiente e simples de implementar é o algoritmo de pontos interiores e direções viáveis (FDIPA), proposto por Herskovits [1, 4, 5]. Este algoritmo requer um ponto inicial pertencente ao interior de uma região definida pelas restrições de desigualdade, gerando uma sequência de pontos no interior desta região. O algoritmo FDIPA resolve dois sistemas de equações lineares com a mesma matriz seguida de uma busca linear, verificando as condições de otimalidade de Karush-Kuhn-Tucker em cada iteração e convergindo globalmente para pontos de Karush-Kuhn-Tucker. Este algoritmo permite resolver um problema generalizado de programação não linear quando são incluídas também as restrições de igualdade.



Por outro lado, um problema de programação semidefinida não linear [6–10] pode ser formulado da seguinte maneira

$$\begin{cases} \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimizar}} & f(\mathbf{x}) \\ \text{sujeito a:} & \mathcal{A}(\mathbf{x}) \preceq \mathbf{0} \end{cases} \quad (2)$$

onde  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é a função objetivo e  $\mathcal{A} : \mathbb{R}^n \rightarrow \mathcal{S}^q$  é a função de restrições matriciais semidefinidas negativas. A função  $\mathcal{A}$  também é considerada uma função suave e continuamente diferenciável.

A gama de aplicações das técnicas das técnicas de programação semidefinida é muito ampla entre elas temos: otimização combinatória, teoria da aproximação, teoria de controle e otimização estrutural. Muitos autores desenvolveram algoritmos para resolver problemas de programação semidefinida linear entre eles temos os trabalhos de Alizadeh [11], Stingl [12], Nesterov e Nemirovskii [13], Antonio e Lu [14], Helmberg et al. [15] e Babynin e Zhadan [16], enquanto que para problemas de programação semidefinida não linear temos os trabalhos de Correa e Ramírez [10], Yamashita et al. [17], Kočvara e Stingl [18], Jarre [19], Fares et al. [20], Huang et al. [21], Kanzow et al. [22] e Lourenço et al. [23].

O algoritmo FDIPA foi estendido para lidar também com este tipo de problema através do algoritmo de pontos interiores e direções viáveis para programação semidefinida não linear (FDIPA-SDP-NS) [24]. As técnicas desenvolvidas neste algoritmo estão baseadas no algoritmo FDIPA-SDP proposto por Aroztegui [8], reformulando os sistemas de equações lineares internos que tem que resolver o algoritmo com a finalidade de preservar a estrutura das restrições matriciais semidefinidas, tornado-se, desta maneira, computacionalmente mais eficiente. Esta reformulação implica a mudança do produto simétrico de Kronecker pelo produto não simétrico de Kronecker cujas definições e algumas propriedades podem ser achadas em [9].

A estrutura das restrições matriciais é muito importante em aplicações práticas da Engenharia como no caso de problemas de otimização estrutural. Em estes problemas, que serão abordados na presente pesquisa como problemas teste, empregam-se as matrizes de rigidez e de massa para realizar a análise estrutural através do método dos elementos finitos. Estas matrizes possuem uma estrutura particular de ser simétricas e semidefinidas positivas e também de ser esparsas.

As técnicas de programação semidefinida são muito eficientes em problemas de otimização estrutural que envolvem restrições globais de frequência natural e de flambagem [9, 25, 26]. Sem embargo podem ser ineficientes em problemas onde são consideradas restrições locais tensão (de tração e de compressão ou de flambagem) e deslocamento nodal [27–29] devido a que cada uma destas restrições tem que ser

inseridas na restrição matricial como uma matriz diagonal. Isto aumenta de modo considerável a dimensão desta matriz limitando desta maneira o número de variáveis de projeto e o número de restrições do problema.

## Contribuições da tese

Para evitar os inconvenientes descritos anteriormente, propõe-se neste trabalho um novo algoritmo de pontos interiores e direções viáveis que combina as técnicas desenvolvidas nos algoritmos FDIPA e FDIPA-SDP-NS para lidar com um problema generalizado de programação semidefinida não linear denominado FDIPA-GSDP sujeito a restrições de desigualdade e de igualdade e também a restrições matriciais semidefinidas. Este tipo de problema pode ser formulado da seguinte maneira

$$\left\{ \begin{array}{l} \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimizar}} \quad f(\mathbf{x}) \\ \text{sujeito a:} \quad \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{g}(\mathbf{x}) \in \mathbb{R}^m \\ \quad \quad \quad \mathbf{h}(\mathbf{x}) = \mathbf{0}, \mathbf{h}(\mathbf{x}) \in \mathbb{R}^p \\ \quad \quad \quad \mathcal{A}(\mathbf{x}) \preceq \mathbf{0}, \mathcal{A}(\mathbf{x}) \in \mathcal{S}^q \end{array} \right. \quad (3)$$

onde  $\mathbf{x} \in \mathbb{R}^n$  é o vetor das variáveis de projeto,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é a função objetivo,  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  é o vetor das  $m$  restrições de desigualdade,  $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$  é o vetor das  $p$  restrições de igualdade e  $\mathcal{A} : \mathbb{R}^n \rightarrow \mathcal{S}^q$  é um mapeamento de  $\mathbb{R}^n$  no espaço  $\mathcal{S}^q$  das matrizes reais simétricas.

Neste caso, as restrições locais de tensão e deslocamento nodal serão consideradas restrições vetoriais de desigualdade e as restrições globais de frequência natural (ou flambagem) serão tratadas como restrições matriciais semidefinidas isto permite resolver problemas de grande porte em comparação com o algoritmo FDIPA-SDP-NS convertendo-se em uma vantagem do método.

Finalmente, apresentam-se problemas algébricos e problemas de otimização estrutural de treliças planas e espaciais para avaliar o desempenho do algoritmo proposto.

## Estado da arte

A seguir, citam-se resumidamente os trabalhos mais relevantes disponíveis na literatura referentes ao estado da arte.

Herskovits [1, 4, 5] desenvolveu o algoritmo de pontos interiores e direções viáveis (FDIPA) para lidar com problemas de otimização não linear com restrições de desigualdade e de igualdade.

Aroztegui [8] utilizou uma técnica de pontos interiores e direções viáveis para programação semidefinida não linear e desenvolveu o algoritmo FDIPA-SDP para resolver problemas de otimização estrutural de materiais lineares elásticos e otimização de estruturas planas onde minimiza-se o volume da estrutura com restrições definidas sobre o tensor de elasticidade.

Köcvara e Stingl [18, 30] desenvolveram o algoritmo PENNON para resolver problemas de programação não linear e semidefinida empregando o método generalizado do lagrangiano aumentado ( $\text{NLP-SDP}_\phi$ ) e apresentaram exemplos algébricos e aplicações em problemas de otimização estrutural topológica. Uma versão livre deste algoritmo é o PENLAB [31].

Yamashita et al. [17] propuseram um método de pontos interiores para resolver problemas de programação semidefinida não linear com restrições de desigualdade e restrições semidefinidas.

Achtzinger e Köcvara [25] apresentaram formulações tradicionais de diferentes problemas de otimização topológica de treliças e suas equivalências mediante a programação semidefinida tais como: a minimização do peso da estrutura para diferentes casos de carregamento, a maximização do mínimo autovalor e a minimização da complacência (maximização da rigidez) estrutural.

Ohsaki *et al.* [32] demonstraram que um problema de otimização estrutural de treliças com multiplicidade de frequências naturais pode ser resolvido empregando as técnicas de programação semidefinida e estabeleceram que este método pode ser estendido a problemas com restrições de flambagem.

## Objetivos

Este trabalho tem os seguintes objetivos:

- Realizar a análise de convergência do algoritmo FDIPA-SDP-NS;
- Formular e implementar computacionalmente o novo algoritmo de pontos interiores e direções viáveis para um problema generalizado de programação não linear semidefinida (FDIPA-GSDP);
- Realizar a análise de convergência do algoritmo FDIPA-GSDP;
- Formular problemas de otimização estrutural para minimizar o peso de estruturas treliçadas sujeitas a restrições estáticas locais de tensão, deslocamento e flambagem e também a restrições dinâmicas globais de frequência natural;
- Mostrar resultados numéricos para problemas algébricos e de otimização estrutural;

## Organização do trabalho

O presente trabalho está organizado em seis capítulos, distribuídos da seguinte maneira:

O capítulo 1 é reservado aos preliminares onde são descritos resumidamente os algoritmos FDIPA e FDIPA-SDP os quais são a base do algoritmo FDIPA-GSDP. Além disso, é descrito o algoritmo PENLAB que será utilizado para comparar o desempenho do algoritmo FDIPA-GSDP;

No Capítulo 2, apresenta-se o algoritmo FDIPA-SDP-NS acrescentando a análise de convergência do mesmo.

No Capítulo 3, apresenta-se o novo algoritmo de pontos interiores e direções viáveis para programação semidefinida generalizada não linear (FDIPA-GSDP) seguido da análise de convergência do algoritmo.

No Capítulo 4, apresentam-se testes numéricos em problemas algébricos e problemas otimização estrutural de treliças para avaliar o comportamento do algoritmo FDIPA-GSDP.

Finalmente, no Capítulo 5, mostram-se as conclusões do presente trabalho e as sugestões para os trabalhos futuros.

# Capítulo 1

## Preliminares

Neste capítulo, apresentam-se resumidamente os algoritmos FDIPA e FDIPA-SDP que foram a base do algoritmo FDIPA-GSDP proposto no presente trabalho. Além disso apresenta-se o algoritmo PENLAB [31] que será utilizado como algoritmo de comparação dos resultados numéricos.

### 1.1 Algoritmo de pontos interiores e direções viáveis (FDIPA)

Nesta seção, apresenta-se a versão básica do algoritmo FDIPA desenvolvido por Herskovits [5] para lidar com problemas de otimização não linear com restrições de desigualdade que é formulado como

$$\begin{cases} \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimizar}} & f(\mathbf{x}) \\ \text{sujeito a:} & \mathbf{g}(\mathbf{x}) \leq 0 \end{cases} \quad (1.1)$$

onde  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é chamada a função objetivo e  $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  é a função das restrições de desigualdade.

#### 1.1.1 Notações

Denote-se por  $n$  o número de variáveis de projeto e  $m$  o número de restrições de desigualdade.

O vetor das variáveis de projeto é denotado por  $\mathbf{x}$  e definido como  $\mathbf{x} = [x_1, \dots, x_n]^T$ , entretanto, o vetor de restrições é denotado por  $\mathbf{g}$  e definido como  $\mathbf{g}(\mathbf{x}) = [g_1(\mathbf{x}), \dots, g_m(\mathbf{x})]^T$ .

O vetor gradiente de uma função  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é denotado por  $\nabla f(\mathbf{x})$  e definido como

$$\nabla f(\mathbf{x}) = \left[ \frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_n} \right]^T \in \mathbb{R}^n \quad (1.2)$$

As derivadas parciais de  $g_i$  em relação a  $x_i$  são agrupadas na matriz  $\nabla \mathbf{g}(\mathbf{x}) \in \mathbb{R}^{n \times m}$  denotada por

$$\nabla \mathbf{g}(\mathbf{x}) = [\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})] \quad (1.3)$$

A Hessiana de  $f$  em  $\mathbf{x}$  é denotada por

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{bmatrix} \quad (1.4)$$

onde as componentes  $\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}$ ,  $i, j = 1, \dots, n$  são as derivadas parciais de segunda ordem da função  $f$ .

A função Lagrangiana  $L : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$  associada ao problema 1.1 é definida como

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x}) = f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) \quad (1.5)$$

O gradiente da função Lagrangiana em relação a  $\mathbf{x}$  é denotado por  $\nabla_{\mathbf{x}} L : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^{n \times m}$  e definido como

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}) = \nabla f(\mathbf{x}) + \sum_{i=1}^m \lambda_i \nabla g_i(\mathbf{x}) \quad (1.6)$$

Finalmente, a Hessiana da função Lagrangiana associada ao problema 1.1 é denotada por  $H : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^{n \times m}$  e definida como

$$H(\mathbf{x}, \boldsymbol{\lambda}) = \nabla^2 f(\mathbf{x}) + \sum_{i=1}^m \lambda_i \nabla^2 g_i(\mathbf{x}) \quad (1.7)$$

## 1.1.2 Definições

A seguir, mostram-se algumas definições que serão utilizadas ao longo do presente trabalho.

**Definição 1.1 (Região viável).** Uma região viável  $\Omega$  é o conjunto de pontos viáveis do problema (1.1) que satisfazem simultaneamente todas as restrições e é definida como

$$\Omega := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) \leq 0\} \quad (1.8)$$

**Definição 1.2 (Conjunto de restrições ativas).** Dado um ponto  $\mathbf{x} \in \Omega$ , o conjunto de restrições de desigualdade ativas está dado por

$$I(\mathbf{x}) := \{g_i : g_i(\mathbf{x}) = 0, \forall \mathbf{x} \in \Omega, i = 1, \dots, m\} \quad (1.9)$$

**Definição 1.3 (Direção de descida).** Um vetor  $\mathbf{d} \in \mathbb{R}^n$  é uma direção de descida de uma função suave  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$  em um ponto  $\mathbf{x} \in \mathbb{R}^n$  se existe  $\psi > 0$  tal que  $\phi(\mathbf{x} + t\mathbf{d}) < \phi(\mathbf{x})$  para todo  $t \in (0, \psi)$ . Na Figura

Na Figura 1.1 pode-se observar uma interpretação geométrica da definição de direção de descida. Se  $\mathbf{d}^T \nabla \phi(\mathbf{x}) < 0$ , então  $\mathbf{d}$  é uma direção de descida de  $\phi$  em um ponto  $\mathbf{x}$ .

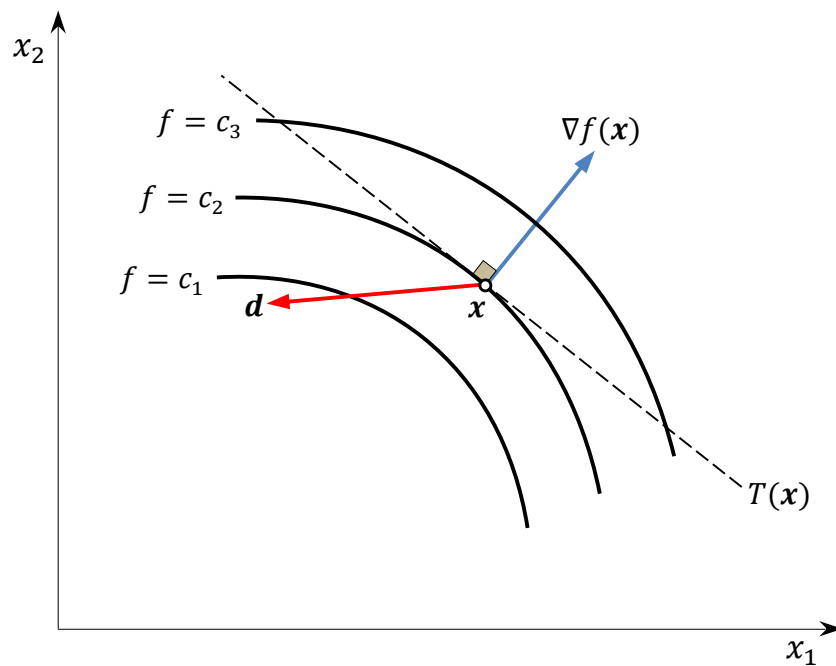


Figura 1.1: Direção de descida [1].

**Definição 1.4 (Direção viável).** Um vetor  $\mathbf{d} \in \mathbb{R}^n$  é uma direção viável na região  $\Omega$  em um ponto  $\mathbf{x} \in \Omega$  se existe  $\theta > 0$  tal que  $\mathbf{x} + t\mathbf{d} \in \Omega$  para todo  $t \in [0, \theta)$  (Ver Figura 1.2).

**Definição 1.5 (Campo de direções uniformemente viáveis).** Um vetor campo  $\mathbf{d}(\mathbf{x})$  definido em  $\Omega \subseteq \mathbb{R}^n$  é chamado um campo de direções uniformemente viáveis do problema (2.1), se existe um tamanho de passo  $\tau > 0$ , tal que  $\mathbf{x} + t\mathbf{d}(\mathbf{x}) \in \Omega$  para todo  $t \in [0, \tau]$  e para todo  $\mathbf{x} \in \Omega$ .

**Definição 1.6 (Mínimo local).** O ponto  $\mathbf{x}^* \in \Omega$  é um mínimo local de  $f$ , se existe um  $\epsilon > 0$  tal que  $f(\mathbf{x}^*) \leq f(\mathbf{x})$  para todo  $\mathbf{x} \in \Omega \cap B(\mathbf{x}^*, \epsilon)$ . Onde  $B(\mathbf{x}^*, \epsilon)$  é a bola aberta de centro  $\mathbf{x}^*$  e raio  $\epsilon$ .

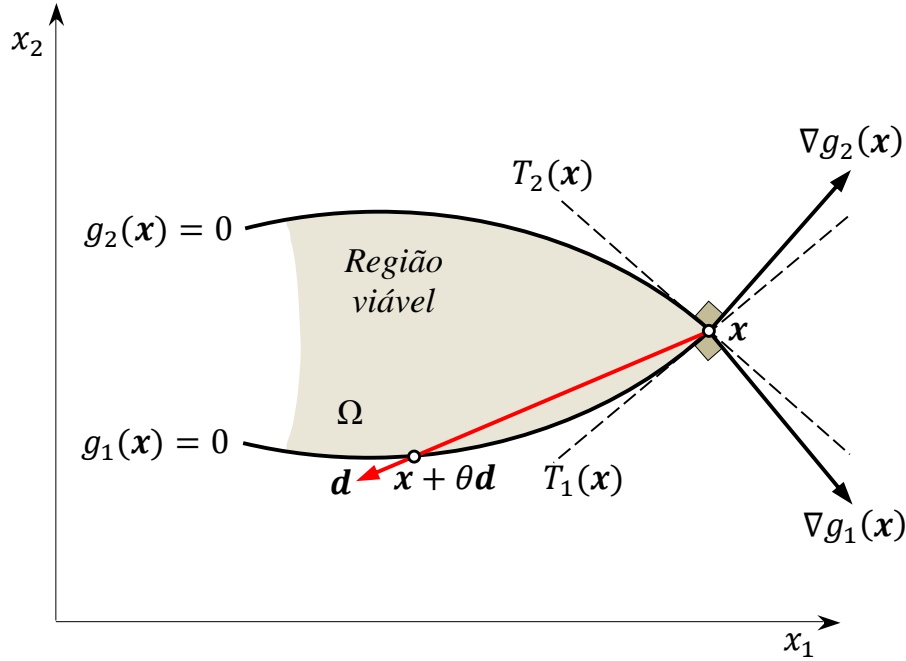


Figura 1.2: Direção viável [1].

**Definição 1.7 (Mínimo local estrito).** O ponto  $\mathbf{x}^* \in \Omega$  é um mínimo local estrito de  $f$ , se existe um  $\epsilon > 0$  tal que  $f(\mathbf{x}^*) < f(\mathbf{x})$  para todo  $\mathbf{x} \in \Omega \cap B(\mathbf{x}^*, \epsilon)$ .

### 1.1.3 Condições de otimalidade de primeira ordem (KKT)

As condições de otimalidade de primeira ordem de Karush-Kuhn-Tucker, correspondentes ao problema (1.1), podem ser escritas da seguinte maneira

$$\nabla f(\mathbf{x}^*) + \nabla \mathbf{g}(\mathbf{x}^*) \boldsymbol{\lambda}^* = 0 \quad (1.10)$$

$$\mathcal{G}(\mathbf{x}^*) \boldsymbol{\gamma}^* = 0 \quad (1.11)$$

$$\boldsymbol{\gamma}^* \geq 0 \quad (1.12)$$

$$\mathbf{g}(\mathbf{x}^*) \leq 0 \quad (1.13)$$

sendo  $\boldsymbol{\gamma} \in \mathbb{R}_+^m$  o vetor dos multiplicadores de Lagrange das restrições de desigualdade e  $\mathcal{G}(\mathbf{x}) \in \mathbb{R}^{m \times m}$  uma matriz diagonal definida como  $\mathcal{G}(\mathbf{x}) = \text{diag}[g_1(\mathbf{x}), \dots, g_m(\mathbf{x})]$ .

A Figura 1.3 [33, 34] mostra uma interpretação geométrica das condições de primeira ordem de KKT. Note-se que o ponto  $\mathbf{x}^*$  é um mínimo da função  $f$ . Tem-se três restrições de desigualdade com duas variáveis de projeto. A restrição  $g_3(\mathbf{x}) \leq 0$  é inativa, isto é  $g_3(\mathbf{x}^*) < 0$ , e portanto,  $\lambda_3 = 0$ . Aplicando a primeira condição de otimalidade de KKT, temos

$$\nabla f(\mathbf{x}^*) + \nabla g_1(\mathbf{x}^*) \lambda_1 + \nabla g_2(\mathbf{x}^*) \lambda_2 = 0, \quad (1.14)$$



que é equivalente a

$$\nabla f(\mathbf{x}^*) = -\nabla g_1(\mathbf{x}^*)\lambda_1 - \nabla g_2(\mathbf{x}^*)\lambda_2, \quad (1.15)$$

onde  $\lambda_1 > 0$  e  $\lambda_2 > 0$ . Pode-se observar que  $\nabla f(\mathbf{x}^*)$  deve ser uma combinação linear dos vetores  $-\nabla g_1(\mathbf{x}^*)$  e  $-\nabla g_2(\mathbf{x}^*)$  com coeficientes positivos. Neste caso, os coeficientes representam os multiplicadores KKT.

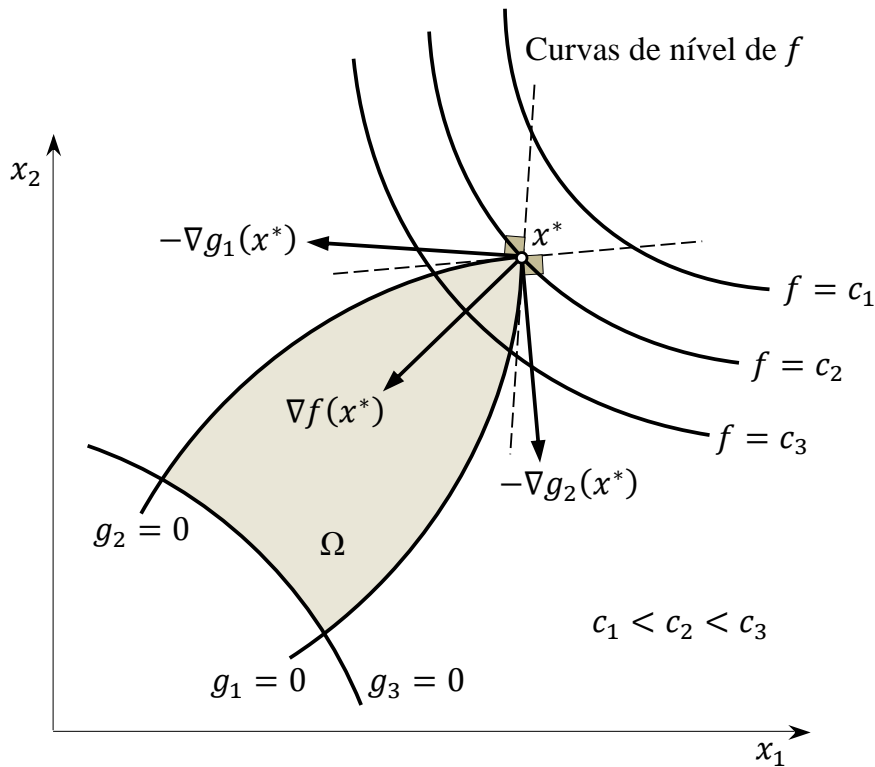


Figura 1.3: Condições de KKT do problema 1.1.

**Definição 1.8 (Ponto regular).** Um ponto  $\mathbf{x}$  é um ponto regular do problema (1.1) se o conjunto  $\{\nabla g_i(\mathbf{x}) : i \in I(\mathbf{x})\}$  é linearmente independente.

**Definição 1.9 (Espaço tangente).** O espaço tangente no ponto regular  $\mathbf{x}$  é dado por

$$T(\mathbf{x}) := \{\mathbf{d} : \nabla g_i^T \mathbf{d} = 0, \forall i \in I(\mathbf{x})\} \quad (1.16)$$

**Definição 1.10 (Par de KKT).** Um vetor  $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$  satisfazendo as Equações (1.10) e (1.11) é chamado um par estacionário do problema 1.1 e um par KKT, se este satisfizer as condições de KKT (1.10)-(1.13).

**Definição 1.11 (Ponto de KKT).** Um vetor  $\mathbf{x}^*$  é um ponto estacionário se existe  $\boldsymbol{\lambda}^*$  tal que as igualdades (1.10) e (1.11) são verdadeiras, e, será um ponto de KKT, se todas as Equações (1.10)-(1.13) forem confirmadas.

### 1.1.4 Algoritmo FDIPA

O algoritmo FDIPA será descrito abaixo de forma resumida, a fim de se entender o seu funcionamento. Na verdade, obtém-se uma família de algoritmos, estabelecendo-se diferentes regras para atualização da matriz quase Newton  $\mathbf{B}$  e do vetor  $\boldsymbol{\lambda}$ .

Isto permite a implementação de diferentes versões, dependendo do problema a ser resolvido, sobre a informação disponível das funções  $f$  e  $g$ , e, sobre a taxa de convergência desejada.

O algoritmo FDIPA consiste no seguinte:

**Parâmetros:**  $\xi \in (0, 1)$ ,  $\eta \in (0, 1)$ ,  $\varphi > 0$  e  $\nu \in (0, 1)$

**Dados Iniciais:**  $\mathbf{x} \in \Omega_a^0$ ,  $0 < \boldsymbol{\lambda} \in \mathbb{R}^m$  e  $\mathbf{B} \in \mathbb{R}^{n \times n}$  simétrica e definida positiva.

**Passo 1: Cálculo da direção de busca**

**Passo 1a:** Resolva o sistema linear para obter  $\mathbf{d}_0$  e  $\boldsymbol{\lambda}_0$

$$\begin{bmatrix} \mathbf{B} & \nabla \mathbf{g}(\mathbf{x}) \\ \boldsymbol{\Lambda} \nabla \mathbf{g}(\mathbf{x})^T & \mathcal{G}(\mathbf{x}) \end{bmatrix} \begin{bmatrix} \mathbf{d}_0 \\ \boldsymbol{\gamma}_0 \end{bmatrix} = \begin{bmatrix} -\nabla f(\mathbf{x}) \\ \mathbf{0} \end{bmatrix} \quad (1.17)$$

Se  $\mathbf{d}_0 = 0$ , parar.

**Passo 1b:** Resolva o sistema linear auxiliar para obter  $\mathbf{d}_1$  e  $\boldsymbol{\lambda}_1$

$$\begin{bmatrix} \mathbf{B} & \nabla \mathbf{g}(\mathbf{x}) \\ \boldsymbol{\Lambda} \nabla \mathbf{g}(\mathbf{x})^T & \mathcal{G}(\mathbf{x}) \end{bmatrix} \begin{bmatrix} \mathbf{d}_1 \\ \boldsymbol{\gamma}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ -\boldsymbol{\gamma} \end{bmatrix} \quad (1.18)$$

**Passo 1c:** Se  $\mathbf{d}_1^T \nabla f(\mathbf{x}) > 0$ , calcule

$$\rho = \min \left\{ \varphi \|\mathbf{d}_0\|^2, (\xi - 1) \frac{\mathbf{d}_0^T \nabla f(\mathbf{x})}{\mathbf{d}_1^T \nabla f(\mathbf{x})} \right\} \quad (1.19)$$

caso contrário, faça:

$$\rho = \varphi \|\mathbf{d}_0\|_2^2 \quad (1.20)$$

**Passo 1d:** Calcule a direção de busca  $\mathbf{d}$  e o vetor  $\bar{\boldsymbol{\gamma}}$

$$\mathbf{d} = \mathbf{d}_0 + \rho \mathbf{d}_1 \quad (1.21)$$

e

$$\bar{\boldsymbol{\gamma}} = \boldsymbol{\gamma}_0 + \rho \boldsymbol{\gamma}_1 \quad (1.22)$$

**Passo 2: Busca Linear**

Calcule o comprimento do passo  $t$ , o primeiro número da série  $\{1, \nu, \nu^2, \nu^3, \dots\}$ ,

tal que

$$f(\mathbf{x} + t\mathbf{d}) \leq f(\mathbf{x}) + t\eta\nabla f(\mathbf{x})^T \mathbf{d} \quad (1.23)$$

e

$$g_i(\mathbf{x} + t\mathbf{d}) < 0, \text{ se } \bar{\lambda}_i \geq 0, \quad i = 1, \dots, m \quad (1.24)$$

ou

$$g_i(\mathbf{x} + t\mathbf{d}) \leq g_i(\mathbf{x}), \text{ se } \bar{\lambda}_i < 0, \quad i = 1, \dots, m \quad (1.25)$$

### Passo 3: Atualizações

**Passo 3a:** Obtenha o novo ponto, como:  $\mathbf{x} := \mathbf{x} + t\mathbf{d}$ ;

**Passo 3b:** Defina novos valores para  $\mathbf{B}$  simétrica definida positiva;

**Passo 3c:** Defina novos valores para  $\boldsymbol{\lambda} > 0$ . A seguinte atualização para  $\boldsymbol{\lambda}$  já foi considerada por Herskovits [5],

$$\lambda_i := \max \left[ \lambda_{0i}, \epsilon \|\mathbf{d}_0\|^2 \right], \quad \epsilon > 0, \quad \text{para } i = 1, \dots, m$$

**Passo 3d:** Retorne ao Passo 1.

## 1.2 Algoritmo de pontos interiores e direções viáveis para programação semidefinida (FDIPA-SDP)

Nesta seção, apresenta-se o algoritmo de pontos interiores e direções viáveis para programação semidefinida não linear (FDIPA-SDP) desenvolvido por Aroztegui [8]. Um problema de programação semidefinida é formulado como

$$\begin{cases} \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimizar}} & f(\mathbf{x}) \\ \text{sujeito a:} & \mathcal{A}(\mathbf{x}) \preceq \mathbf{0} \end{cases} \quad (1.26)$$

onde,  $\mathbf{x} \in \mathbb{R}^n$  é o vetor das variáveis de projeto,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é a função objetivo e  $\mathcal{A} : \mathbb{R}^n \rightarrow \mathcal{S}^q$  é um mapeamento de  $\mathbb{R}^n$  no espaço  $\mathcal{S}^q$  das matrizes simétricas de ordem  $q \times q$ . As funções  $f$  e  $\mathcal{A}$  são consideradas suaves e pelo menos uma delas é não linear.

Cabe indicar que neste algoritmo é empregado o símbolo  $\otimes_s$  que representa o produto simétrico de Kronecker [9] de duas matrizes simétricas  $\mathbf{A} \in \mathbb{R}^{q \times q}$  e  $\mathbf{B} \in \mathbb{R}^{q \times q}$  definido como

$$(\mathbf{A} \otimes_s \mathbf{B}) \mathbf{u} = \frac{1}{2} \text{svec}(\mathbf{B}\mathbf{U}\mathbf{A}^T + \mathbf{A}\mathbf{U}\mathbf{B}^T) = \frac{1}{2} \text{svec}(\mathbf{B}\mathbf{U}\mathbf{A} + \mathbf{A}\mathbf{U}\mathbf{B}) \quad (1.27)$$

sendo  $\mathbf{U} \in \mathbb{R}^{q \times q}$  uma matriz simétrica e  $\mathbf{u} = \text{svec}(\mathbf{U})$ . Aqui,  $\text{svec} : \mathbb{R}^{q \times q} \rightarrow \mathbb{R}^{q(q+1)/2}$  representa o mapeamento definido por

$$\text{svec}(\mathbf{U}) = [u_{11}, \sqrt{2}u_{21}, \dots, u_{22}, \sqrt{2}u_{32}, \dots, \sqrt{2}u_{q2}, u_{qq}]^T \quad (1.28)$$

### 1.2.1 Algoritmo FDIPA-SDP

O algoritmo FDIPA-SDP consiste no seguinte:

**Parâmetros de controle:**  $\xi \in (0, 1)$ ,  $\eta \in (0, 1)$ ,  $\varphi > 0$ , e  $\nu \in (0, 1)$ .

**Dados Iniciais:**  $\mathbf{x} \in \text{int}(\Omega)$ ,  $\mathbf{B} \in \mathcal{S}_{++}^n$  e  $\mathbf{\Lambda} \in \mathcal{S}_{++}^q$  comutando com  $\mathcal{A}(\mathbf{x})$ .

**Passo 1:** Cálculo da direção de busca

**Passo 1a:** Resolva o sistema linear para obter  $\mathbf{d}_0 \in \mathbb{R}^n$  e  $\boldsymbol{\lambda}_0 \in \mathbb{R}^m$ .

$$\begin{bmatrix} \mathbf{B} & \nabla \mathcal{A}(\mathbf{x}) \\ [\mathbf{\Lambda} \otimes_s \mathbf{I}_q] \nabla \mathcal{A}(\mathbf{x})^T & \mathbf{I}_q \otimes_s \mathcal{A}(\mathbf{x}) \end{bmatrix} \begin{bmatrix} \mathbf{d}_0 \\ \boldsymbol{\lambda}_0 \end{bmatrix} = \begin{bmatrix} -\nabla f(\mathbf{x}) \\ \mathbf{0} \end{bmatrix} \quad (1.29)$$

Se  $\mathbf{d}_0 = \mathbf{0}$ , parar.

**Passo 1b:** Resolva o sistema linear para obter  $\mathbf{d}_1 \in \mathbb{R}^n$  e  $\boldsymbol{\lambda}_1 \in \mathbb{R}^m$

$$\begin{bmatrix} \mathbf{B} & \nabla \mathcal{A}(\mathbf{x}) \\ [\mathbf{\Lambda} \otimes_s \mathbf{I}_q] \nabla \mathcal{A}(\mathbf{x})^T & \mathbf{I}_q \otimes_s \mathcal{A}(\mathbf{x}) \end{bmatrix} \begin{bmatrix} \mathbf{d}_1 \\ \boldsymbol{\lambda}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ -\boldsymbol{\lambda} \end{bmatrix} \quad (1.30)$$

**Passo 1c:** Se  $\mathbf{d}_1^T \nabla f(\mathbf{x}) > 0$ , calcule

$$\rho = \min \left\{ \varphi \|\mathbf{d}_0\|^2, (\xi - 1) \frac{\mathbf{d}_0^T \nabla f(\mathbf{x})}{\mathbf{d}_1^T \nabla f(\mathbf{x})} \right\} \quad (1.31)$$

caso contrário, calcule

$$\rho = \varphi \|\mathbf{d}_0\|^2 \quad (1.32)$$

**Passo 1d:** Calcule a direção de busca  $\mathbf{d}$  e  $\bar{\mathbf{\Lambda}}$

$$\mathbf{d} = \mathbf{d}_0 + \rho \mathbf{d}_1 \quad (1.33)$$

e

$$\bar{\mathbf{\Lambda}} = \mathbf{\Lambda}_0 + \rho \mathbf{\Lambda}_1 \quad (1.34)$$

**Passo 2:** Busca Linear

Calcule o comprimento do passo  $t$ , o primeiro número da série  $\{1, \nu, \nu^2, \nu^3, \dots\}$ ,

tal que

$$f(\mathbf{x} + t\mathbf{d}) \leq f(\mathbf{x}) + t\eta\nabla f(\mathbf{x})^T \mathbf{d} \quad (1.35)$$

e

$$\mathcal{A}(\mathbf{x} + t\mathbf{d}) \prec 0 \quad (1.36)$$

### Passo 3 Atualizações

**Passo 3a:** Obtenha o novo ponto, como:  $\mathbf{x} := \mathbf{x} + t\mathbf{d}$ ;

**Passo 3b:** Defina novos valores para  $\mathbf{B} \in \mathcal{S}_{++}^m$ ;

**Passo 3c:** Defina novos valores para  $\mathbf{\Lambda} \in \mathcal{S}_{++}^m$  comutando com  $\mathcal{A}(\mathbf{x})$ . A seguinte regra de atualização para  $\mathbf{\Lambda}$  foi considerada por Aroztegui [8, 35],

(i) Calcule  $\alpha_{0min}$ , o minimo autovalor de  $sym(\mathbf{\Lambda}_0)$ ;

(ii) Se  $\alpha_{0min} \geq \lambda^I$ , faça

$$\mathbf{\Lambda} = sym(\mathbf{\Lambda}_0); \quad (1.37)$$

(iii) Caso contrário, faça

$$\mathbf{\Lambda} = sym(\mathbf{\Lambda}_0) + (\lambda^I - \lambda_{0min})\mathbf{I}_q. \quad (1.38)$$

**Passo 3d:** Retorne ao Passo 1.

## 1.3 Algoritmo PENLAB

PENLAB é um algoritmo de código aberto para otimização linear, não linear e semidefinida não linear. Este algoritmo permite resolver um problema de otimização com uma função objetivo não linear sujeito a restrições de igualdade e desigualdade não lineares e também a restrições matriciais semidefinidas formulado como

$$\left\{ \begin{array}{ll} \underset{\mathbf{x} \in \mathbb{R}^n, \mathbf{Y}_1 \in \mathcal{S}^{p_1}, \dots, \mathbf{Y}_k \in \mathcal{S}^{p_k}}{\text{minimizar}} & f(\mathbf{x}, \mathbf{Y}) \\ \text{sujeito a:} & g_i(\mathbf{x}, \mathbf{Y}) \leq 0, \quad i = 1, \dots, m_g \\ & h_i(\mathbf{x}, \mathbf{Y}) = 0, \quad i = 1, \dots, m_h \\ & \mathcal{A}_i(\mathbf{x}, \mathbf{Y}) \preceq \mathbf{0}, \quad i = 1, \dots, m_A \\ & \underline{\lambda}_i \mathbf{I} \preceq \mathbf{Y}_i \preceq \bar{\lambda}_i \mathbf{I}, \quad i = 1, \dots, k \end{array} \right. \quad (1.39)$$

onde  $\mathbf{x} \in \mathbb{R}^n$  é o vetor de variáveis;  $\mathbf{Y}_1 \in \mathcal{S}^{p_1}, \dots, \mathbf{Y}_k \in \mathcal{S}^{p_k}$  são as variáveis matriciais, denotando-se por  $\mathbf{Y} = [\mathbf{Y}_1, \dots, \mathbf{Y}_k]$ ;  $f$ ,  $g_i$  e  $h_i$  são as funções  $C^2$  de  $\mathbb{R}^n \times \mathcal{S}^{p_1} \times \dots \times \mathcal{S}^{p_k} \rightarrow \mathbb{R}$ ;  $\underline{\lambda}_i$  e  $\bar{\lambda}_i$  são os limites inferior e superior, respectivamente dos autovalores de  $\mathbf{Y}_i$ , para  $i = 1, \dots, k$  e  $\mathcal{A}_i(\mathbf{x}, \mathbf{Y})$  são os mapeamentos duas vezes diferenciáveis de  $\mathbb{R}^n \times \mathcal{S}^{p_1} \times \dots \times \mathcal{S}^{p_k} \rightarrow \mathcal{S}^{p_{A_i}}$  sendo  $p_{A_i}$ , para  $i = 1, \dots, m_A$ , inteiros

positivos. Aqui, a restrição de igualdade  $h_i$  é formulada como duas restrições de desigualdade.

Este algoritmo baseia-se em uma escolha de duas funções de penalidade/barreira  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  que penalize as restrições de desigualdade e  $\phi : \mathcal{S}^p \rightarrow \mathcal{S}^p$  que penalize as matrizes semidefinidas. Estas funções satisfazem certo número de propriedades que garantem que para qualquer  $\pi > 0$  e  $\Pi > 0$ , têm-se

$$\mathbf{z}(\mathbf{x}) \leq \mathbf{0} \iff \pi\varphi(\mathbf{z}(\mathbf{x})/\pi) \leq \mathbf{0}, \mathbf{z} \in C^2(\mathbb{R}^n \rightarrow \mathbb{R}) \quad (1.40)$$

e

$$\mathbf{Z}(\mathbf{x}) \preceq \mathbf{0} \iff \Pi\phi(\mathbf{z}/\Pi) \preceq \mathbf{0}, \mathbf{Z} \in \mathcal{S}^p \quad (1.41)$$

Isto significa que, para qualquer  $\pi > 0$  e  $\Pi > 0$ , o problema (1.39) tem a mesma solução do que o seguinte problema aumentado

$$\left\{ \begin{array}{l} \underset{\mathbf{x} \in \mathbb{R}^n, \mathbf{Y}_1 \in \mathcal{S}^{p_1}, \dots, \mathbf{Y}_k \in \mathcal{S}^{p_k}}{\text{minimizar}} \quad f(\mathbf{x}, \mathbf{Y}) \\ \text{sujeito a:} \quad \varphi_\pi(g_i(\mathbf{x}, \mathbf{Y})) \leq 0, \quad i = 1, \dots, m_g \\ \quad \quad \quad \phi_\Pi(\mathcal{A}_i(\mathbf{x}, \mathbf{Y})) \preceq \mathbf{0}, \quad i = 1, \dots, m_A \\ \quad \quad \quad \phi_\Pi(\underline{\lambda}_i \mathbf{I} - \mathbf{Y}_i) \preceq \mathbf{0}, \quad i = 1, \dots, k \\ \quad \quad \quad \phi_\Pi(\mathbf{Y}_i - \bar{\lambda}_i \mathbf{I}) \preceq \mathbf{0}, \quad i = 1, \dots, k \\ \quad \quad \quad h_i(\mathbf{x}, \mathbf{Y}) = 0, \quad i = 1, \dots, m_h \end{array} \right. \quad (1.42)$$

onde  $\varphi_\pi = \pi\varphi(\cdot/\pi)$  e  $\phi_\Pi = \Pi\phi(\cdot/\Pi)$ .

Desta maneira, o Lagrangiano do problema (1.42) pode ser tratado como uma generalização do Lagrangiano aumentado do problema (1.39) tal que

$$\begin{aligned} L(\mathbf{x}, \mathbf{Y}, \mathbf{u}, \underline{\Xi}, \underline{\mathbf{U}}, \bar{\mathbf{U}}, \mathbf{v}, \pi, \Pi) &= f(\mathbf{x}, \mathbf{Y}) + \sum_{i=1}^{m_g} u_i \varphi_\pi(g_i(\mathbf{x}, \mathbf{Y})) + \sum_{i=1}^{m_A} \langle \underline{\Xi}_i, \phi_\Pi(\mathcal{A}_i(\mathbf{x}, \mathbf{Y})) \rangle \\ &+ \sum_{i=1}^k \langle \underline{\mathbf{U}}_i, \phi_\Pi(\underline{\lambda}_i \mathbf{I} - \mathbf{Y}_i) \rangle + \sum_{i=1}^k \langle \bar{\mathbf{U}}_i, \phi_\Pi(\mathbf{Y}_i - \bar{\lambda}_i \mathbf{I}) \rangle + v^T \mathbf{h}(\mathbf{x}, \mathbf{Y}) \end{aligned} \quad (1.43)$$

onde  $\mathbf{u} \in \mathbb{R}^{m_g}$  é o vetor dos multiplicadores de Lagrange associados com as restrições de desigualdade;  $\underline{\Xi} = [\underline{\Xi}_1, \dots, \underline{\Xi}_{m_A}]$ ,  $\underline{\Xi}_i \in \mathcal{S}^{p_{A_i}}$ , e  $\underline{\mathbf{U}} = [\underline{\mathbf{U}}_1, \dots, \underline{\mathbf{U}}_k]$ ,  $\bar{\mathbf{U}} = [\bar{\mathbf{U}}_1, \dots, \bar{\mathbf{U}}_k]$ ,  $\underline{\mathbf{U}}_i, \bar{\mathbf{U}}_i \in \mathcal{S}^{p_i}$ , são os multiplicadores de Lagrange associados com as restrições matriciais semidefinidas e  $\mathbf{v} \in \mathbb{R}^{m_h}$  é o vetor dos multiplicadores de Lagrange associados com as restrições de igualdade.

O algoritmo PENLAB combina ideias dos métodos de penalidade (exterior) e de barreira (interior) com o método do Lagrangiano aumentado que consiste no seguinte:

**Parâmetros:** Dados  $\pi^1 > 0$ ,  $\Pi^1 > 0$  e  $\alpha^1$ .

**Dados Iniciais:**  $\mathbf{x}^1$ ,  $\mathbf{Y}^1$  e  $\mathbf{u}^1$ ,  $\underline{\Xi}^1$ ,  $\underline{\mathbf{U}}^1$ ,  $\bar{\mathbf{U}}^1$  e  $\mathbf{v}^1$ . Para  $l = 1, \dots$ , repetir até atingir um critério de parada fazer:

**Passo 1:** Calcular  $\mathbf{x}^{l+1}$ ,  $\mathbf{Y}^{l+1}$  e  $\mathbf{v}^{l+1}$  tal que

$$\left\| \nabla_{\mathbf{x}, \mathbf{Y}} L \left( \mathbf{x}^{l+1}, \mathbf{Y}^{l+1}, \mathbf{u}^l, \underline{\Xi}^l, \underline{\mathbf{U}}^l, \bar{\mathbf{U}}^l, \mathbf{v}^{l+1}, \pi^l, \Pi^l \right) \right\| \leq \alpha^l \quad (1.44)$$

e

$$\left\| \mathbf{h}(\mathbf{x}^{l+1}, \mathbf{Y}^{l+1}) \right\| \leq \alpha^l \quad (1.45)$$

**Passo 2:** Atualizar os multiplicadores de Lagrange

$$u_i^{l+1} = u_i^l \varphi'_{\pi^l} \left( g_i(\mathbf{x}^{l+1}, \mathbf{Y}^{l+1}) \right), \quad i = 1, \dots, m_g \quad (1.46)$$

$$\underline{\Xi}_i^{l+1} = D_{\mathcal{A}} \phi_{\Pi^l} \left( \mathcal{A}_i(\mathbf{x}^{l+1}, \mathbf{Y}^{l+1}); \underline{\Xi}_i^l \right), \quad i = 1, \dots, m_A \quad (1.47)$$

$$\underline{\mathbf{U}}_i^{l+1} = D_{\mathcal{A}} \phi_{\Pi^l} \left( (\lambda_i \mathbf{I} - \mathbf{Y}_i); \underline{\mathbf{U}}_i^l \right), \quad i = 1, \dots, k \quad (1.48)$$

e

$$\bar{\mathbf{U}}_i^{l+1} = D_{\mathcal{A}} \phi_{\Pi^l} \left( (\mathbf{Y}_i - \bar{\lambda}_i \mathbf{I}); \bar{\mathbf{U}}_i^l \right), \quad i = 1, \dots, k \quad (1.49)$$

Nestas expressões  $D\phi(\mathbf{X}; \mathbf{Y})$  representa a derivada direcional de  $\phi$  em relação a  $X$  na direção  $Y$ .

**Passo 3** Atualizar os parâmetros de penalidade

$$\pi^{l+1} < \pi^l \quad (1.50)$$

$$\Pi^{l+1} < \Pi^l \quad (1.51)$$

e

$$\alpha^{l+1} < \alpha^l \quad (1.52)$$

### 1.3.1 Observações

- No passo 1 do algoritmo pretende-se encontrar uma solução aproximada do seguinte sistema

$$\nabla_{\mathbf{x}, \mathbf{Y}} L(\mathbf{x}, \mathbf{Y}, \mathbf{u}, \underline{\Xi}, \underline{\mathbf{U}}, \bar{\mathbf{U}}, \mathbf{v}, \pi, \Pi) = \mathbf{0} \quad (1.53)$$

$$\mathbf{h}(\mathbf{x}, \mathbf{Y}) = \mathbf{0} \quad (1.54)$$

onde os parâmetros de penalidade  $\pi$  e  $\Pi$  assim como os multiplicadores de Lagrange  $\mathbf{u}$ ,  $\underline{\Xi}$ ,  $\underline{\mathbf{U}}$  e  $\bar{\mathbf{U}}$  são fixos. Para resolver este sistema é empregado o

método de Newton com amortecimento.

- O comprimento do passo é obtido usando a função de mérito do Lagrangiano aumentado definida como

$$L(\mathbf{x}, \mathbf{Y}, \mathbf{u}, \Xi, \underline{\mathbf{U}}, \bar{\mathbf{U}}, \mathbf{v}, \pi, \Pi) + \frac{1}{2\mu} \|\mathbf{h}(\mathbf{x}, \mathbf{Y})\|_2^2 \quad (1.55)$$

conjuntamente com uma regra de Armijo.

- Os parâmetros de penalidade  $\pi$  e  $\Pi$  no passo 3 do algoritmo são atualizados por algum fator constante que depende dos parâmetros de penalidade iniciais  $\pi^1$  e  $\Pi^1$ . O processo de atualização termina quando  $\pi_{eps}$  (por defeito  $10^{-6}$ ) é atingido.
- O algoritmo para quando um critério baseado no erro de KKT é satisfeito então são verificadas simultaneamente as seguintes desigualdades

$$\frac{|f(\mathbf{x}^l, \mathbf{Y}^l) - L(\mathbf{x}^l, \mathbf{Y}^l, \mathbf{u}^l, \Xi^l, \underline{\mathbf{U}}^l, \bar{\mathbf{U}}^l, \mathbf{v}^l, \pi^l, \Pi^l)|}{1 + |f(\mathbf{x}^l, \mathbf{Y}^l)|} < \epsilon \quad (1.56)$$

e

$$\frac{|f(\mathbf{x}^l, \mathbf{Y}^l) - f(\mathbf{x}^{l-1}, \mathbf{Y}^{l-1})|}{1 + |f(\mathbf{x}^l, \mathbf{Y}^l)|} < \epsilon \quad (1.57)$$

onde  $\epsilon$  tem um valor por defeito de  $10^{-6}$ .



# Capítulo 2

## Algoritmo de pontos interiores e direções viáveis para programação semidefinida (FDIPA-SDP-NS)

### 2.1 Introdução

No presente capítulo, apresenta-se o algoritmo de de pontos interiores e direções viáveis para programação semidefinida (FDIPA-SDP-NS) [24] acrescentando a análise de convergência do algoritmo.

As técnicas desenvolvidas no algoritmo FDIPA-SDP-NS estão baseadas no algoritmo FDIPA-SDP proposto por Aroztegui [8] onde foram reformulando os sistemas de equações lineares internos que tem que resolver o algoritmo com a finalidade de preservar a estrutura das restrições matriciais semidefinidas, tornando-se, desta maneira, computacionalmente mais eficiente. Esta reformulação implica a mudança do produto simétrico de Kronecker pelo produto não simétrico de Kronecker cujas definições e algumas propriedades podem ser achadas em [9]. A sigla NS foi empregada para diferenciar ambos os algoritmos.

Desta maneira, o algoritmo FDIPA-DIPA-NS permite resolver problemas de programação semidefinida não linear definidos como

$$\begin{cases} \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimizar}} & f(\mathbf{x}) \\ \text{sujeito a:} & \mathcal{A}(\mathbf{x}) \preceq 0 \end{cases} \quad (2.1)$$

As definições, conceitos e procedimentos descritos a continuação serão estendidos a um problema generalizado de programação semidefinida não linear no seguinte capítulo.

## 2.2 Conceitos básicos

Nesta seção apresentam-se alguns conceitos básicos e resultados teóricos associados ao Problema (2.1). Alguns destes conceitos foram introduzidos por Shapiro [6, 7].

### 2.2.1 Notação e terminologia

A notação empregada ao longo do presente trabalho é a seguinte:

- Denote-se por  $\Omega = \{\mathbf{x} \in \mathbb{R}^n : \mathcal{A}(\mathbf{x}) \preceq \mathbf{0}\}$  o conjunto de pontos viáveis do problema (2.1) e seu interior é denotado por  $int(\Omega)$ ;
- Para as matrizes  $\mathbf{A}, \mathbf{B} \in \mathcal{S}^q$ , a notação  $\mathbf{A} \succcurlyeq \mathbf{B}$  (ou  $\mathbf{A} \preceq \mathbf{B}$ ) significa que a matriz  $\mathbf{A} - \mathbf{B}$  é semidefinida positiva (ou semidefinida negativa). Da mesma forma, a notação  $\mathbf{A} \succ \mathbf{B}$  (ou  $\mathbf{A} \prec \mathbf{B}$ ) significa que a matriz  $\mathbf{A} - \mathbf{B}$  é definida positiva (ou definida negativa) [36];
- O conjunto de matrizes simétricas semidefinidas positivas (ou semidefinidas negativas) é denotado por  $\mathcal{S}_+^q$  (ou  $\mathcal{S}_-^q$ ) e o conjunto de matrizes simétricas definidas positivas (ou definidas negativas) é denotado por  $\mathcal{S}_{++}^q$  (ou  $\mathcal{S}_{--}^q$ ).
- Os autovalores de uma matriz  $\mathbf{A} \in \mathcal{S}^q$  são denotados por  $\lambda_1(\mathbf{A}) \leq \dots \leq \lambda_q(\mathbf{A})$ , dispostos na ordem crescente, sendo  $\lambda_{\min}(\mathbf{A}) = \lambda_1(\mathbf{A})$  e  $\lambda_{\max}(\mathbf{A}) = \lambda_q(\mathbf{A})$  os autovalores mínimo e máximo da matriz  $\mathbf{A}$ , respectivamente;
- Para toda matriz simétrica  $\mathbf{A} \in \mathbb{R}^{q \times q}$  consideramos os mapeamentos  $vec : \mathbb{R}^{q \times q} \rightarrow \mathbb{R}^{q^2}$  e  $mat : \mathbb{R}^{q^2} \rightarrow \mathbb{R}^{q \times q}$  que são definidos como

$$vec(\mathbf{A}) = [a_{11}, \dots, a_{n1}, a_{12}, \dots, a_{n2}, a_{1n}, \dots, a_{nn}]^T, \quad (2.2)$$

e  $mat$  a inversa do mapeamento  $vec$  [9].

- O produto interno de duas matrizes  $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{q \times q}$  é denotado por  $\langle \mathbf{A}, \mathbf{B} \rangle$ , e definido como

$$\langle \mathbf{A}, \mathbf{B} \rangle = tr(\mathbf{A}^T \mathbf{B}) = vec(\mathbf{A})^T vec(\mathbf{B}) \quad (2.3)$$

- Denote-se por  $\mathbf{A} \otimes \mathbf{B}$  o produto de Kronecker de duas matrizes  $\mathbf{A} \in \mathbb{R}^{p \times q}$  e  $\mathbf{B} \in \mathbb{R}^{r \times s}$  definido como

$$\mathbf{A} \otimes \mathbf{B} = [a_{ij} \mathbf{B}] = \begin{bmatrix} a_{11} \mathbf{B} & \cdots & a_{1q} \mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{p1} \mathbf{B} & \cdots & a_{pq} \mathbf{B} \end{bmatrix}, \quad i = 1, \dots, p, \quad j = 1, \dots, q \quad (2.4)$$

**Lema 2.1.** *As seguintes propriedades do produto de Kronecker são verdadeiras (Ver [9]):*

$$(i) \forall \mathbf{A} \in \mathbb{R}^{p \times q}, \mathbf{B} \in \mathbb{R}^{r \times s}, \mathbf{C} \in \mathbb{R}^{s \times q}$$

$$(\mathbf{A} \otimes \mathbf{B})\text{vec}(\mathbf{C}) = \text{vec}(\mathbf{BCA}^T) \quad (2.5)$$

$$(ii) \forall \mathbf{A} \in \mathbb{R}^{p \times q}, \mathbf{B} \in \mathbb{R}^{r \times s}$$

$$(\mathbf{A} \otimes \mathbf{B})^T = \mathbf{A}^T \otimes \mathbf{B}^T \quad (2.6)$$

$$(iii) \forall \mathbf{A} \in \mathbb{R}^{p \times q}, \mathbf{B} \in \mathbb{R}^{r \times s}, \mathbf{C} \in \mathbb{R}^{q \times k}, \mathbf{D} \in \mathbb{R}^{s \times l}$$

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD}) \quad (2.7)$$

$$(iv) \forall \mathbf{A} \in \mathbb{R}^{p \times p}, \mathbf{B} \in \mathbb{R}^{r \times r}$$

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1} \quad (2.8)$$

*Demonstração.* Ver [9]. □

- As matrizes das derivadas parciais do mapeamento  $\mathcal{A}(\mathbf{x})$  em relação a  $x_k$ , com  $k = 1, \dots, n$  é denotado por  $\mathcal{A}_k(\mathbf{x}) \in \mathbb{R}^{q \times q}$ , e definida como

$$\mathcal{A}_k(\mathbf{x}) = \frac{\partial \mathcal{A}(\mathbf{x})}{\partial x_k} = \begin{bmatrix} \frac{\partial a_{11}(\mathbf{x})}{\partial x_k} & \cdots & \frac{\partial a_{1q}(\mathbf{x})}{\partial x_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial a_{1q}(\mathbf{x})}{\partial x_k} & \cdots & \frac{\partial a_{qq}(\mathbf{x})}{\partial x_k} \end{bmatrix} \quad (2.9)$$

- Denote-se por  $\nabla \mathcal{A}(\mathbf{x}) \in \mathbb{R}^{n \times q^2}$  a derivada do mapeamento  $\mathcal{A}(\cdot)$  em  $\mathbf{x}$ , isto é,  $\nabla \mathcal{A}(\mathbf{x})$  é uma transformação linear de  $\mathbb{R}^n$  em  $\mathcal{S}^q$  definida por:

$$\nabla \mathcal{A}(\mathbf{x}) = \begin{bmatrix} \text{vec}(\mathcal{A}_1(\mathbf{x}))^T \\ \vdots \\ \text{vec}(\mathcal{A}_q(\mathbf{x}))^T \end{bmatrix} = \begin{bmatrix} \text{vec}\left(\frac{\partial \mathcal{A}(\mathbf{x})}{\partial x_1}\right)^T \\ \vdots \\ \text{vec}\left(\frac{\partial \mathcal{A}(\mathbf{x})}{\partial x_n}\right)^T \end{bmatrix} \quad (2.10)$$

- A derivada direcional do mapeamento  $\mathcal{A}(\cdot)$  na direção  $\mathbf{d} \in \mathbb{R}^n$ , denotada por  $\nabla \mathcal{A}(\mathbf{x})\mathbf{d} \in \mathbb{R}^{n \times q^2}$ , é definida como

$$D\mathcal{A}(\mathbf{x})\mathbf{d} = \sum_{p=1}^n d_p \frac{\partial \mathcal{A}(\mathbf{x})}{\partial x_p} \quad (2.11)$$

Aplicando a definição do mapeamento  $vec$  na Equação (2.11) resulta

$$vec(D\mathcal{A}(\mathbf{x})\mathbf{d}) = \nabla \mathcal{A}(\mathbf{x})^T \mathbf{d} \quad (2.12)$$

- A função Lagrangiana  $L : \mathbb{R}^n \times \mathbb{R}^q \rightarrow \mathbb{R}$  associada ao problema (2.1) é definida como

$$L(\mathbf{x}, \mathbf{\Lambda}) = f(\mathbf{x}) + \langle \mathbf{\Lambda}, \mathcal{A}(\mathbf{x}) \rangle \quad (2.13)$$

onde  $\mathbf{\Lambda} \in \mathcal{S}_+^q$ .

Empregando a definição do produto interno de duas matrizes fornecida pela Equação (2.3) e a definição do mapeamento  $vec$ , tal que  $\boldsymbol{\lambda} = vec(\mathbf{\Lambda})$ , tem-se

$$\langle \mathbf{\Lambda}, \mathcal{A}(\mathbf{x}) \rangle = vec(\mathbf{\Lambda})vec(\mathcal{A}(\mathbf{x})) = vec(\mathcal{A}(\mathbf{x}))^T \boldsymbol{\lambda} \quad (2.14)$$

Desta maneira, a função Lagrangiana pode ser redefinida como

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + vec(\mathcal{A}(\mathbf{x}))^T \boldsymbol{\lambda} \quad (2.15)$$

- O gradiente da função Lagrangiana em relação a  $\mathbf{x}$  é denotado por  $\nabla_{\mathbf{x}}L : \mathbb{R}^n \times \mathbb{R}^{q^2} \rightarrow \mathbb{R}^n$  e definido como

$$\nabla_{\mathbf{x}}L(\mathbf{x}, \boldsymbol{\lambda}) = \nabla f(\mathbf{x}) + \nabla \mathcal{A}(\mathbf{x})\boldsymbol{\lambda} \quad (2.16)$$

- A Hessiana da função Lagrangiana é denotada por  $\mathbf{H} : \mathbb{R}^n \times \mathbb{R}^{q^2} \rightarrow \mathbb{R}^{n \times n}$  e definida como

$$\mathbf{H}(\mathbf{x}, \boldsymbol{\lambda}) = \nabla^2 f(\mathbf{x}) + \nabla^2 \mathcal{A}(\mathbf{x})\boldsymbol{\lambda} \quad (2.17)$$

## 2.2.2 Definições

Este algoritmo emprega as mesmas definições do algoritmo FDIPA descritas na Seção 1.1.2 do Capítulo 1, tais como: direção de descida, direção viável e campo de direções uniformemente viáveis.

A seguinte proposição, introduzida por Aroztegui [8], estabelece uma condição para que uma direção  $\mathbf{d}$  verifique a Definição 1.4.

**Proposição 2.1.** *Seja a função  $\mathcal{A}$  diferenciável no ponto  $\mathbf{x} \in \Omega$  e  $\{b_1, \dots, b_r\}$  uma base ortonormal do  $\ker(\mathcal{A}(\mathbf{x}))$ . Seja também  $\mathbf{E}_0 = [\mathbf{b}_1, \dots, \mathbf{b}_r] \in \mathbb{R}^{q \times r}$ .*

Se  $\mathbf{E}_0^T \mathbf{\Pi} \mathbf{E}_0 \prec \mathbf{0}$  então,  $\mathbf{d} \in \mathbb{R}^n$  é uma direção viável do problema (2.1), onde  $\mathbf{\Pi} = \sum_{p=1}^n d_p \frac{\partial \mathcal{A}(\mathbf{x})}{\partial x_p}$ .

*Demonstração.* Ver [35]. □

### 2.2.3 Resultados auxiliares

A seguir, serão mostrados alguns resultados auxiliares [35, 37] referentes às propriedades de matrizes simétricas que serão utilizados na análise de convergência do presente algoritmo.

**Lema 2.2.** *Sejam as matrizes  $\mathbf{A} \in \mathcal{S}_{++}^q$  e  $\mathbf{B} \in \mathcal{S}^q$ . Se  $\mathbf{AB} + \mathbf{BA} \prec \mathbf{0}$  então  $\mathbf{B} \prec \mathbf{0}$ .*

*Demonstração.* Ver [35]. □

**Lema 2.3.** *Se  $\mathbf{A} \in \mathcal{S}_{++}^q$  e  $\mathbf{B} \in \mathcal{S}^q$  então a matriz  $\mathbf{AB}$  tem autovalores reais.*

*Demonstração.* Ver [37]. □

**Lema 2.4.** *Sejam  $\mathbf{A} \in \mathcal{S}_{++}^q$  e  $\mathbf{B} \in \mathcal{S}_-^q$ . Se elas comutam, então*

$$\mathbf{y}^T \mathbf{A} \mathbf{B} \mathbf{y} = 0 \iff \mathbf{B} \mathbf{y} = \mathbf{0} \quad (2.18)$$

*Demonstração.* Ver [35]. □

**Lema 2.5.** *Sejam  $\mathbf{A} \in \mathcal{S}_{++}^q$  e  $\mathbf{B} \in \mathcal{S}_-^q$ . Se elas comutam, então*

$$[\mathbf{A} \otimes \mathbf{I}_q]^{-1} [\mathbf{B} \otimes \mathbf{I}_q] \in \mathcal{S}_-^{q^2} \quad (2.19)$$

*Demonstração.* Sendo  $\mathbf{A} \succ \mathbf{0}$  e  $\mathbf{B} \preceq \mathbf{0}$  e elas comutam, então existe uma matriz não singular  $\mathbf{P} \in \mathbb{R}^{q \times q}$  tal que

$$\mathbf{A} = \mathbf{P} \mathbf{D}_A \mathbf{P}^{-1} \quad (2.20)$$

$$\mathbf{B} = \mathbf{P} \mathbf{D}_B \mathbf{P}^{-1} \quad (2.21)$$

Em vista das propriedades do produto de Kronecker do Lema 2.1, temos

$$\mathbf{A} \otimes \mathbf{I}_q = \mathbf{P}_P \mathbf{D}_{AI} \mathbf{P}_P^{-1} \quad (2.22)$$

$$\mathbf{B} \otimes \mathbf{I}_q = \mathbf{P}_P \mathbf{D}_{BI} \mathbf{P}_P^{-1} \quad (2.23)$$

onde  $\mathbf{P}_P = \mathbf{P} \otimes \mathbf{P}$ ,  $\mathbf{D}_{AI} = \mathbf{D}_A \otimes \mathbf{I}_q$  e  $\mathbf{D}_{BI} = \mathbf{D}_B \otimes \mathbf{I}_q$ . Aqui,  $\mathbf{D}_{AI}$  e  $\mathbf{D}_{BI}$  são matrizes diagonais e  $\mathbf{P}_P$  é uma matriz ortonormal.

Logo, da Equação (2.19), obtém-se

$$[\mathbf{A} \otimes \mathbf{I}_q]^{-1} [\mathbf{B} \otimes \mathbf{I}_q] = \mathbf{P}_P \mathbf{D}_{AI} \mathbf{D}_{BI} \mathbf{P}_P^T \quad (2.24)$$

Finalmente, sendo  $\mathbf{D}_{AI} \mathbf{D}_{BI} \prec \mathbf{0}$ , então  $\mathbf{P}_P \mathbf{D}_{AI} \mathbf{D}_{BI} \mathbf{P}_P^T \prec \mathbf{0}$  e, portanto,  $[\mathbf{A} \otimes \mathbf{I}_q]^{-1} [\mathbf{B} \otimes \mathbf{I}_q] \in \mathcal{S}_-^{q^2}$ .  $\square$

## 2.3 Condições de otimalidade de primeira ordem

As condições de otimalidade de primeira ordem de Karush-Kuhn-Tucker, introduzidas por Shapiro [6], correspondentes ao problema (2.1) podem ser formuladas da seguinte maneira

$$\nabla f(\mathbf{x}) + \nabla \mathcal{A}(\mathbf{x}) \boldsymbol{\lambda} = \mathbf{0} \quad (2.25a)$$

$$\boldsymbol{\Lambda} \mathcal{A}(\mathbf{x}) = \mathbf{0} \quad (2.25b)$$

$$\mathcal{A}(\mathbf{x}) \preceq \mathbf{0} \quad (2.25c)$$

$$\boldsymbol{\Lambda} \succeq \mathbf{0} \quad (2.25d)$$

A Equação (2.25a) é denominada a condição lagrangiana e a Equação (2.25b) é denominada a condição de complementariedade. Por outro lado, as Equações (2.25c) e (2.25d) são denominadas as condições de viabilidade.

Sem embargo, devido a que  $\mathcal{A}(\mathbf{x}) \preceq \mathbf{0}$  e  $\boldsymbol{\Lambda} \succeq \mathbf{0}$ , a condição de complementariedade  $\boldsymbol{\Lambda} \mathcal{A}(\mathbf{x}) = \mathbf{0}$ , fornecida pela Equação (2.25b), pode ser reescrita como

$$\langle \boldsymbol{\Lambda}, \mathcal{A}(\mathbf{x}) \rangle = \text{tr} (\boldsymbol{\Lambda} \mathcal{A}(\mathbf{x})) = 0, \quad (2.26)$$

a qual é conhecida como a condição de complementariedade padrão da restrição matricial.

A condição (2.26) implica que as matrizes  $\boldsymbol{\Lambda}$  e  $\mathcal{A}(\mathbf{x})$  comutam, ou seja, têm os mesmos autovetores. Por outro lado, verifica-se que  $\alpha_i \beta_i = 0$ , para  $i = 1, \dots, q$ , sendo  $\alpha_i \geq 0$  e  $\beta_i \leq 0$  os autovalores das matrizes  $\boldsymbol{\Lambda}$  e  $\mathcal{A}(\mathbf{x})$ , respectivamente [6].

A seguir são mostradas algumas definições referentes as condições de otimalidade de primeira ordem de Karush-Kuhn-Tucker estendidas a um problema de programação semidefinida [6].

**Definição 2.1 (Ponto regular).** Um ponto  $\mathbf{x}$  é um ponto regular do problema (2.1) se os vetores  $\nabla \mathcal{A}(\mathbf{x}) \mathbf{b}_i(\mathbf{x})$ , para  $i = 1, \dots, r$ , são linearmente independentes. Neste problema,  $\{\mathbf{b}_1(\mathbf{x}), \dots, \mathbf{b}_r(\mathbf{x})\} \subset \mathbb{R}^{q^2}$  representa uma base ortonormal do  $\ker (\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q)$ .

**Definição 2.2 (Ponto estacionário).** Um ponto regular  $\mathbf{x}$  é um ponto estacionário do problema (2.1) se existe uma matriz  $\Lambda \in \mathbb{R}^{q \times q}$  tal que as seguintes condições são verificadas

$$\begin{aligned}\nabla f(\mathbf{x}) + \nabla \mathcal{A}(\mathbf{x})\boldsymbol{\lambda} &= \mathbf{0} \\ \Lambda \mathcal{A}(\mathbf{x}) &= \mathbf{0} \\ \mathcal{A}(\mathbf{x}) &\preceq \mathbf{0}\end{aligned}$$

**Definição 2.3 (Ponto de Karush-Kuhn-Tucker).** Um ponto estacionário  $\mathbf{x}$  é um ponto de Karush-Kuhn-Tucker do problema (2.1) se todas as Equações (2.25a)-(2.25d) são verificadas simultaneamente.

## 2.4 Hipóteses sobre o problema

As hipóteses acerca das funções  $f$  e  $\mathcal{A}$  que foram consideradas no problema (2.1) são as seguintes:

**Hipótese 2.1.** *Existe um número real  $a$  tal que o conjunto  $\Omega_a = \{\mathbf{x} \in \Omega : f(\mathbf{x}) \leq a\}$  é compacto e tem um interior não vazio denotado por  $\text{int}(\Omega_a)$ .*

**Hipótese 2.2.** *Para qualquer ponto  $\mathbf{x} \in \text{int}(\Omega_a)$  verifica-se que  $\mathcal{A}(\mathbf{x}) \prec \mathbf{0}$ .*

**Hipótese 2.3.** *As funções  $f$  e  $\mathcal{A}$  são continuamente diferenciáveis em  $\Omega_a$  e suas derivadas  $\nabla f$  e  $\partial \mathcal{A} / \partial x_p$ , para  $p = 1, \dots, n$ , satisfazem a condição de Lipschitz.*

**Hipótese 2.4.** *Qualquer ponto  $\mathbf{x} \in \text{int}(\Omega_a)$  é um ponto regular do problema (2.1).*

**Hipótese 2.5.** *Em cada iteração os valores correntes de  $\Lambda$  e  $\mathcal{A}(\mathbf{x})$  comutam.*

Estas hipóteses são uma generalização das hipóteses empregadas por Herskovits [1] em um problema de otimização não linear onde as restrições de desigualdade podem ser representadas por uma restrição matricial diagonal.

## 2.5 Descrição do algoritmo

A descrição do algoritmo FDIPA-SDP-NS resume-se nos seguintes itens:

- Determinação da matriz do sistema linear.
- Cálculo da direção de busca (de descida e viável).
- Busca linear.

### 2.5.1 Matriz do sistema linear

A matriz do sistema linear é formada empregando a condição lagrangiana e a condição de complementariedade da restrição matricial, fornecidas pelas Equações (2.25a) e (2.25b), e definindo o vetor  $\mathbf{y} = (\mathbf{x}, \boldsymbol{\lambda})^T$  e a função vetorial  $\boldsymbol{\Psi} : \mathbb{R}^{n+q^2} \rightarrow \mathbb{R}^{n+q^2}$ , tal que

$$\boldsymbol{\Psi}(\mathbf{x}, \boldsymbol{\lambda}) = \begin{bmatrix} \Psi_l(\mathbf{x}, \boldsymbol{\lambda}) \\ \Psi_c(\mathbf{x}, \boldsymbol{\lambda}) \end{bmatrix} \quad (2.27)$$

sendo

$$\Psi_l(\mathbf{x}, \boldsymbol{\lambda}) = \nabla f(\mathbf{x}) + \nabla \mathcal{A}(\mathbf{x}) \boldsymbol{\lambda} \quad (2.28)$$

$$\Psi_c(\mathbf{x}, \boldsymbol{\lambda}) = \boldsymbol{\Lambda} \mathcal{A}(\mathbf{x}) \quad (2.29)$$

O gradiente da função vetorial  $\boldsymbol{\Psi}(\mathbf{x}, \boldsymbol{\lambda})$  em relação às variáveis primais e duais é denotado por  $\nabla \boldsymbol{\Psi}(\mathbf{x}, \boldsymbol{\lambda})$  e definido como

$$\nabla \boldsymbol{\Psi}(\mathbf{x}, \boldsymbol{\lambda}) = \begin{bmatrix} \nabla_{\mathbf{x}} \Psi_l(\mathbf{x}, \boldsymbol{\lambda}) & \nabla_{\boldsymbol{\lambda}} \Psi_l(\mathbf{x}, \boldsymbol{\lambda}) \\ \nabla_{\mathbf{x}} \Psi_c(\mathbf{x}, \boldsymbol{\lambda}) & \nabla_{\boldsymbol{\lambda}} \Psi_c(\mathbf{x}, \boldsymbol{\lambda}) \end{bmatrix}, \quad (2.30)$$

que representa a matriz Jacobiana de  $\boldsymbol{\Psi}$ .

Derivando a Equação (2.28) em relação às variáveis  $\mathbf{x}$  e  $\boldsymbol{\lambda}$ , resulta

$$\nabla_{\mathbf{x}} \Psi_l(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{H}(\mathbf{x}, \boldsymbol{\lambda}) \quad (2.31)$$

$$\nabla_{\boldsymbol{\lambda}} \Psi_l(\mathbf{x}, \boldsymbol{\lambda}) = \nabla \mathcal{A}(\mathbf{x}) \quad (2.32)$$

Para obter as derivadas de  $\Psi_c$  em relação às variáveis  $\mathbf{x}$  e  $\boldsymbol{\lambda}$  é necessário, primeiramente, estabelecer a seguinte relação

$$\boldsymbol{\Lambda} \mathcal{A}(\mathbf{x}) \mathbf{I}_q = \mathbf{I}_q \boldsymbol{\Lambda} \mathcal{A}(\mathbf{x}) \quad (2.33)$$

Empregando a definição do mapeamento *vec* fornecida pela Equação (2.2) a ambos os lados da Equação (2.33), obtém-se

$$\text{vec}(\boldsymbol{\Lambda} \mathcal{A}(\mathbf{x}) \mathbf{I}_q) = \text{vec}(\mathbf{I}_q \boldsymbol{\Lambda} \mathcal{A}(\mathbf{x})) \quad (2.34)$$

Logo, aplicando as propriedades do Produto de Kronecker do Lema 2.1 na Equação (2.34) temos as seguintes expressões

$$\text{vec}(\boldsymbol{\Lambda} \mathcal{A}(\mathbf{x}) \mathbf{I}_q) = [\mathbf{I}_q^T \otimes \boldsymbol{\Lambda}] \text{vec}(\mathcal{A}(\mathbf{x})) = [\mathbf{I}_q \otimes \boldsymbol{\Lambda}] \text{vec}(\mathcal{A}(\mathbf{x})) \quad (2.35)$$

$$\text{vec}(\mathcal{A}(\mathbf{x}) \mathbf{I}_q \boldsymbol{\Lambda}) = [\mathcal{A}(\mathbf{x})^T \otimes \mathbf{I}_q] \text{vec}(\boldsymbol{\Lambda}) = [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \text{vec}(\boldsymbol{\Lambda}) \quad (2.36)$$



Derivando as Equações (2.35) e (2.36) em relação às variáveis  $\mathbf{x}$  e  $\boldsymbol{\lambda}$ , obtém-se as seguintes expressões

$$\nabla_{\mathbf{x}} \Psi_c(\mathbf{x}, \boldsymbol{\lambda}) = [\mathbf{I}_q \otimes \boldsymbol{\Lambda}] \nabla \mathcal{A}(\mathbf{x})^T \quad (2.37)$$

$$\nabla_{\boldsymbol{\lambda}} \Psi_c(\mathbf{x}, \boldsymbol{\lambda}) = \mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q \quad (2.38)$$

Assim, a Equação (2.30) pode ser reescrita como

$$\nabla \Psi(\mathbf{x}, \boldsymbol{\lambda}) = \begin{bmatrix} \mathbf{H}(\mathbf{x}, \boldsymbol{\lambda}) & \nabla \mathcal{A}(\mathbf{x}) \\ [\mathbf{I}_q \otimes \boldsymbol{\Lambda}] \nabla \mathcal{A}(\mathbf{x})^T & \mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q \end{bmatrix} \quad (2.39)$$

Finalmente, substitui-se a Hessiana da função lagrangiana  $\mathbf{H}(\mathbf{x}^k, \boldsymbol{\lambda}^k)$  por uma matriz  $\mathbf{B}^k$  simétrica definida positiva que pode ser considerada como uma aproximação da Hessiana obtida por alguma técnica quase-Newton [38] ou por uma matriz identidade. Desta maneira, obtém-se a seguinte matriz inversível  $\mathbf{W} \in \mathbb{R}^{(n+q^2) \times (n+q^2)}$

$$\mathbf{W}(\mathbf{x}, \mathbf{B}, \boldsymbol{\Lambda}) = \begin{bmatrix} \mathbf{B} & \nabla \mathcal{A}(\mathbf{x}) \\ [\mathbf{I}_q \otimes \boldsymbol{\Lambda}] \nabla \mathcal{A}(\mathbf{x})^T & \mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q \end{bmatrix}, \quad (2.40)$$

denominada matriz do sistema linear.

## 2.5.2 Cálculo da direção de busca

Em cada iteração do algoritmo é calculada uma direção de busca  $\mathbf{d}^k$ . Pretende-se que  $\mathbf{d}^k$  seja uma direção viável para o problema (2.1) e também de descida para a função objetivo  $f$  no ponto  $\mathbf{x}^k$ .

Para isso, emprega-se uma iteração de Newton com o ponto corrente  $\mathbf{y}^k = (\mathbf{x}^k, \boldsymbol{\lambda}^k)^T$  para obter um novo ponto  $\mathbf{y}^{k+1} = (\mathbf{x}^{k+1}, \boldsymbol{\lambda}_0^{k+1})^T$  através da seguinte expressão

$$\nabla \Psi(\mathbf{y}^k) (\mathbf{y}^{k+1} - \mathbf{y}^k) = -\Psi(\mathbf{y}^k) \quad (2.41)$$

Substituindo as Equações (2.27) e (2.39) na Equação (2.41), obtém-se o seguinte sistema linear

$$\begin{bmatrix} \mathbf{B}^k & \nabla \mathcal{A}(\mathbf{x}^k) \\ [\mathbf{I}_q \otimes \boldsymbol{\Lambda}^k] \nabla \mathcal{A}(\mathbf{x}^k)^T & \mathcal{A}(\mathbf{x}^k) \otimes \mathbf{I}_q \end{bmatrix} \begin{bmatrix} \mathbf{x}^{k+1} - \mathbf{x}^k \\ \boldsymbol{\lambda}_0^{k+1} - \boldsymbol{\lambda}^k \end{bmatrix} = - \begin{bmatrix} \nabla f(\mathbf{x}^k) + \nabla \mathcal{A}(\mathbf{x}^k) \boldsymbol{\lambda}^k \\ \text{vec}(\boldsymbol{\Lambda}^k \mathcal{A}(\mathbf{x}^k)) \end{bmatrix} \quad (2.42)$$

sendo  $\boldsymbol{\lambda}^k = \text{vec}(\boldsymbol{\Lambda}^k)$  e  $\boldsymbol{\lambda}_0^{k+1} = \text{vec}(\boldsymbol{\Lambda}_0^{k+1})$ .

Define-se a direção  $\mathbf{d}_0^k$  no espaço primal como  $\mathbf{d}_0^k = \mathbf{x}^{k+1} - \mathbf{x}^k$ . Desta maneira, o sistema linear da Equação (2.42) nas variáveis  $\mathbf{d}_0^k$  e  $\boldsymbol{\lambda}_0^{k+1}$  pode ser reescrito como

$$\mathbf{B}^k \mathbf{d}_0^k + \nabla \mathcal{A}(\mathbf{x}^k) \boldsymbol{\lambda}_0^{k+1} = -\nabla f(\mathbf{x}^k) \quad (2.43a)$$

$$\left[ \mathbf{I}_q \otimes \boldsymbol{\Lambda}^k \right] \nabla \mathcal{A}(\mathbf{x}^k)^T \mathbf{d}_0^k + \left[ \mathcal{A}(\mathbf{x}^k) \otimes \mathbf{I}_q \right] \boldsymbol{\lambda}_0^{k+1} = 0 \quad (2.43b)$$

onde a solução deste sistema fornece uma direção  $\mathbf{d}_0^k$  no espaço primal e uma nova estimativa dos multiplicadores de Lagrange  $\boldsymbol{\lambda}_0^{k+1}$ .

A seguir, analisam-se duas possíveis soluções da direção  $\mathbf{d}_0^k$  obtidas pelo sistema (2.43). No primeiro caso assume-se que a direção  $\mathbf{d}_0^k = \mathbf{0}$  e no segundo caso que  $\mathbf{d}_0^k \neq \mathbf{0}$ .

- **Caso 1:  $\mathbf{d}_0^k = \mathbf{0}$**

Se  $\mathbf{d}_0 = \mathbf{0}$ , o sistema de equações (2.43) resulta

$$\nabla f(\mathbf{x}^k) + \nabla \mathcal{A}(\mathbf{x}^k) \boldsymbol{\lambda}_0^{k+1} = \mathbf{0} \quad (2.44a)$$

$$\left[ \mathcal{A}(\mathbf{x}^k) \otimes \mathbf{I}_q \right] \boldsymbol{\lambda}_0^{k+1} = 0 \quad (2.44b)$$

Como  $\mathcal{A}(\mathbf{x}^k) \prec \mathbf{0}$ , então a matriz  $\left[ \mathcal{A}(\mathbf{x}^k) \otimes \mathbf{I}_q \right]$  é não singular o que implica que  $\boldsymbol{\lambda}_0^{k+1} = \mathbf{0}$  e também  $\boldsymbol{\Lambda}_0^{k+1} = \text{smat}(\boldsymbol{\lambda}_0^{k+1}) = \mathbf{0}$ . Desta maneira, verifica-se a condição de complementariedade  $\boldsymbol{\Lambda}_0^{k+1} \mathcal{A}(\mathbf{x}^k) = \mathbf{0}$ .

Logo, da Equação (2.44a) tem-se que  $\nabla f(\mathbf{x}^k) = \mathbf{0}$  e, portanto, pode-se concluir que o ponto  $\mathbf{x}^k$  é um ponto estacionário do problema (2.1). Se satisfaz também as condições de viabilidade então será um ponto KKT do problema (2.1) e, desta maneira, o algoritmo para.

- **Caso 2:  $\mathbf{d}_0^k \neq \mathbf{0}$**

Para o caso que  $\mathbf{d}_0 \neq \mathbf{0}$ , prova-se que a direção  $\mathbf{d}_0^k$  obtida pelas Equações (2.43a) e (2.43b) é uma direção de descida da função  $f(\mathbf{x})$  se  $\mathbf{d}_0^{kT} \nabla f(\mathbf{x}^k) < 0$ .

Sem embargo,  $\mathbf{d}_0^k$  pode não ser uma direção viável se um ponto está na fronteira da região viável  $\Omega$ , ou seja, quando alguns dos autovalores da matriz  $\mathcal{A}(\mathbf{x}^k)$  se aproximam de zero. Para tal caso, existe uma base ortonormal do  $\ker(\mathcal{A}(\mathbf{x}^k))$  definida pelo conjunto  $\{\mathbf{b}_1, \dots, \mathbf{b}_r\}$ . Seja a matriz  $\mathbf{E}_0 = [\mathbf{b}_1, \dots, \mathbf{b}_r] \in \mathbb{R}^{q \times r}$ , logo, empregando a Equação (2.43b) e aplicando a definição do mapeamento  $\text{vec}$ , obtêm-se as seguintes equivalências

$$\left[ \mathbf{I}_q \otimes \boldsymbol{\Lambda}^k \right] \text{vec} \left( \sum_{i=1}^n \mathbf{d}_{0i}^k \mathcal{A}_i(\mathbf{x}^k) \right) + \text{vec} \left( \boldsymbol{\Lambda}_0^k \mathcal{A}(\mathbf{x}^k) \right) = \mathbf{0} \quad (2.45)$$

e

$$\text{vec} \left( \Lambda^k \sum_{i=1}^n \mathbf{d}_{0i}^k \mathcal{A}_i(\mathbf{x}^k) + \Lambda_0^k \mathcal{A}(\mathbf{x}^k) \right) = \mathbf{0} \quad (2.46)$$

Aplicando a definição de *smat* à Equação (2.46), obtém-se

$$\Lambda^k \sum_{i=1}^n \mathbf{d}_{0i}^k \mathcal{A}_i(\mathbf{x}^k) + \Lambda_0^k \mathcal{A}(\mathbf{x}^k) = \mathbf{0} \quad (2.47)$$

Pré-multiplicando por  $\mathbf{E}_0^T$  e pós-multiplicando por  $\mathbf{E}_0$  ambos os termos da Equação (2.47) resulta

$$\mathbf{E}_0^T \Lambda^k \sum_{i=1}^n \mathbf{d}_{0i}^k \mathcal{A}_i(\mathbf{x}^k) \mathbf{E}_0 + \mathbf{E}_0^T \Lambda_0^k \mathcal{A}(\mathbf{x}^k) \mathbf{E}_0 = \mathbf{0} \quad (2.48)$$

Como  $\mathbf{E}_0^T \Lambda_0^k \mathcal{A}(\mathbf{x}^k) \mathbf{E}_0 = \mathbf{0}$ , então

$$\mathbf{E}_0^T \Lambda^k \sum_{i=1}^n \mathbf{d}_{0i}^k \mathcal{A}_i(\mathbf{x}^k) \mathbf{E}_0 = \mathbf{0} \quad (2.49)$$

Assumindo que  $\Lambda^k = \mu \mathbf{I}_q \succ \mathbf{0}$ , obtém-se

$$\mathbf{E}_0^T \sum_{i=1}^n \mathbf{d}_{0i}^k \mathcal{A}_i(\mathbf{x}^k) \mathbf{E}_0 = \mathbf{0}, \quad (2.50)$$

Portanto, pela Proposição 2.1, pode-se concluir que a Equação (2.50) não garante que a direção  $\mathbf{d}_0^k$  seja viável.

### 2.5.3 Direção viável

Com a finalidade de obter uma direção viável, Herskovits [1] propõe uma pequena deflexão na direção  $\mathbf{d}_0^k$  de tal maneira que esta nova direção fique sempre dentro da região viável. Esta deflexão é feita adicionando um termo negativo no lado direito da Equação (2.43b) do primeiro sistema linear conservando a mesma matriz  $\mathbf{W}(\mathbf{x}, \mathbf{B}, \Lambda)$  deste sistema.

Desta maneira, define-se um segundo sistema linear nas variáveis  $\mathbf{d}^k$  e  $\bar{\boldsymbol{\lambda}}^{k+1}$ , assim temos

$$\mathbf{B}^k \mathbf{d}^k + \nabla \mathcal{A}(\mathbf{x}^k) \bar{\boldsymbol{\lambda}}^{k+1} = -\nabla f(\mathbf{x}^k) \quad (2.51a)$$

$$\left[ \mathbf{I}_q \otimes \Lambda^k \right] \nabla \mathcal{A}(\mathbf{x}^k)^T \mathbf{d}^k + \left[ \mathcal{A}(\mathbf{x}^k) \otimes \mathbf{I}_q \right] \bar{\boldsymbol{\lambda}}^{k+1} = -\rho^k \boldsymbol{\lambda}^k \quad (2.51b)$$

onde  $\rho^k$  é um número real positivo. Este novo sistema fornece uma nova direção de busca  $\mathbf{d}^k$  e as novas estimativas dos multiplicadores de Lagrange  $\bar{\boldsymbol{\lambda}}^{k+1}$ .

A seguir, mostra-se que  $\mathbf{d}^k$  é uma direção viável considerando novamente o caso quando alguns dos autovalores da matriz  $\mathcal{A}(\mathbf{x}^k)$  se aproximam de zero e aplicando o procedimento descrito anteriormente.

Da Equação (2.51b) e aplicando a definição do mapeamento *vec*, tem-se

$$\mathbf{\Lambda}^k \sum_{i=1}^n \mathbf{d}_i^k \mathcal{A}_i(\mathbf{x}^k) + \bar{\mathbf{\Lambda}}^k \mathcal{A}(\mathbf{x}^k) = -\rho^k \mathbf{\Lambda}^k \quad (2.52)$$

Pré-multiplicando por  $\mathbf{E}_0^T$  e pós-multiplicando por  $\mathbf{E}_0$  ambos os termos da Equação (2.52), obtém-se

$$\mathbf{E}_0^T \sum_{i=1}^n \mathbf{d}_i^k \mathcal{A}_i(\mathbf{x}^k) \mathbf{E}_0 + \mathbf{E}_0^T \bar{\mathbf{\Lambda}}^k \mathcal{A}(\mathbf{x}^k) \mathbf{E}_0 = -\rho^k \mathbf{E}_0^T \mathbf{\Lambda}^k \mathbf{E}_0 \quad (2.53)$$

como  $\mathbf{E}_0^T \bar{\mathbf{\Lambda}}^k \mathcal{A}(\mathbf{x}^k) \mathbf{E}_0 = \mathbf{0}$ , então

$$\mathbf{E}_0^T \sum_{i=1}^n \mathbf{d}_i^k \mathcal{A}_i(\mathbf{x}^k) \mathbf{E}_0 = -\rho^k \mathbf{E}_0^T \mathbf{\Lambda}^k \mathbf{E}_0 \quad (2.54)$$

Assumindo que  $\mathbf{\Lambda}^k = \mu \mathbf{I}_q \succ \mathbf{0}$  e  $\rho^k > 0$ , então

$$\mathbf{E}_0^T \sum_{i=1}^n \mathbf{d}_i^k \mathcal{A}_i(\mathbf{x}^k) \mathbf{E}_0 = -\rho^k \mathbf{I}_q \prec \mathbf{0} \quad (2.55)$$

e em virtude da Proposição 2.1, pode-se concluir que a direção  $\mathbf{d}^k$  é viável, faltando provar que seja também uma direção de descida.

O cálculo de  $\mathbf{d}^k$  e  $\bar{\mathbf{\Lambda}}^{k+1}$  também pode ser realizado resolvendo o seguinte sistema linear auxiliar

$$\mathbf{B}^k \mathbf{d}_1^k + \nabla \mathbf{g}(\mathbf{x}^k) \gamma_1^{k+1} + \nabla \mathcal{A}(\mathbf{x}^k) \boldsymbol{\lambda}_1^{k+1} = \mathbf{0} \quad (2.56a)$$

$$\left[ \mathbf{I}_q \otimes \mathbf{\Lambda}^k \right] \nabla \mathcal{A}(\mathbf{x}^k)^T \mathbf{d}_1^k + \left[ \mathcal{A}(\mathbf{x}^k) \otimes \mathbf{I}_q \right] \boldsymbol{\lambda}_1^{k+1} = -\boldsymbol{\lambda}^k \quad (2.56b)$$

que fornece a direção  $\mathbf{d}_1^k$  e as novas estimativas dos multiplicadores de Lagrange  $\boldsymbol{\lambda}_1^{k+1}$ , e tomando as seguintes seguintes relações

$$\mathbf{d}^k = \mathbf{d}_0^k + \rho^k \mathbf{d}_1^k \quad (2.57)$$

$$\bar{\boldsymbol{\lambda}}^{k+1} = \boldsymbol{\lambda}_0^{k+1} + \rho^k \boldsymbol{\lambda}_1^{k+1} \quad (2.58)$$

Pode-se observar na Equação (2.57) que a direção de busca  $\mathbf{d}^k$  é uma composição das direções  $\mathbf{d}_0^k$  e  $\mathbf{d}_1^k$ .

Por outro lado, a deflexão da direção  $\mathbf{d}_0^k$  é proporcional ao parâmetro  $\rho$  o qual controla o grau de deflexão. Como  $\mathbf{d}_0^k$  é uma direção de descida da função objetivo

no ponto  $\mathbf{x}^k$ , então é necessário estabelecer os limites deste parâmetro que assegurem que a direção  $\mathbf{d}^k$  seja também uma direção de descida. Isto é possível, definindo um novo parâmetro  $\xi \in (0, 1)$  e impondo que

$$\mathbf{d}^{kT} \nabla f(\mathbf{x}^k) \leq \xi \mathbf{d}_0^{kT} \nabla f(\mathbf{x}^k), \quad (2.59)$$

onde  $\xi > 0$ .

Como  $\mathbf{d}_0^{kT} \nabla f(\mathbf{x}^k) < 0$  isto implica que  $\mathbf{d}^{kT} \nabla f(\mathbf{x}^k) < 0$  e, portanto, garante que  $\mathbf{d}^{kT}$  é uma direção de descida.

Em geral, a taxa de decréscimo da função objetivo ao longo de  $\mathbf{d}^k$  é  $1/\xi$  vezes menor que a longo de  $\mathbf{d}_0^k$ , constituindo-se no custo a pagar para que a nova direção de busca calculada seja viável. A escolha inicial do parâmetro  $\xi$  poderia influir na velocidade de convergência do algoritmo. Se o valor de  $\xi$  é próximo de 0 então a deflexão será muito grande e a velocidade de convergência será mais lenta, entretanto, um valor próximo de 1 originará uma velocidade de convergência mais rápida.

Substituindo a Equação (2.57) na Equação (2.59), pode-se escrever

$$\mathbf{d}_0^{kT} \nabla f(\mathbf{x}^k) + \rho^k \mathbf{d}_1^{kT} \nabla f(\mathbf{x}^k) \leq \rho^k \mathbf{d}_0^{kT} \nabla f(\mathbf{x}^k) \quad (2.60)$$

E assim, pode-se definir a cota superior de  $\rho$  da forma

$$\rho^k \leq (\xi - 1) \frac{\mathbf{d}_0^{kT} \nabla f(\mathbf{x}^k)}{\mathbf{d}_1^{kT} \nabla f(\mathbf{x}^k)} \quad (2.61)$$

se  $\mathbf{d}_1^{kT} \nabla f(\mathbf{x}^k) > 0$ . Caso contrário, se  $\mathbf{d}_1^{kT} \nabla f(\mathbf{x}^k) \leq 0$ , assume-se que

$$\rho^k \leq \varphi \left\| \mathbf{d}_0^k \right\|^2 \quad (2.62)$$

para algum parâmetro fixo  $\varphi > 0$ , definido em [4].

#### 2.5.4 Busca linear

Para obter um novo ponto  $\mathbf{x}^{k+1}$  que seja viável realiza-se uma minimização inexata de Armijo [38] ao longo da direção  $\mathbf{d}^k$  de maneira que  $\mathbf{x}^{k+1} \in \text{int}(\Omega_a)$  e proporcione um decréscimo da função objetivo.

A busca linear de Armijo consiste em achar um comprimento de passo  $t$ , ou seja, o primeiro elemento da sequência  $\{1, \nu, \nu^2, \nu^3, \dots\}$ , que reduza o valor da função objetivo e mantenha a viabilidade das restrições, tal que

$$f(\mathbf{x}^k + t^k \mathbf{d}^k) \leq f(\mathbf{x}^k) + t^k \eta \nabla f(\mathbf{x}^k)^T \mathbf{d}^k \quad (2.63)$$

e

$$\mathcal{A}(\mathbf{x}^k + t^k \mathbf{d}^k) \prec \mathbf{0} \quad (2.64)$$

onde  $\eta \in (0, 1)$  e  $\nu \in (0, 1)$  são parâmetros fixos.

## 2.6 Algoritmo FDIPA-SDP-NS

O algoritmo FDIPA-SDP-NS [24] consiste no seguinte:

**Parâmetros de controle:**  $\xi \in (0, 1)$ ,  $\eta \in (0, 1)$ ,  $\varphi > 0$ , e  $\nu \in (0, 1)$ .

**Dados Iniciais:**  $\mathbf{x} \in \text{int}(\Omega)$ ,  $\mathbf{B} \in \mathcal{S}_{++}^n$  e  $\Lambda \in \mathcal{S}_{++}^q$  comutando com  $\mathcal{A}(\mathbf{x})$ .

**Passo 1: Cálculo da direção de busca**

**Passo 1a:** Resolva o sistema linear para obter  $\mathbf{d}_0 \in \mathbb{R}^n$  e  $\boldsymbol{\lambda}_0 \in \mathbb{R}^m$ .

$$\begin{bmatrix} \mathbf{B} & \nabla \mathcal{A}(\mathbf{x}) \\ [\mathbf{I}_q \otimes \Lambda] \nabla \mathcal{A}(\mathbf{x})^T & \mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q \end{bmatrix} \begin{bmatrix} \mathbf{d}_0 \\ \boldsymbol{\lambda}_0 \end{bmatrix} = \begin{bmatrix} -\nabla f(\mathbf{x}) \\ \mathbf{0} \end{bmatrix} \quad (2.65)$$

Se  $\mathbf{d}_0 = \mathbf{0}$ , parar.

**Passo 1b:** Resolva o sistema linear auxiliar para obter  $\mathbf{d}_1 \in \mathbb{R}^n$  e  $\boldsymbol{\lambda}_1 \in \mathbb{R}^m$

$$\begin{bmatrix} \mathbf{B} & \nabla \mathcal{A}(\mathbf{x}) \\ [\mathbf{I}_q \otimes \Lambda] \nabla \mathcal{A}(\mathbf{x})^T & \mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q \end{bmatrix} \begin{bmatrix} \mathbf{d}_1 \\ \boldsymbol{\lambda}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ -\boldsymbol{\lambda} \end{bmatrix} \quad (2.66)$$

**Passo 1c:** Se  $\mathbf{d}_1^T \nabla f(\mathbf{x}) > 0$ , calcule

$$\rho = \min \left\{ \varphi \|\mathbf{d}_0\|^2, (\xi - 1) \frac{\mathbf{d}_0^T \nabla f(\mathbf{x})}{\mathbf{d}_1^T \nabla f(\mathbf{x})} \right\} \quad (2.67)$$

caso contrário, calcule

$$\rho = \varphi \|\mathbf{d}_0\|^2 \quad (2.68)$$

**Passo 1d:** Calcule a direção de busca  $\mathbf{d}$  e  $\bar{\Lambda}$

$$\mathbf{d} = \mathbf{d}_0 + \rho \mathbf{d}_1 \quad (2.69)$$

e

$$\bar{\Lambda} = \Lambda_0 + \rho \Lambda_1 \quad (2.70)$$

**Passo 2: Busca Linear**

Calcule o comprimento do passo  $t$ , o primeiro número da série  $\{1, \nu, \nu^2, \nu^3, \dots\}$ ,

tal que:

$$f(\mathbf{x} + t\mathbf{d}) \leq f(\mathbf{x}) + t\eta\nabla f(\mathbf{x})^T \mathbf{d} \quad (2.71)$$

e

$$\mathcal{A}(\mathbf{x} + t\mathbf{d}) \prec \mathbf{0} \quad (2.72)$$

### Passo 3 Atualizações

**Passo 3a:** Obtenha o novo ponto, como:  $\mathbf{x} := \mathbf{x} + t\mathbf{d}$ ;

**Passo 3b:** Defina novos valores para  $\mathbf{B} \in \mathcal{S}_{++}^n$ ;

**Passo 3c:** Defina novos valores para  $\mathbf{\Lambda} \in \mathcal{S}_{++}^m$  comutando com  $\mathcal{A}(\mathbf{x})$ ;

**Passo 3d:** Retorne ao Passo 1.

## 2.6.1 Hipóteses do algoritmo

As seguintes hipóteses acerca de  $\mathbf{\Lambda}$  e  $\mathbf{B}$  são necessárias para provar a convergência global do algoritmo.

**Hipótese 2.6.** *Existem números positivos  $\lambda^I$  e  $\lambda^S$  tal que*

$$\lambda^I \mathbf{I} \preceq \mathbf{\Lambda} \preceq \lambda^S \mathbf{I}$$

**Hipótese 2.7.** *Existem números positivos  $\sigma_1$  e  $\sigma_2$  tal que*

$$\sigma_1 \|\mathbf{d}\|^2 \leq \mathbf{d}^T \mathbf{B} \mathbf{d} \leq \sigma_2 \|\mathbf{d}\|^2, \quad \mathbf{d} \in \mathbb{R}^n$$

## 2.6.2 Atualizações

Realizam-se certas operações que preparam os dados para uma nova iteração, tais como:

### (i) Rotina de atualização de $\mathbf{\Lambda}$

A rotina atualização da matriz dos multiplicadores de Lagrange  $\mathbf{\Lambda}^{k+1}$  que tem que comutar com  $\mathcal{A}(\mathbf{x}^{k+1})$  foi definida na Subseção 1.2.1 [8] a qual que deve verificar a Hipótese 2.6. Sem embargo esta hipótese é verificada somente no limite.

### (ii) Rotina de atualização de $\mathbf{B}$

A rotina de atualização da matriz quase-Newton  $\mathbf{B}^k$  emprega a fórmula BFGS descrita em Luenberger [38] que define uma nova matriz simétrica definida positiva a qual deve satisfazer a Hipótese 2.7.

## 2.7 Análise de convergência global

Nesta seção, realiza-se a prova de convergência global do algoritmo FDIPA-GSDP-NS que consiste no seguinte:

- Primeiramente, o Lema 2.6 garante que os sistemas (2.43) e (2.56) são bem definidos, ou seja, têm soluções únicas e, portanto, o algoritmo não falha;
- No Lema 2.7 prova-se que se o algoritmo para depois de um número finito de iterações, então, ele gera um ponto de KKT do problema (2.1).
- Nos Lemas 2.8 e 2.9 provam-se que as direções  $\mathbf{d}_0$  e  $\mathbf{d}$  calculadas no passo 1 do algoritmo são direções de descida da função objetivo.
- Finalmente, no Lema 2.13 prova-se que qualquer ponto de acumulação gerado pelo algoritmo é um ponto estacionário do problema (2.1).

**Lema 2.6.** *Assume-se que  $\mathbf{x} \in \Omega_a$  é um ponto regular do problema (2.1),  $\mathbf{B} \in \mathcal{S}_{++}^n$  e  $\Lambda \in \mathcal{S}_{++}^q$ , tal que  $\Lambda$  e  $\mathcal{A}(\mathbf{x})$  comutam. Então, a seguinte matriz*

$$\mathbf{W}(\mathbf{x}, \mathbf{B}, \Lambda) = \begin{bmatrix} \mathbf{B} & \nabla \mathcal{A}(\mathbf{x}) \\ [\mathbf{I}_q \otimes \Lambda] \nabla \mathcal{A}(\mathbf{x})^T & \mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q \end{bmatrix} \quad (2.73)$$

é não singular.

*Demonstração.* Temos que provar que, se para algum  $\mathbf{v} \in \mathbb{R}^{n+q^2}$ , a solução do sistema  $\mathbf{W}(\mathbf{x}, \mathbf{B}, \Lambda)\mathbf{v} = \mathbf{0}$  admite somente a solução trivial, ou seja,  $\mathbf{v} = \mathbf{0}$ .

Seja  $\mathbf{v}^T = [\mathbf{v}_1^T, \mathbf{v}_2^T]$ , onde  $\mathbf{v}_1 \in \mathbb{R}^n$  e  $\mathbf{v}_2 \in \mathbb{R}^{q^2}$ . Assim, tem-se o seguinte sistema

$$\mathbf{B}\mathbf{v}_1 + \nabla \mathcal{A}(\mathbf{x})\mathbf{v}_2 = \mathbf{0} \quad (2.74a)$$

$$[\mathbf{I}_q \otimes \Lambda] \nabla \mathcal{A}(\mathbf{x})^T \mathbf{v}_1 + [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \mathbf{v}_2 = \mathbf{0} \quad (2.74b)$$

Da Equação (2.74a), tem-se

$$\mathbf{v}_1 = -\mathbf{B}^{-1} \nabla \mathcal{A}(\mathbf{x})\mathbf{v}_2 \quad (2.75)$$

Substituindo  $\mathbf{v}_1$  na Equação (2.74b), resulta

$$-[\mathbf{I}_q \otimes \Lambda] \nabla \mathcal{A}(\mathbf{x})^T \mathbf{B}^{-1} \nabla \mathcal{A}(\mathbf{x})\mathbf{v}_2 + [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \mathbf{v}_2 = \mathbf{0} \quad (2.76)$$

Reescrevendo a Equação (2.76)

$$\nabla \mathcal{A}(\mathbf{x})^T \mathbf{B}^{-1} \nabla \mathcal{A}(\mathbf{x})\mathbf{v}_2 - [\mathbf{I}_q \otimes \Lambda]^{-1} [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \mathbf{v}_2 = \mathbf{0} \quad (2.77)$$



Logo, fatorando o termo  $\mathbf{v}_2$  da equação anterior, obtém-se

$$\left[ \nabla \mathcal{A}(\mathbf{x})^T \mathbf{B}^{-1} \nabla \mathcal{A}(\mathbf{x}) - [\mathbf{I}_q \otimes \boldsymbol{\Lambda}]^{-1} [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \right] \mathbf{v}_2 = 0 \quad (2.78)$$

Chamando  $\mathbf{D} = \left[ \nabla \mathcal{A}(\mathbf{x})^T \mathbf{B}^{-1} \nabla \mathcal{A}(\mathbf{x}) - [\mathbf{I}_q \otimes \boldsymbol{\Lambda}]^{-1} [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \right]$  o complemento de Schur de  $\mathbf{B}$  em  $\mathbf{M}$  (Ver Anexo A), onde

$$\mathbf{M} = \begin{bmatrix} \mathbf{B} & \nabla \mathcal{A}(\mathbf{x}) \\ \nabla \mathcal{A}(\mathbf{x})^T & [\mathbf{I}_q \otimes \boldsymbol{\Lambda}]^{-1} [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \end{bmatrix} \quad (2.79)$$

Realizando o produto  $\mathbf{v}_2^T \mathbf{D} \mathbf{v}_2$ , resulta

$$\mathbf{v}_2^T \mathbf{D} \mathbf{v}_2 = \mathbf{v}_2^T \left[ \nabla \mathcal{A}(\mathbf{x})^T \mathbf{B}^{-1} \nabla \mathcal{A}(\mathbf{x}) - [\mathbf{I}_q \otimes \boldsymbol{\Lambda}]^{-1} [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \right] \mathbf{v}_2 \quad (2.80)$$

E assim

$$\mathbf{v}_2^T \mathbf{D} \mathbf{v}_2 = \mathbf{v}_2^T \nabla \mathcal{A}(\mathbf{x})^T \mathbf{B}^{-1} \nabla \mathcal{A}(\mathbf{x}) \mathbf{v}_2 - \mathbf{v}_2^T [\mathbf{I}_q \otimes \boldsymbol{\Lambda}]^{-1} [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \mathbf{v}_2 \quad (2.81)$$

Sendo  $\mathbf{B} \succ \mathbf{0}$ , então  $\mathbf{B}^{-1}$  é definida positiva, portanto

$$\mathbf{v}_2^T \nabla \mathcal{A}(\mathbf{x})^T \mathbf{B}^{-1} \nabla \mathcal{A}(\mathbf{x}) \mathbf{v}_2 \geq 0 \quad (2.82)$$

E pelo Lema 2.5, tem-se

$$\mathbf{v}_2^T [\mathbf{I}_q \otimes \boldsymbol{\Lambda}]^{-1} [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \mathbf{v}_2 \leq 0 \quad (2.83)$$

Portanto, das Equações (2.82) e (2.83) pode-se concluir que  $\mathbf{v}_2^T \mathbf{D} \mathbf{v}_2 \geq 0$ .

Assume-se que  $\mathbf{v}_2^T \mathbf{D} \mathbf{v}_2 = \mathbf{0}$ , então tem-se que provar que  $\mathbf{v}_2 = \mathbf{0}$ . Para isso, utiliza-se a Equação (2.81), então tem-se a seguinte igualdade

$$\mathbf{v}_2^T \nabla \mathcal{A}(\mathbf{x})^T \mathbf{B}^{-1} \nabla \mathcal{A}(\mathbf{x}) \mathbf{v}_2 = \mathbf{v}_2^T (\mathbf{I}_q \otimes \boldsymbol{\Lambda})^{-1} (\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q) \mathbf{v}_2 \quad (2.84)$$

Como  $\mathbf{v}_2^T \nabla \mathcal{A}(\mathbf{x})^T \mathbf{B}^{-1} \nabla \mathcal{A}(\mathbf{x}) \mathbf{v}_2 \geq 0$  e  $\mathbf{v}_2^T (\mathbf{I}_q \otimes \boldsymbol{\Lambda})^{-1} (\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q) \mathbf{v}_2 \leq 0$ , então têm-se

$$\mathbf{v}_2^T \nabla \mathcal{A}(\mathbf{x})^T \mathbf{B}^{-1} \nabla \mathcal{A}(\mathbf{x}) \mathbf{v}_2 = 0 \quad (2.85)$$

e

$$\mathbf{v}_2^T [\mathbf{I}_q \otimes \boldsymbol{\Lambda}]^{-1} [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \mathbf{v}_2 = 0 \quad (2.86)$$

Na Equação (2.85),  $\mathbf{B}^{-1}$  é definida positiva, portanto tem-se

$$\nabla \mathcal{A}(\mathbf{x}) \mathbf{v}_2 = 0 \quad (2.87)$$

Pelo Lema 2.4 e da Equação (2.86), temos

$$\mathbf{v}_2^T [\mathbf{I}_q \otimes \mathbf{\Lambda}]^{-1} [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \mathbf{v}_2 = 0 \quad (2.88)$$

Portanto

$$[\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \mathbf{v}_2 = \mathbf{0} \quad (2.89)$$

Como  $\mathbf{x}$  é um ponto regular, então na Equação (2.87) os vetores  $\nabla \mathcal{A}(\mathbf{x}) \mathbf{v}_2$  são linearmente independentes, desta maneira tem-se que  $\mathbf{v}_2 = \mathbf{0}$ .

Finalmente, substituindo  $\mathbf{v}_2$  na Equação (2.75) resulta  $\mathbf{v}_1 = \mathbf{0}$ . Portanto, a solução do sistema é  $(\mathbf{v}_1, \mathbf{v}_2) = (\mathbf{0}, \mathbf{0})$ .  $\square$

**Lema 2.7.** *Se o algoritmo para no ponto  $\mathbf{x}^k$  e  $\mathbf{d}_0^k = \mathbf{0}$ , então  $\mathbf{x}^k$  é um ponto de KKT do problema (2.1).*

*Demonstração.* Se  $\mathbf{d}_0^k = \mathbf{0}$ , o sistema (2.43) resulta

$$\nabla f(\mathbf{x}^k) + \nabla \mathcal{A}(\mathbf{x}^k) \boldsymbol{\lambda}_0^{k+1} = \mathbf{0} \quad (2.90a)$$

$$[\mathcal{A}(\mathbf{x}^k) \otimes \mathbf{I}_q] \boldsymbol{\lambda}_0^{k+1} = \mathbf{0} \quad (2.90b)$$

Pela Hipótese 2.2, como o algoritmo é viável em cada iteração, então  $\mathcal{A}(\mathbf{x}^k) \prec \mathbf{0}$ , e sendo a matriz  $[\mathcal{A}(\mathbf{x}^k) \otimes \mathbf{I}_q]$  não singular, então  $\boldsymbol{\lambda}_0^{k+1} = \mathbf{0}$  e  $\boldsymbol{\Lambda}_0^{k+1} = \mathbf{0}$ . Com isso  $\nabla f(\mathbf{x}^k) = \mathbf{0}$  e, portanto, o ponto  $\mathbf{x}^k$  é um ponto de KKT do problema (2.1).  $\square$

**Lema 2.8.** *O vetor  $\mathbf{d}_0$  definido no passo 1 do algoritmo é uma direção de descida que verifica*

$$\mathbf{d}_0^T \nabla f(\mathbf{x}) \leq -\mathbf{d}_0^T \mathbf{B} \mathbf{d}_0 \quad (2.91)$$

*Demonstração.* Multiplicando a Equação (2.43a) por  $\mathbf{d}_0$  temos

$$\mathbf{d}_0^T \nabla f(\mathbf{x}) = -\mathbf{d}_0^T \mathbf{B} \mathbf{d}_0 - \mathbf{d}_0^T \nabla \mathcal{A}(\mathbf{x}) \boldsymbol{\lambda}_0 \quad (2.92)$$

Da Equação (2.43b):

$$[\mathbf{I}_q \otimes \mathbf{\Lambda}] \nabla \mathcal{A}(\mathbf{x})^T \mathbf{d}_0 = -[\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \boldsymbol{\lambda}_0 \quad (2.93)$$

Então

$$\nabla \mathcal{A}(\mathbf{x})^T \mathbf{d}_0 = -[\mathbf{I}_q \otimes \mathbf{\Lambda}]^{-1} [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \boldsymbol{\lambda}_0 \quad (2.94)$$

Transpondo a Equação (2.94) e considerando as seguintes igualdades  $[\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q]^T = [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q]$  e  $[\mathbf{I}_q \otimes \mathbf{\Lambda}]^T = [\mathbf{I}_q \otimes \mathbf{\Lambda}]$ , obtém-se

$$\mathbf{d}_0^T \nabla \mathcal{A}(\mathbf{x}) = -\boldsymbol{\lambda}_0^T [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] [\mathbf{I}_q \otimes \mathbf{\Lambda}]^{-1} \quad (2.95)$$

Substituindo a Equação (2.95) na Equação (2.92)

$$\mathbf{d}_0^T \nabla f(\mathbf{x}) = -\mathbf{d}_0^T \mathbf{B} \mathbf{d}_0 + \boldsymbol{\lambda}_0^T [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] [\mathbf{I}_q \otimes \boldsymbol{\Lambda}]^{-1} \boldsymbol{\lambda}_0 \quad (2.96)$$

Como  $\mathbf{B} \succ \mathbf{0}$  e  $\boldsymbol{\Lambda} \succ \mathbf{0}$ , então tem-se que  $\mathbf{d}_0^T \mathbf{B} \mathbf{d}_0 > 0$  e pelo Lema 2.5 tem-se que  $[\mathbf{I}_q \otimes \boldsymbol{\Lambda}]^{-1} [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \prec \mathbf{0}$ , então  $\boldsymbol{\lambda}_0^T [\mathbf{I}_q \otimes \boldsymbol{\Lambda}]^{-1} [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \boldsymbol{\lambda}_0 \prec \mathbf{0}$ . Portanto, pode-se concluir que  $\mathbf{d}_0^T \nabla f(\mathbf{x}) \leq -\mathbf{d}_0^T \mathbf{B} \mathbf{d}_0$ .  $\square$

Como uma consequência do Lema 2.8,  $\mathbf{d}_0$  é uma direção de descida de  $f$  em  $\mathbf{x}$ , já que

$$\mathbf{d}_0^T \nabla f(\mathbf{x}) \leq -\mathbf{d}_0^T \mathbf{B} \mathbf{d}_0 < 0 \quad (2.97)$$

**Lema 2.9.** *A direção de busca  $\mathbf{d}$ , definida no passo 1 do algoritmo, é uma direção de descida que verifica*

$$\mathbf{d}^T \nabla f(\mathbf{x}) \leq \xi \mathbf{d}_0^T \nabla f(\mathbf{x}) \quad (2.98)$$

*Demonstração.* Ver [35].  $\square$

**Lema 2.10.** *Existe  $\tau > 0$  tal que, para qualquer  $\mathbf{x} \in \Omega_a$  e  $\mathbf{d}$  calculado no passo 1b do algoritmo, as condições (2.71) e (2.72) são verificadas para qualquer  $t \in [0, \tau]$ .*

*Demonstração.* Ver [35].  $\square$

**Lema 2.11.** *Existem  $\delta \geq 1$  e  $\varphi_0 > 0$  tal que a direção de busca  $\mathbf{d}$  e o parâmetro  $\rho$  calculado pelo algoritmo verifica*

$$\varphi_0 \|\mathbf{d}_0\|^2 \leq \rho \leq \varphi \|\mathbf{d}_0\|^2 \quad (2.99)$$

e

$$\|\mathbf{d}\| \leq \delta \|\mathbf{d}_0\| \quad (2.100)$$

*Demonstração.* Ver [35].  $\square$

**Lema 2.12.** *Da Hipótese (2.3) existe um número real positivo  $L$  tal que*

$$\mathcal{A}(\mathbf{y}) \preceq \mathcal{A}(\mathbf{x}) + D\mathcal{A}(\mathbf{x})(\mathbf{y} - \mathbf{x}) + L \|\mathbf{y} - \mathbf{x}\|^2 \mathbf{I}_q \quad (2.101)$$

onde  $\mathbf{x}, \mathbf{y} \in \Omega$ .

*Demonstração.* Ver [35].  $\square$

**Proposição 2.2.** *Para todo  $\mathbf{x} \in \Omega_a$  tal que  $\|\mathbf{d}(\mathbf{x})\| \geq M > 0$  então existe  $\tau > 0$  tal que*

$$\mathcal{A}(\mathbf{x} + t\mathbf{d}(\mathbf{x})) \preceq \mathbf{0} \quad (2.102)$$

para todo  $t \in [0, \tau]$ .

*Demonstração.* Ver [35]. □

**Lema 2.13.** *Qualquer ponto de acumulação  $\mathbf{x}^*$  da sequencia  $\{\mathbf{x}^k\}$  gerada pelo algoritmo é um ponto estacionário do problema (2.1).*

*Demonstração.* Considere-se a sequencia  $\{\mathbf{x}^k\}_{k \in \mathbb{K}}$ , onde  $\mathbb{K} \subset \mathbb{N} = \{1, 2, \dots\}$ . Dados  $\boldsymbol{\lambda}$ ,  $\mathbf{B}$  e  $\rho$  limitados em  $\mathbb{K}$ , então existe uma subsequencia dada pelo algoritmo  $\{\mathbf{x}^k, \boldsymbol{\lambda}^k, \mathbf{B}^k, \rho^k\}_{k \in \mathbb{K}_1}$ , sendo  $\mathbb{K}_1 \subset \mathbb{K}$ , que converge para  $\mathbf{x}^*, \boldsymbol{\lambda}^*, \mathbf{B}^*, \rho^*$ .

Considerando que  $\mathbf{d}$  depende continuamente de  $\mathbf{x}$ ,  $\boldsymbol{\lambda}$ ,  $\mathbf{B}$  e  $\rho$ , então

$$\{\mathbf{d}^k\}_{k \in \mathbb{K}_1} \rightarrow \mathbf{d}^* \quad (2.103)$$

Agora, considerando a seguinte subsequencia  $\{t^k\}_{k \in \mathbb{K}_2}$ , sendo  $\mathbb{K}_2 \subset \mathbb{K}_1$ , que converge para  $t^*$ . Do Lema 2.10 mostra-se que  $t^* > 0$ .

A Equação (2.71) pode ser reescrita como

$$f(\mathbf{x}^k + t\mathbf{d}) \leq f(\mathbf{x}^k) + \eta t^k \mathbf{d}^{kT} \nabla f(\mathbf{x}) \quad (2.104)$$

onde assume-se que  $k \in \mathbb{K}_2$  e  $k + 1$  não necessariamente pertence a  $\mathbb{K}_2$ .

Tomando limites em ambos os lados da expressão anterior quando  $k \rightarrow \infty$

$$f(\mathbf{x}^*) \leq f(\mathbf{x}^*) + \eta t^* \mathbf{d}^{*T} \nabla f(\mathbf{x}^*) \quad (2.105)$$

e assim tem-se

$$\eta t^* \mathbf{d}^{*T} \nabla f(\mathbf{x}^*) \geq 0 \quad (2.106)$$

como  $\eta \in (0, 1)$  e  $t^* > 0$ , então

$$\mathbf{d}^{*T} \nabla f(\mathbf{x}^*) \geq 0 \quad (2.107)$$

Empregando o Lema 2.9 e tomando limites para ambos os lados da Equação (2.98) quando  $k \rightarrow \infty$ , tem-se

$$\mathbf{d}^{*T} \nabla f(\mathbf{x}^*) \leq \xi \mathbf{d}_0^{*T} \nabla f(\mathbf{x}^*) \quad (2.108)$$

como  $\xi \in (0, 1)$ , então

$$\mathbf{d}^{*T} \nabla f(\mathbf{x}^*) \leq 0 \quad (2.109)$$

Das Equações (2.107) e (2.109) pode-se concluir que

$$\mathbf{d}^{*T} \nabla f(\mathbf{x}^*) = 0 \quad (2.110)$$

E, portanto

$$\mathbf{d}^* = \mathbf{0} \quad (2.111)$$

Empregando novamente o Lema 2.9, obtém-se

$$\mathbf{d}_0^{*T} \nabla f(\mathbf{x}^*) \geq 0 \quad (2.112)$$

Do Lema 2.8 e tomando limites para ambos os lados da Equação (2.91) quando  $k \rightarrow \infty$ , tem-se

$$\mathbf{d}_0^{*T} \nabla f(\mathbf{x}^*) \leq -\mathbf{d}_0^{*T} \mathbf{B}^* \mathbf{d}_0^* \quad (2.113)$$

Substituindo a Equação (2.110) na Equação (2.113), tem-se

$$\mathbf{d}_0^{*T} \mathbf{B}^* \mathbf{d}_0^* \leq 0 \quad (2.114)$$

Como a matriz  $\mathbf{B}^*$  é definida positiva, então

$$\mathbf{d}_0^{*T} \mathbf{B}^* \mathbf{d}_0^* \geq 0 \quad (2.115)$$

De (2.114) e (2.115), tem-se que

$$\mathbf{d}_0^{*T} \mathbf{B}^* \mathbf{d}_0^* = 0 \quad (2.116)$$

E, portanto

$$\mathbf{d}_0^* = \mathbf{0} \quad (2.117)$$

Tomando o limite quando  $k \rightarrow \infty$  no sistema (2.43), obtém-se o seguinte sistema linear

$$\mathbf{B}^* \mathbf{d}_0^* + \nabla \mathcal{A}(\mathbf{x}^*) \boldsymbol{\lambda}_0^* = -\nabla f(\mathbf{x}^*) \quad (2.118a)$$

$$[\mathbf{I}_q \otimes \boldsymbol{\Lambda}^*] \nabla \mathcal{A}(\mathbf{x}^*)^T \mathbf{d}_0^* + [\mathcal{A}(\mathbf{x}^*) \otimes \mathbf{I}_q] \boldsymbol{\lambda}_0^* = 0 \quad (2.118b)$$

Como  $\mathbf{d}_0^* = \mathbf{0}$ , então o sistema (2.118) resulta

$$\nabla f(\mathbf{x}^*) + \nabla \mathcal{A}(\mathbf{x}^*) \boldsymbol{\lambda}_0^* = \mathbf{0} \quad (2.119a)$$

$$[\mathcal{A}(\mathbf{x}^*) \otimes \mathbf{I}_q] \boldsymbol{\lambda}_0^* = 0 \quad (2.119b)$$

Sendo  $\boldsymbol{\Lambda}_0^* = \text{mat}(\boldsymbol{\lambda}_0^*)$ , então a Equação (2.119b) é equivalente a

$$\boldsymbol{\Lambda}_0^* \mathcal{A}(\mathbf{x}^*) = \mathbf{0} \quad (2.120)$$

Portanto, pode-se concluir que o ponto  $\mathbf{x}^*$  é um ponto estacionário do problema (2.1).  $\square$

## 2.8 Detalhes da implementação do algoritmo

### 2.8.1 Resolução dos sistemas lineares internos do algoritmo FDIPA-SDP-NS

A seguir serão mostrados os custos computacionais para a solução dos sistemas lineares internos do algoritmo FDIPA-SDP-NS tendo em consideração a estrutura da matriz  $\mathcal{A}$ .

No caso particular da otimização estrutural as matrizes de rigidez e massa tem uma estrutura particular que é abordada de forma muito eficiente por códigos de elementos finitos.

Da equação (2.43a) temos

$$\boldsymbol{\lambda}_0 = -[\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q]^{-1} [\mathbf{I}_q \otimes \boldsymbol{\Lambda}] \nabla \mathcal{A}(\mathbf{x})^T \mathbf{d}_0 \quad (2.121)$$

Substituindo a Equação (2.121) na Equação 2.43b) temos

$$[\mathbf{B} - \nabla \mathcal{A}(\mathbf{x}) [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q]^{-1} [\mathbf{I}_q \otimes \boldsymbol{\Lambda}] \nabla \mathcal{A}(\mathbf{x})^T] \mathbf{d}_0 = -\nabla f(\mathbf{x}) \quad (2.122)$$

Seja

$$\mathbf{U} = [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q]^{-1} [\mathbf{I}_q \otimes \boldsymbol{\Lambda}] \nabla \mathcal{A}(\mathbf{x})^T \quad (2.123)$$

onde  $\mathbf{U} \in \mathbb{R}^{q^2 \times n}$ .

Então, obtém-se o seguinte sistema linear simétrico

$$[\mathbf{B} - \nabla \mathcal{A}(\mathbf{x}) \mathbf{U}] \mathbf{d}_0 = -\nabla f(\mathbf{x}) \quad (2.124)$$

Pelo Lema 2.1 e aplicando as propriedades do produto de Kronecker, o termo  $[\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q]^{-1}$  da Equação (2.123) é equivalente a

$$[\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q]^{-1} = \mathcal{A}(\mathbf{x})^{-1} \otimes \mathbf{I}_q^{-1} \quad (2.125)$$

Substituindo a Equação (2.125) na Equação (2.123) resulta

$$\mathbf{U} = [\mathcal{A}(\mathbf{x})^{-1} \otimes \mathbf{I}_q^{-1}] [\mathbf{I}_q \otimes \boldsymbol{\Lambda}] \nabla \mathcal{A}(\mathbf{x})^T \quad (2.126)$$

Aplicando novamente as propriedades do produto de Kronecker do Lema 2.1 a Equação (2.126) é equivalente a

$$\mathbf{U} = [\mathcal{A}(\mathbf{x})^{-1} \mathbf{I}_q] \otimes [\mathbf{I}_q^{-1} \boldsymbol{\Lambda}] \nabla \mathcal{A}(\mathbf{x})^T \quad (2.127)$$

Portanto

$$\mathbf{U} = [\mathcal{A}(\mathbf{x})^{-1} \otimes \mathbf{\Lambda}] \nabla \mathcal{A}(\mathbf{x})^T \quad (2.128)$$

Seja  $U_k$  a  $k$ -ésima coluna de  $\mathbf{U}$ . Então, empregando a Equação (2.128) e aplicando as definições do mapeamento  $vec$  e  $\nabla \mathcal{A}(\mathbf{x})$  tem-se

$$U_k = vec \left( \mathbf{\Lambda} \frac{\partial \mathcal{A}}{\partial x_k}(\mathbf{x}) \mathcal{A}(\mathbf{x})^{-1} \right) \quad (2.129)$$

Agora, aplicando a definição do mapeamento  $mat$  a ambos os lados da Equação (2.129), então  $U_k$  pode ser obtida resolvendo os seguintes  $q$  sistemas lineares

$$\mathcal{A}(\mathbf{x}) mat(U_k) = \frac{\partial \mathcal{A}}{\partial x_k}(\mathbf{x}) \mathbf{\Lambda} \quad (2.130)$$

Desta maneira, para calcular  $\mathbf{U}$  é necessário resolver  $q \times n$  sistemas lineares com a mesma matriz simétrica definida negativa  $\mathcal{A}(\mathbf{x})$ .

A fatoração para resolver o sistema simétrico (2.124) requer  $\frac{1}{6}n^3$  operações de pontos flutuantes (flops).

No caso quando  $\mathcal{A}(\mathbf{x})$  é uma matriz banda com comprimento de banda  $b$ , a fatoração de Cholesky necessita cerca de  $\frac{1}{2}b(b+3)q$  flops (Ver [39]). Note que em aplicações de otimização estrutural em geral tem-se que  $b \ll q$ . Uma vez que a substituição para frente e para trás para cada sistema requer cerca de  $(2b+1)q$  flops, para calcular  $\mathbf{U}$  necessita-se cerca de  $q^2n(2b+1) + \frac{1}{2}b(b+3)q$  flops. O cálculo de  $\boldsymbol{\lambda}_0$  na Equação (2.121) requer cerca de  $2nq^2$  flops.

O sistema resolvido no algoritmo FDIPA-SDP não é simétrico e de dimensão  $\bar{l} = n + \frac{1}{2}q(q+1)$ . Neste algoritmo a matriz  $\mathbf{\Lambda}$  não é esparsa e, portanto, em geral, o sistema linear resultante não é esparso e a fatoração requereria cerca de  $\frac{1}{3}(\bar{l}^3 - \bar{l})$  operações [39].

Desta maneira, o algoritmo FDIPA-SDP não toma vantagem da estrutura da matriz  $\mathcal{A}(\mathbf{x})$  e, portanto, o número de operações requeridas em cada passo do algoritmo é muito maior do que do algoritmo FDIPA-SDP-NS.

## 2.8.2 Critérios de parada

Como  $\mathbf{d}(\mathbf{x}) = \mathbf{0}$  e  $\nabla L(\mathbf{x}, \mathbf{\Lambda}_0) = \mathbf{0}$  em um ponto de KKT então pode-se empregar as seguintes condições:  $\mathbf{d}_0(\mathbf{x}^k) \leq \delta$  e/ou  $\nabla L(\mathbf{x}^k, \mathbf{\Lambda}_0^k) \leq \delta$  como critérios de parada do algoritmo. Aqui  $\delta$  representa um indicador da proximidade da solução o qual tem um valor muito pequeno. Também podem-se dar valores de tolerância para as condições de complementariedade.

### 2.8.3 Busca de um ponto inicial viável

Quando a restrição matricial não é verificada no ponto inicial, um ponto inicial viável tem que ser encontrado empregando o seguinte programa matemático auxiliar

$$\begin{cases} \underset{\mathbf{x}, z}{\text{minimizar}} & z \\ \text{sujeito a:} & \mathcal{A}(\mathbf{x}) - z\mathbf{I}_q \preceq 0 \end{cases}, \quad (2.131)$$

onde  $z$  é uma variável escalar auxiliar.

Depois de realizar certo número de iterações para resolver este problema auxiliar obtém-se um ponto inicial viável uma vez que  $z$  seja negativo.



# Capítulo 3

## Algoritmo de pontos interiores e direções viáveis para programação semidefinida generalizada (FDIPA-GSDP)

### 3.1 Introdução

Neste capítulo apresenta-se um novo algoritmo de pontos interiores e direções viáveis para programação semidefinida generalizada não linear denominado FDIPA-GSDP para lidar com problemas de otimização não linear com restrições vetoriais de desigualdade e igualdade e também com restrições matriciais semidefinidas.

Um problema de programação semidefinida generalizada [17, 18] pode ser formulado da seguinte maneira

$$\left\{ \begin{array}{l} \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimizar}} \quad f(\mathbf{x}) \\ \text{sujeito a:} \quad \mathbf{g}(\mathbf{x}) \leq 0; \quad \mathbf{g}(\mathbf{x}) \in \mathbb{R}^m \\ \quad \quad \quad \mathbf{h}(\mathbf{x}) = 0; \quad \mathbf{h}(\mathbf{x}) \in \mathbb{R}^p \\ \quad \quad \quad \mathcal{A}(\mathbf{x}) \preceq 0; \quad \mathcal{A}(\mathbf{x}) \in \mathcal{S}^q \end{array} \right. \quad (3.1)$$

sendo  $\mathbf{x} \in \mathbb{R}^n$  o vetor das variáveis de projeto,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  a função objetivo,  $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  a função das restrições de desigualdade e  $\mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^p$  a função das restrições de igualdade. Assume-se que as funções  $f$ ,  $\mathbf{g}$ ,  $\mathbf{h}$  e  $\mathcal{A}$  são suaves e pelo menos uma delas é não linear.

Este novo algoritmo, proposto neste trabalho, é uma generalização do problema de programação semidefinida não linear que foi discutido e analisado no capítulo anterior. As técnicas empregadas neste algoritmo são baseadas nas técnicas clássicas

de programação não linear assim como nas técnicas de programação semidefinida não linear desenvolvidas nos algoritmos desenvolvidas nos algoritmos FDIPA [1, 4, 5] e FDIPA-SDP-NS [24, 40], respectivamente. Desta maneira, o algoritmo FDIPA-GSDP é capaz de resolver problemas de programação semidefinida não linear (tirando as restrições de desigualdade e igualdade), problemas de programação não linear (tirando as restrições semidefinidas negativas) e problemas de programação semidefinida generalizada não linear (considerando todas as restrições).

A análise teórica e a prova de convergência do algoritmo FDIPA-GSDP foi realizada em duas etapas. Na primeira etapa, analisa-se um problema de otimização considerando restrições de desigualdade e na segunda etapa são incluídas as restrições de igualdade.

## 3.2 Problema de otimização com restrições de desigualdade

Neste problema são consideradas as restrições de desigualdade e as restrições matriciais semidefinidas. Este problema pode ser formulado como

$$\begin{cases} \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimizar}} & f(\mathbf{x}) \\ \text{sujeito a:} & \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \quad \mathbf{g}(\mathbf{x}) \in \mathbb{R}^m \\ & \mathcal{A}(\mathbf{x}) \preceq \mathbf{0}, \quad \mathcal{A}(\mathbf{x}) \in \mathcal{S}^q \end{cases} \quad (3.2)$$

### 3.2.1 Notações

- O conjunto de pontos viáveis do problema (3.2) em relação às restrições de desigualdade e restrições matriciais semidefinidas negativas é denotado por  $\Omega = \{\mathbf{x} \in \mathbb{R}^n : \mathcal{A}(\mathbf{x}) \preceq \mathbf{0}, \mathbf{g}(\mathbf{x}) \leq \mathbf{0}\}$ . Seu interior é denotado por  $\text{int}(\Omega)$ .
- O vetor gradiente da função objetivo  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é denotado por

$$\nabla f(\mathbf{x}) = \left[ \frac{\partial f(\mathbf{x})}{\partial x_1}, \dots, \frac{\partial f(\mathbf{x})}{\partial x_n} \right]^T \in \mathbb{R}^n \quad (3.3)$$

- As derivadas parciais das funções  $g_i$  em relação a  $x_i$  são agrupadas na matriz  $\nabla \mathbf{g}(\mathbf{x}) \in \mathbb{R}^{n \times m}$ , denotada por

$$\nabla \mathbf{g}(\mathbf{x}) = \left[ \nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x}) \right] \quad (3.4)$$

- O gradiente da restrição matricial semidefinida  $\nabla \mathcal{A}(\mathbf{x})$  foi definida no capítulo anterior.

- A função Lagrangiana  $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{q \times q} \rightarrow \mathbb{R}$  associada ao problema (3.2) é

$$L(\mathbf{x}, \boldsymbol{\gamma}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \text{vec}(\mathbf{A}(\mathbf{x}))^T \boldsymbol{\lambda} \quad (3.5)$$

- O gradiente da função Lagrangiana em relação a  $\mathbf{x}$  é denotado por  $\nabla_{\mathbf{x}}L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{q^2} \rightarrow \mathbb{R}^n$  e definido como

$$\nabla_{\mathbf{x}}L(\mathbf{x}, \boldsymbol{\gamma}, \boldsymbol{\lambda}) = \nabla f(\mathbf{x}) + \sum_{i=1}^m \lambda_i \nabla g_i(\mathbf{x}) + \nabla \mathbf{A}(\mathbf{x}) \boldsymbol{\lambda} \quad (3.6)$$

- A Hessiana da função Lagrangiana, associada ao problema (3.2), é denotada por  $\mathbf{H} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{q^2} \rightarrow \mathbb{R}^{n \times n}$  e definida como

$$\mathbf{H}(\mathbf{x}, \boldsymbol{\gamma}, \boldsymbol{\lambda}) = \nabla^2 f(\mathbf{x}) + \sum_{i=1}^m \lambda_i \nabla^2 g_i(\mathbf{x}) + \nabla^2 \mathbf{A}(\mathbf{x}) \boldsymbol{\lambda} \quad (3.7)$$

### 3.2.2 Definições

Este algoritmo também emprega as mesmas definições do algoritmo FDIPA descritas na Seção 1.1.2 do Capítulo 1.

A seguinte proposição [1] estabelece uma condição para que uma direção  $\mathbf{d}$  verifique a Definição 1.4.

**Proposição 3.1.** *Se a função  $\mathbf{g}$  é diferenciável em  $\mathbf{x} \in \Omega$ , então  $\mathbf{d} \in \mathbb{R}^n$  é uma direção viável em relação às restrições de desigualdade, se  $\nabla g_i(\mathbf{x}^k) \mathbf{d} < 0$ , para todo  $i$  tal que  $g_i(\mathbf{x}) = 0$ .*

*Demonstração.* Ver [5]. □

### 3.2.3 Condições de otimalidade de primeira ordem

As condições de otimalidade de primeira ordem de Karush-Kuhn-Tucker para este problema são

$$\nabla f(\mathbf{x}) + \nabla \mathbf{g}(\mathbf{x}) \boldsymbol{\gamma} + \nabla \mathbf{A}(\mathbf{x}) \boldsymbol{\lambda} = 0 \quad (3.8a)$$

$$\mathcal{G}(\mathbf{x}) \boldsymbol{\gamma} = 0 \quad (3.8b)$$

$$\boldsymbol{\Lambda} \mathbf{A}(\mathbf{x}) = 0 \quad (3.8c)$$

$$\mathbf{g}(\mathbf{x}) \leq 0 \quad (3.8d)$$

$$\mathbf{A}(\mathbf{x}) \preceq 0 \quad (3.8e)$$

$$\boldsymbol{\gamma} \geq 0 \quad (3.8f)$$

$$\boldsymbol{\Lambda} \succeq 0 \quad (3.8g)$$

onde  $\boldsymbol{\gamma} \in \mathbb{R}_+^m$  é o vetor dos multiplicadores de Lagrange das restrições de desigualdade. Os demais termos destas equações foram descritos no capítulo anterior.

A Equação (3.8a) é denominada a condição lagrangiana e as Equações (3.8b) e (3.8c) são denominadas as condições de complementariedade das restrições de desigualdade e semidefinida, respectivamente. Neste caso, as Equações (3.8d)-(3.8g) representam as condições de viabilidade.

As definições referentes as condições de otimalidade de primeira ordem de Karush-Kuhn-Tucker são descritas a continuação.

**Definição 3.1 (Ponto regular).** Um ponto  $\mathbf{x}$  é um ponto regular do problema (3.2) se os vetores  $\nabla g_i(\mathbf{x})$ ,  $i \in I(\mathbf{x})$  e  $\nabla \mathcal{A}(\mathbf{x})\mathbf{b}_i(\mathbf{x}) : i = 1, \dots, r$  e, são linearmente independentes. Sendo  $I(\mathbf{x})$  o conjunto dos índices das restrições ativas (Ver Definição 1.2).

**Definição 3.2 (Ponto estacionário).** Um ponto regular  $\mathbf{x}$  é um ponto estacionário do problema (3.2) se existem um vetor  $\boldsymbol{\gamma} \in \mathbb{R}^m$  e uma matriz  $\boldsymbol{\Lambda} \in \mathbb{R}^{q \times q}$  tal que as seguintes condições são verificadas

$$\begin{aligned}\nabla f(\mathbf{x}) + \nabla \mathbf{g}(\mathbf{x})\boldsymbol{\gamma} + \nabla \mathcal{A}(\mathbf{x})\boldsymbol{\lambda} &= 0 \\ \mathcal{G}(\mathbf{x})\boldsymbol{\gamma} &= 0 \\ \boldsymbol{\Lambda}\mathcal{A}(\mathbf{x}) &= 0 \\ \mathcal{A}(\mathbf{x}) &\preceq 0\end{aligned}$$

**Definição 3.3 (Ponto de Karush-Kuhn-Tucker).** Um ponto estacionário  $\mathbf{x}$  é um ponto de Karush-Kuhn-Tucker do problema (3.2) se todas as Equações (3.8a)-(3.8g) são verificadas simultaneamente.

### 3.2.4 Hipóteses sobre o problema

As hipóteses acerca das funções  $f$ ,  $\mathbf{g}$  e  $\mathcal{A}$  que foram consideradas no problema (3.2) são as seguintes:

**Hipótese 3.1.** *Existe um número real  $a$  tal que o conjunto  $\Omega_a = \{\mathbf{x} \in \Omega : f(\mathbf{x}) \leq a\}$  é compacto e tem um interior não vazio denotado por  $\text{int}(\Omega_a)$ .*

**Hipótese 3.2.** *Para qualquer ponto  $\mathbf{x} \in \text{int}(\Omega_a)$ , verificam-se que  $\mathbf{g}(\mathbf{x}) < \mathbf{0}$  e  $\mathcal{A}(\mathbf{x}) \prec \mathbf{0}$ .*

**Hipótese 3.3.** *As funções  $f$ ,  $\mathbf{g}$  e  $\mathcal{A}$  são continuamente diferenciáveis em  $\Omega_a$  e suas derivadas  $\nabla f$ ,  $\nabla \mathbf{g}$  e  $\frac{\partial \mathcal{A}}{\partial x_p}$ , para  $p = 1, \dots, n$ , satisfazem a condição de Lipschitz.*

**Hipótese 3.4.** *Qualquer ponto  $\mathbf{x} \in \Omega_a$  é um ponto regular do problema (3.2).*

### 3.2.5 Descrição do algoritmo FDIPA-GSDP

O algoritmo FDIPA-GSDP, proposto no presente trabalho, é um método de pontos interiores que resolve um problema generalizado de otimização não linear semidefinida fazendo iterações nas variáveis de projeto  $\mathbf{x}$  (variáveis primais) e nos multiplicadores de Lagrange (variáveis duais) para resolver as condições de otimalidade de Karush-Kuhn-Tucker (KKT).

Dado um ponto interior inicial, o algoritmo gera uma sequência  $\{\mathbf{x}^k\}$  de pontos interiores, tal que  $f(\mathbf{x}^{k+1}) < f(\mathbf{x}^k)$ ,  $\mathbf{g}(\mathbf{x}^k) < 0$  e  $\mathcal{A}(\mathbf{x}^k) \prec 0$ , que converge para o conjunto de pontos de KKT do problema. Em cada iteração uma direção de busca  $\mathbf{d}^k$  é definida de maneira que seja de descida da função objetivo e viável de  $\Omega$ . Uma busca linear nessa direção determina um novo ponto interior garantindo uma convergência global.

- **Matriz do sistema linear**

Como já foi descrito no capítulo anterior, a matriz do sistema linear é formada empregando a condição lagrangiana e as condições de complementariedade das restrições fornecidas pelas Equações (3.8a)-(3.8c). Neste caso, o vetor  $\mathbf{y} = (\mathbf{x}, \boldsymbol{\gamma}, \boldsymbol{\lambda})^T$  e a função vetorial  $\boldsymbol{\Psi} : \mathbb{R}^{n+m+q^2} \rightarrow \mathbb{R}^{n+m+q^2}$  é definida como

$$\boldsymbol{\Psi}(\mathbf{x}, \boldsymbol{\gamma}, \boldsymbol{\lambda}) = \begin{bmatrix} \boldsymbol{\Psi}_l(\mathbf{x}, \boldsymbol{\gamma}, \boldsymbol{\lambda}) \\ \boldsymbol{\Psi}_g(\mathbf{x}, \boldsymbol{\gamma}, \boldsymbol{\lambda}) \\ \boldsymbol{\Psi}_A(\mathbf{x}, \boldsymbol{\gamma}, \boldsymbol{\lambda}) \end{bmatrix} \quad (3.9)$$

sendo

$$\boldsymbol{\Psi}_l(\mathbf{x}, \boldsymbol{\gamma}, \boldsymbol{\lambda}) = \nabla f(\mathbf{x}) + \nabla \mathbf{g}(\mathbf{x})\boldsymbol{\gamma} + \nabla \mathcal{A}(\mathbf{x})\boldsymbol{\lambda} \quad (3.10)$$

$$\boldsymbol{\Psi}_g(\mathbf{x}, \boldsymbol{\gamma}, \boldsymbol{\lambda}) = \mathcal{G}(\mathbf{x})\boldsymbol{\gamma} \quad (3.11)$$

$$\boldsymbol{\Psi}_A(\mathbf{x}, \boldsymbol{\gamma}, \boldsymbol{\lambda}) = \boldsymbol{\Lambda}\mathcal{A}(\mathbf{x}) \quad (3.12)$$

Empregando o procedimento do capítulo anterior, obtém-se a matriz do sistema linear  $\mathbf{W} \in \mathbb{R}^{(n+m+q^2) \times (n+m+q^2)}$  definida como

$$\mathbf{W}(\mathbf{x}, \mathbf{B}, \boldsymbol{\gamma}, \boldsymbol{\Lambda}) = \begin{bmatrix} \mathbf{B} & \nabla \mathbf{g}(\mathbf{x}) & \nabla \mathcal{A}(\mathbf{x}) \\ \boldsymbol{\Gamma} \nabla \mathbf{g}(\mathbf{x})^T & \mathcal{G}(\mathbf{x}) & \mathbf{0} \\ [\mathbf{I}_q \otimes \boldsymbol{\Lambda}] \nabla \mathcal{A}(\mathbf{x})^T & \mathbf{0} & \mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q \end{bmatrix}, \quad (3.13)$$

sendo esta também inversível.

- **Cálculo da direção de busca**

Em cada iteração do algoritmo é calculada uma direção de busca  $\mathbf{d}^k$  a qual

pretende-se que seja uma direção de descida para a função  $f$  no ponto  $\mathbf{x}^k$  e viável para o problema (3.2).

Logo de empregar uma iteração de Newton como no algoritmo anterior e definindo a direção  $\mathbf{d}_0^k$  no espaço primal como  $\mathbf{d}_0^k = \mathbf{x}^{k+1} - \mathbf{x}^k$ , obtém-se o seguinte sistema linear

$$\mathbf{B}^k \mathbf{d}_0^k + \nabla \mathbf{g}(\mathbf{x}^k) \boldsymbol{\gamma}_0^{k+1} + \nabla \mathcal{A}(\mathbf{x}^k) \boldsymbol{\lambda}_0^{k+1} = -\nabla f(\mathbf{x}^k) \quad (3.14a)$$

$$\boldsymbol{\Gamma}^k \nabla \mathbf{g}(\mathbf{x}^k) \mathbf{d}_0^k + \mathcal{G}(\mathbf{x}^k) \boldsymbol{\gamma}_0^{k+1} = \mathbf{0} \quad (3.14b)$$

$$\left[ \mathbf{I}_q \otimes \boldsymbol{\Lambda}^k \right] \nabla \mathcal{A}(\mathbf{x}^k)^T \mathbf{d}_0^k + \left[ \mathcal{A}(\mathbf{x}^k) \otimes \mathbf{I}_q \right] \boldsymbol{\lambda}_0^{k+1} = \mathbf{0} \quad (3.14c)$$

A solução do sistema (3.14) fornece uma direção  $\mathbf{d}_0^k$  no espaço primal e uma nova estimativa dos multiplicadores de Lagrange  $\boldsymbol{\gamma}_0^{k+1}$  e  $\boldsymbol{\lambda}_0^{k+1}$ .

Sem embargo  $\mathbf{d}_0^k$  pode não ser uma direção viável se um ponto está na fronteira de  $\Omega$ , como já foi mostrado no capítulo anterior para o caso quando alguns dos autovalores da matriz semidefinida se aproximam de zero. Isto também poderia acontecer se alguma restrição de desigualdade torna-se próxima de zero, ou seja, quando a restrição tende a ser ativa e desta maneira  $\mathbf{d}_0^k$  torna-se tangente à região viável.

Escrevendo a Equação (3.14b) para a  $i$ -ésima restrição de desigualdade, tem-se a seguinte expressão

$$\gamma_i^k \nabla g_i(\mathbf{x}^k)^T \mathbf{d}_0^k + g_i(\mathbf{x}^k) \gamma_{0i}^{k+1} = 0, \quad i = 1, \dots, m, \quad (3.15)$$

se  $g_i(\mathbf{x}^k) = 0$ , para  $i \in I(\mathbf{x})$ , então, a Equação (3.15) resulta

$$\gamma_i^k \nabla g_i(\mathbf{x}^k)^T \mathbf{d}_0^k = 0, \quad (3.16)$$

além disso,  $\gamma_i \neq 0$ , o que implica que  $\nabla g_i(\mathbf{x}^k)^T \mathbf{d}_0^k = 0$ , e portanto, os vetores  $\nabla g_i(\mathbf{x}^k)$  e  $\mathbf{d}_0^k$  são ortogonais.

Uma interpretação geométrica desta análise pode ser observado na Figura 3.1, na qual somente foi considerada uma restrição de desigualdade. Neste caso, os vetores  $\nabla g_1(\mathbf{x}^k)$  e  $\mathbf{d}_0^k$  são ortogonais produzindo que  $\mathbf{d}_0^k$  seja tangente ao conjunto viável e portanto, tornando-se não viável.

- **Direção viável**

Para obter uma direção viável adiciona-se um termo negativo no lado direito das Equações (3.14b) e (3.14c) do primeiro sistema linear conservando a mesma matriz  $\mathbf{W}(\mathbf{x}, \mathbf{B}, \boldsymbol{\Lambda})$  deste sistema. Desta maneira, define-se o seguinte sistema

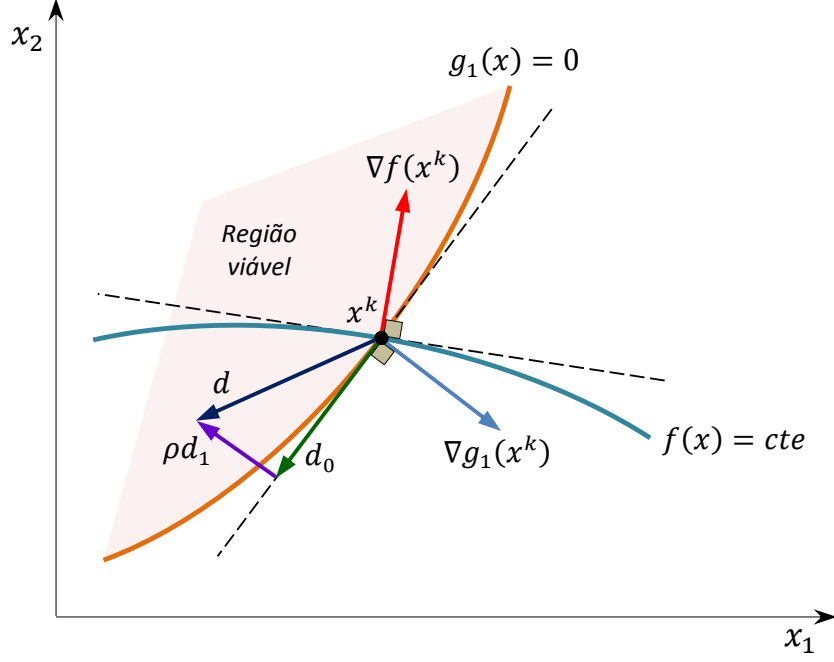


Figura 3.1: Direção de busca não viável no algoritmo FDIPA-GSDP.

linear

$$\mathbf{B}^k \mathbf{d}^k + \nabla \mathbf{g}(\mathbf{x}^k) \bar{\gamma}^{k+1} + \nabla \mathcal{A}(\mathbf{x}^k) \bar{\boldsymbol{\lambda}}^{k+1} = -\nabla f(\mathbf{x}^k) \quad (3.17a)$$

$$\mathbf{\Gamma}^k \nabla \mathbf{g}(\mathbf{x}^k) \mathbf{d}^k + \mathcal{G}(\mathbf{x}^k) \bar{\gamma}^{k+1} = -\rho^k \gamma^k \quad (3.17b)$$

$$[\mathbf{I}_q \otimes \boldsymbol{\Lambda}^k] \nabla \mathcal{A}(\mathbf{x}^k)^T \mathbf{d}^k + [\mathcal{A}(\mathbf{x}^k) \otimes \mathbf{I}_q] \bar{\boldsymbol{\lambda}}^{k+1} = -\rho^k \boldsymbol{\lambda}^k \quad (3.17c)$$

onde  $\rho^k$  é um número real positivo. Este novo sistema linear fornece a nova direção de busca  $\mathbf{d}^k$  e as novas estimativas dos multiplicadores de Lagrange  $\bar{\gamma}^{k+1}$  e  $\bar{\boldsymbol{\lambda}}^{k+1}$ .

Neste caso, o sistema auxiliar está dado por

$$\mathbf{B}^k \mathbf{d}_1^k + \nabla \mathbf{g}(\mathbf{x}^k) \gamma_1^{k+1} + \nabla \mathcal{A}(\mathbf{x}^k) \boldsymbol{\lambda}_1^{k+1} = 0 \quad (3.18a)$$

$$\mathbf{\Gamma}^k \nabla \mathbf{g}(\mathbf{x}^k) \mathbf{d}_1^k + \mathcal{G}(\mathbf{x}^k) \gamma_1^{k+1} = -\gamma^k \quad (3.18b)$$

$$[\mathbf{I}_q \otimes \boldsymbol{\Lambda}^k] \nabla \mathcal{A}(\mathbf{x}^k)^T \mathbf{d}_1^k + [\mathcal{A}(\mathbf{x}^k) \otimes \mathbf{I}_q] \boldsymbol{\lambda}_1^{k+1} = -\boldsymbol{\lambda}^k \quad (3.18c)$$

que fornece a direção  $\mathbf{d}_1^k$  e os multiplicadores de Lagrange  $\gamma_1^{k+1}$  e  $\boldsymbol{\lambda}_1^{k+1}$  e toma as seguintes seguintes relações

$$\mathbf{d}^k = \mathbf{d}_0^k + \rho^k \mathbf{d}_1^k \quad (3.19)$$

$$\bar{\gamma} = \gamma_0 + \rho \gamma_1 \quad (3.20)$$

$$\bar{\boldsymbol{\lambda}}^{k+1} = \boldsymbol{\lambda}_0^{k+1} + \rho^k \boldsymbol{\lambda}_1^{k+1} \quad (3.21)$$

As relações mostradas nas Equações (3.19)-(3.21) foram obtidas manipulando as equações dos sistemas lineares (3.14) e (3.18). Para isso, resolve-se o sistema (3.14) retirando os valores de  $\boldsymbol{\gamma}_0$  e  $\boldsymbol{\lambda}_0$  das Equações (3.14b) e (3.14c), respectivamente, assim, obtêm-se as seguintes expressões

$$\boldsymbol{\Gamma}_0 = -\mathcal{G}(\mathbf{x})^{-1} \boldsymbol{\Gamma} \nabla \mathbf{g}(\mathbf{x})^T \mathbf{d}_0 \quad (3.22)$$

$$\boldsymbol{\lambda}_0 = -[\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q]^{-1} [\mathbf{I}_q \otimes \boldsymbol{\Lambda}] \nabla \mathcal{A}(\mathbf{x})^T \mathbf{d}_0 \quad (3.23)$$

Logo, substituindo as Equações (3.22) e (3.23) na Equação (3.14a), tem-se que

$$\begin{aligned} & \left[ \mathbf{B} - \nabla \mathbf{g}(\mathbf{x}) \mathcal{G}(\mathbf{x})^{-1} \boldsymbol{\Gamma} \nabla \mathbf{g}(\mathbf{x})^T - \right. \\ & \left. \nabla \mathcal{A}(\mathbf{x}) [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q]^{-1} [\mathbf{I}_q \otimes \boldsymbol{\Lambda}] \nabla \mathcal{A}(\mathbf{x})^T \right] \mathbf{d}_0 = -\nabla f(\mathbf{x}) \end{aligned} \quad (3.24)$$

Define-se a matriz  $\mathbf{M}$  mediante a seguinte expressão

$$\mathbf{M} = \mathbf{B} - \nabla \mathbf{g}(\mathbf{x}) \mathcal{G}(\mathbf{x})^{-1} \boldsymbol{\Gamma} \nabla \mathbf{g}(\mathbf{x})^T - \nabla \mathcal{A}(\mathbf{x}) [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q]^{-1} [\mathbf{I}_q \otimes \boldsymbol{\Lambda}] \nabla \mathcal{A}(\mathbf{x})^T \quad (3.25)$$

Então, obtêm-se a seguinte relação

$$\mathbf{d}_0 \mathbf{M} = -\nabla f(\mathbf{x}) \quad (3.26)$$

Empregando o mesmo procedimento para o sistema (3.17) obtêm-se as seguintes expressões para  $\bar{\boldsymbol{\gamma}}$  e  $\bar{\boldsymbol{\lambda}}$

$$\bar{\boldsymbol{\gamma}} = -\mathcal{G}(\mathbf{x})^{-1} \left[ \boldsymbol{\Gamma} \nabla \mathbf{g}(\mathbf{x})^T \mathbf{d} + \rho \boldsymbol{\gamma} \right] \quad (3.27)$$

$$\bar{\boldsymbol{\lambda}} = -[\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q]^{-1} \left[ [\mathbf{I}_q \otimes \boldsymbol{\Lambda}] \nabla \mathcal{A}(\mathbf{x})^T \mathbf{d} + \rho \boldsymbol{\lambda} \right] \quad (3.28)$$

Substituindo as Equações (3.22) e (3.28) na Equação (3.17a) resulta

$$\begin{aligned} & \left[ \mathbf{B} - \nabla \mathbf{g}(\mathbf{x}) \mathcal{G}(\mathbf{x})^{-1} \boldsymbol{\Gamma} \nabla \mathbf{g}(\mathbf{x})^T - \nabla \mathcal{A}(\mathbf{x}) [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q]^{-1} [\mathbf{I}_q \otimes \boldsymbol{\Lambda}] \nabla \mathcal{A}(\mathbf{x})^T \right] \mathbf{d} = \\ & -\nabla f(\mathbf{x}) + \nabla \mathbf{g}(\mathbf{x}) \rho \boldsymbol{\gamma} + \nabla \mathcal{A}(\mathbf{x}) [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q]^{-1} \rho \boldsymbol{\lambda} \end{aligned} \quad (3.29)$$

Substituindo as Equações (3.25) e (3.26) na Equação (3.29) e multiplicando ambos os termos por  $\mathbf{M}^{-1}$  obtêm-se

$$\mathbf{d} = -\mathbf{d}_0 + \rho \mathbf{M}^{-1} \left[ \nabla \mathbf{g}(\mathbf{x}) \boldsymbol{\gamma} + \nabla \mathcal{A}(\mathbf{x}) [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q]^{-1} \boldsymbol{\lambda} \right] \quad (3.30)$$



Denote-se por  $\mathbf{d}_1$  ao termo  $\mathbf{M}^{-1} [\nabla \mathbf{g}(\mathbf{x})\boldsymbol{\gamma} + \nabla \mathcal{A}(\mathbf{x}) [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q]^{-1} \boldsymbol{\lambda}]$ , obtém-se a relação entre as direções  $\mathbf{d}_0$  e  $\mathbf{d}$  dada pela Equação (3.19). De forma análoga podem-se encontrar as outras duas relações.

Finalmente, Para definir a cota superior do parâmetro  $\rho$  emprega-se a mesma rotina definida na Seção 2.5.3 do capítulo anterior.

- **Busca linear**

Para obter um novo ponto  $\mathbf{x}^{k+1}$  que seja viável, realiza-se uma minimização inexata de Armijo [38] ao longo da direção  $\mathbf{d}^k$  de maneira que o ponto  $\mathbf{x}^{k+1}$  esteja dentro da região viável.

A busca linear de Armijo consiste em achar o passo  $t$ , ou seja, o primeiro elemento da sequência  $\{1, \nu, \nu^2, \nu^3, \dots\}$ , tal que

$$f(\mathbf{x}^k + t^k \mathbf{d}^k) \leq f(\mathbf{x}^k) + t^k \eta \nabla f(\mathbf{x}^k)^T \mathbf{d}^k \quad (3.31)$$

$$g_i(\mathbf{x}^k + t^k \mathbf{d}^k) < 0, \quad i = 1, \dots, m \quad (3.32)$$

e

$$\mathcal{A}(\mathbf{x}^k + t^k \mathbf{d}^k) < 0 \quad (3.33)$$

onde  $\eta \in (0, 1)$  e  $\nu \in (0, 1)$  são parâmetros fixos.

- **Atualizações**

Para definir as atualizações dos valores de  $\boldsymbol{\gamma}$ ,  $\boldsymbol{\Lambda}$  e  $\mathbf{B}$  é necessário estabelecer as seguintes hipóteses adicionais:

**Hipótese 3.5.** *Existem números positivos  $\gamma^I$ ,  $\gamma^S$ , com  $\gamma^S > \gamma^I$ , e  $\bar{g}$  tal que  $0 < \gamma_i \leq \gamma^S$ ,  $i = 1, \dots, m$ , e  $\gamma_i \geq \gamma^I$  para qualquer  $i$  tal que  $g_i(\mathbf{x}) \geq -\bar{g}$ .*

**Hipótese 3.6.** *Existem números positivos  $\lambda^I$ ,  $\lambda^S$  tal que*

$$\lambda^I \mathbf{I}_q \preceq \boldsymbol{\Lambda} \preceq \lambda^S \mathbf{I}_q$$

**Hipótese 3.7.** *Existem números positivos  $\sigma_1$  e  $\sigma_2$  tal que*

$$\sigma_1 \|\mathbf{d}\|^2 \leq \mathbf{d}^T \mathbf{B} \mathbf{d} \leq \sigma_2 \|\mathbf{d}\|^2, \quad \mathbf{d} \in \mathbb{R}^n$$

Logo, realizam-se certas operações que preparam os dados para uma nova iteração, tais como:

- (i) Atualização do ponto:  $\mathbf{x}^{k+1} = \mathbf{x}^k + t^k \mathbf{d}^k$ ;
- (ii) Atualização da matriz quase-Newton  $\mathbf{B}^k$  com uma nova matriz simétrica definida positiva empregando a fórmula BFGS [38];

- (iii) Atualização dos valores de  $\gamma^{k+1}$  empregando o método descrito no algoritmo FDIPA [5];
- (iv) Atualização dos valores de  $\Lambda^{k+1}$  empregando o método descrito no algoritmo FDIPA-SDP [8].

### 3.2.6 Análise de convergência global

A análise de convergência do algoritmo FDIPA-GSDP segue o mesmo esquema do realizado no algoritmo FDIPA-SDP-NS, ou seja, prova-se o seguinte:

- Primeiramente, o Lema 3.1 garante que os sistemas (3.14) e (3.17) são bem definidos e têm soluções únicas e, portanto, o algoritmo não falha.
- No Lema 3.2 prova-se que se o algoritmo para depois de um número finito de iterações, então, ele gera um ponto de KKT do problema (3.2).
- Nos Lemas 3.3 e 3.4 provam-se que as direções  $\mathbf{d}_0$  e  $\mathbf{d}$  calculadas no passo 1 do algoritmo são direções de descida da função objetivo.
- Finalmente, no Lema 3.6 prova-se que qualquer ponto de acumulação gerado pelo algoritmo é um ponto estacionário do problema (3.2).

**Lema 3.1.** *Dados qualquer ponto regular  $\mathbf{x} \in \Omega_a$  do problema (2.1), qualquer matriz  $B$  simétrica definida positiva, qualquer  $\gamma \in \mathbb{R}^m$  tal que  $\gamma > 0$  se  $g_i(\mathbf{x}) = 0$  e qualquer matriz  $\Lambda \in \mathcal{S}_{++}^q$  tal que  $\Lambda$  e  $\mathcal{A}(\mathbf{x})$  comutam, a matriz*

$$\mathbf{W}(\mathbf{x}, \mathbf{B}, \gamma, \Lambda) = \begin{bmatrix} \mathbf{B} & \nabla \mathbf{g}(\mathbf{x}) & \nabla \mathcal{A}(\mathbf{x}) \\ \Gamma \nabla \mathbf{g}(\mathbf{x})^T & \mathcal{G}(\mathbf{x}) & \mathbf{0} \\ [\mathbf{I}_q \otimes \Lambda] \nabla \mathcal{A}(\mathbf{x})^T & \mathbf{0} & \mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q \end{bmatrix},$$

é não singular.

*Demonstração.* Para provar este teorema basta mostrar que o sistema

$$\mathbf{B}\mathbf{v}_1 + \nabla \mathbf{g}(\mathbf{x})\mathbf{v}_2 + \nabla \mathcal{A}(\mathbf{x})\mathbf{v}_3 = \mathbf{0} \quad (3.34a)$$

$$\Gamma \nabla \mathbf{g}(\mathbf{x})\mathbf{v}_1 + \mathcal{G}(\mathbf{x})\mathbf{v}_2 = \mathbf{0} \quad (3.34b)$$

$$[\mathbf{I}_q \otimes \Lambda] \nabla \mathcal{A}(\mathbf{x})^T \mathbf{v}_1 + [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \mathbf{v}_3 = \mathbf{0} \quad (3.34c)$$

para  $\mathbf{v}_1 \in \mathbb{R}^n$ ,  $\mathbf{v}_2 \in \mathbb{R}^m$  e  $\mathbf{v}_3 \in \mathbb{R}^{q^2}$ , tem solução única  $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)^T = (\mathbf{0}, \mathbf{0}, \mathbf{0})^T$ .

Pré-multiplicando a Equação (3.34a) por  $\mathbf{v}_1^T$  temos que

$$\mathbf{v}_1^T \mathbf{B} \mathbf{v}_1 + \sum_{i=1}^m v_{2i} \nabla g_i(\mathbf{x})^T \mathbf{v}_1 + \mathbf{v}_1^T \nabla \mathcal{A}(\mathbf{x}) \mathbf{v}_3 = \mathbf{0} \quad (3.35)$$

Das Equações (3.34c) e (3.34b), obtêm-se

$$\nabla g_i(\mathbf{x}) \mathbf{v}_1 = -\frac{v_{2i}}{\gamma_i} g_i(\mathbf{x}) \quad (3.36)$$

e

$$\mathbf{v}_1^T \nabla \mathcal{A}(\mathbf{x}) = -\mathbf{v}_3^T [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] [\mathbf{I}_q \otimes \mathbf{\Lambda}]^{-1} \quad (3.37)$$

Substituindo as Equações (3.37) e (3.36) na Equação (3.35), tem-se

$$\mathbf{v}_1^T \mathbf{B} \mathbf{v}_1 = \sum_{i=1}^m \frac{v_{2i}^2}{\gamma_i} g_i(\mathbf{x}) + \mathbf{v}_3^T [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] [\mathbf{I}_q \otimes \mathbf{\Lambda}]^{-1} \mathbf{v}_3 \quad (3.38)$$

Como  $\mathbf{\Lambda} \succcurlyeq \mathbf{0}$  e  $\mathcal{A}(\mathbf{x}) \prec \mathbf{0}$  e também pelo Lema 2.5 tem-se que a matriz  $[\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] [\mathbf{I}_q \otimes \mathbf{\Lambda}]^{-1}$  é semidefinida negativa. Por outro lado, como  $g_i(\mathbf{x}) \leq 0$  e  $\gamma_i > 0$ , tem-se que  $\mathbf{v}_1^T \mathbf{B} \mathbf{v}_1 \leq 0$ , e sendo a matriz  $\mathbf{B}$  definida positiva, então  $\mathbf{v}_1 = \mathbf{0}$ .

Da Equação (3.34c) resulta

$$[\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \mathbf{v}_3 = \mathbf{0}, \quad (3.39)$$

Como  $\mathcal{A}(\mathbf{x}) \prec \mathbf{0}$ , então a matriz  $[\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q]$  é definida negativa e, portanto, obtêm-se  $\mathbf{v}_3 = \mathbf{0}$ .

Substituindo os valores de  $\mathbf{v}_1 = \mathbf{v}_3 = \mathbf{0}$  na Equação (3.38), tem-se que

$$\sum_{i=1}^m \frac{v_{2i}^2}{\gamma_i} g_i(\mathbf{x}) = 0 \quad (3.40)$$

Logo, para  $g_i(\mathbf{x}) < 0$ , tem-se que  $v_{2i} = 0$ ,  $\forall i \notin I(\mathbf{x})$ , ou seja, para as restrições de desigualdade que não são ativas.

Da Equação (3.35)

$$\sum_{i=1}^m v_{2i} \nabla g_i(\mathbf{x}) = \mathbf{0}, \quad (3.41)$$

Como os vetores  $\nabla g_i(\mathbf{x})$ , para  $i \in I(\mathbf{x})$ , são linearmente independentes, então, tem-se que  $v_{2i} = 0$ .

Portanto, a solução única do sistema é  $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = (\mathbf{0}, \mathbf{0}, \mathbf{0})$ .  $\square$

**Lema 3.2.** *Se o algoritmo para no ponto  $\mathbf{x}^k$  e  $\mathbf{d}_0^k = \mathbf{0}$ , então  $\mathbf{x}^k$  é um ponto de KKT do problema (3.2).*

*Demonstração.* Se  $\mathbf{d}_0^k = \mathbf{0}$ , o sistema (3.14) resulta

$$\nabla f(\mathbf{x}^k) + \nabla \mathbf{g}(\mathbf{x}^k) \boldsymbol{\gamma}_0^{k+1} + \nabla \mathcal{A}(\mathbf{x}^k) \boldsymbol{\lambda}_0^{k+1} = \mathbf{0} \quad (3.42)$$

$$\mathcal{G}(\mathbf{x}^k) \boldsymbol{\gamma}_0^{k+1} = \mathbf{0} \quad (3.43)$$

$$[\mathcal{A}(\mathbf{x}^k) \otimes \mathbf{I}_q] \boldsymbol{\lambda}_0^{k+1} = \mathbf{0} \quad (3.44)$$

Pela Hipótese 3.2, o algoritmo verifica em cada iteração que  $g_i(\mathbf{x}^k) < 0$ , o qual implica que  $\gamma_{0i}^{k+1} = 0$ . Além disso,  $\mathcal{A}(\mathbf{x}^k) \prec \mathbf{0}$  e sendo a matriz  $[\mathcal{A}(\mathbf{x}^k) \otimes \mathbf{I}_q]$  não singular, então  $\boldsymbol{\lambda}_0^{k+1} = \mathbf{0}$ . Com isso tem-se que  $\nabla f(\mathbf{x}^k) = \mathbf{0}$  e desta maneira o ponto  $\mathbf{x}^k$  representa um ponto de KKT do problema (3.2).  $\square$

**Lema 3.3.** *O vetor  $\mathbf{d}_0$  definido no sistema linear (3.14) do algoritmo é uma direção de descida que verifica*

$$\mathbf{d}_0^T \nabla f(\mathbf{x}) \leq -\mathbf{d}_0^T \mathbf{B} \mathbf{d}_0 \quad (3.45)$$

*Demonstração.* Multiplicando a Equação (3.14a) por  $\mathbf{d}_0$  tem-se que

$$-\mathbf{d}_0^T \nabla f(\mathbf{x}) = \mathbf{d}_0^T \mathbf{B} \mathbf{d}_0 + \mathbf{d}_0^T \nabla \mathbf{g}(\mathbf{x}) \boldsymbol{\gamma}_0 + \mathbf{d}_0^T \nabla \mathbf{h}(\mathbf{x}) \boldsymbol{\mu}_0 + \mathbf{d}_0^T \nabla \mathcal{A}(\mathbf{x}) \boldsymbol{\lambda}_0 \quad (3.46)$$

Da Equação (3.14b), temos que

$$\boldsymbol{\Gamma} \nabla \mathbf{g}(\mathbf{x}) \mathbf{d}_0 = -\mathcal{G}(\mathbf{x}) \boldsymbol{\gamma}_0 \quad (3.47)$$

Portanto

$$\mathbf{d}_0^T \nabla \mathbf{g}(\mathbf{x}) = -\boldsymbol{\gamma}_0^T \boldsymbol{\Gamma}^{-1} \mathcal{G}(\mathbf{x}) \quad (3.48)$$

Da Equação (3.14c), tem-se que

$$\mathbf{d}_0^T \nabla \mathcal{A}(\mathbf{x}) = -\boldsymbol{\lambda}_0^T [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}] [\mathbf{I} \otimes \boldsymbol{\Lambda}]^{-1} \boldsymbol{\lambda}_0 \quad (3.49)$$

Substituindo as Equações (3.48) e (3.49) na Equação (3.46), tem-se que

$$\mathbf{d}_0^T \nabla f(\mathbf{x}) = -\mathbf{d}_0^T \mathbf{B} \mathbf{d}_0 + \boldsymbol{\gamma}_0^T \boldsymbol{\Gamma}^{-1} \mathcal{G}(\mathbf{x}) \boldsymbol{\gamma}_0 + \boldsymbol{\lambda}_0^T [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}] [\mathbf{I} \otimes \boldsymbol{\Lambda}]^{-1} \boldsymbol{\lambda}_0 \quad (3.50)$$

Na Equação (3.50), a matriz  $\mathbf{B}$  é definida positiva, então o termo  $\mathbf{d}_0^T \mathbf{B} \mathbf{d}_0$  é positivo. Além disso, a matriz  $\boldsymbol{\Gamma}$  é definida positiva e sendo  $g_i(\mathbf{x}) < 0$  implica que a matriz  $\mathcal{G}(\mathbf{x})$  é definida negativa, portanto, a matriz  $\boldsymbol{\Gamma}^{-1} \mathcal{G}(\mathbf{x})$  também é definida negativa e desta maneira o termo  $\boldsymbol{\gamma}_0^T \boldsymbol{\Gamma}^{-1} \mathcal{G}(\mathbf{x}) \boldsymbol{\gamma}_0$  é negativo. Por outro lado, sendo  $\mathcal{A}(\mathbf{x}) \prec \mathbf{0}$  e  $\boldsymbol{\lambda}_0 > \mathbf{0}$ , então a matriz  $\boldsymbol{\Lambda}$  é definida positiva e pelo Lema 2.5 tem-se que a matriz  $[\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}] [\mathbf{I} \otimes \boldsymbol{\Lambda}]^{-1}$  é definida negativa e, portanto, o termo  $\boldsymbol{\lambda}_0^T [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}] [\mathbf{I} \otimes \boldsymbol{\Lambda}]^{-1} \boldsymbol{\lambda}_0$  é negativo.

Finalmente, excluindo os termos  $\boldsymbol{\gamma}_0^T \boldsymbol{\Gamma}^{-1} \mathcal{G}(\mathbf{x}) \boldsymbol{\gamma}_0$  e  $\boldsymbol{\lambda}_0^T [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}] [\mathbf{I} \otimes \boldsymbol{\Lambda}]^{-1} \boldsymbol{\lambda}_0$  da Equação (3.50), pode-se concluir que  $\mathbf{d}_0^T \nabla f(\mathbf{x}) \leq -\mathbf{d}_0^T \mathbf{B} \mathbf{d}_0$ .  $\square$

**Lema 3.4.** *A direção de busca  $\mathbf{d}$  definido no passo 1 do algoritmo é uma direção de descida que verifica*

$$\mathbf{d}^T \nabla f(\mathbf{x}) \leq \xi \mathbf{d}_0^T \nabla f(\mathbf{x}) \quad (3.51)$$

*Demonstração.* Ver [8].  $\square$

**Lema 3.5.** *Existe  $\tau > 0$  tal que, para qualquer  $\mathbf{x} \in \Omega_a$  e  $\mathbf{d}$  calculado no passo 1b do algoritmo, as condições (3.31)-(3.33) são verificadas para qualquer  $t \in [0, \tau]$ .*

*Demonstração.* Ver [4] e [35]. □

**Lema 3.6.** *Qualquer ponto de acumulação  $\mathbf{x}^*$  da sequência  $\{\mathbf{x}^k\}$  gerada pelo algoritmo é um ponto estacionário do problema (3.2).*

*Demonstração.* Considere-se a sequência  $\{\mathbf{x}^k\}_{k \in \mathbb{K}}$ , onde  $\mathbb{K} \subset \mathbb{N} = \{1, 2, \dots\}$ . Dados  $\gamma$ ,  $\lambda$ ,  $\mathbf{B}$  e  $\rho$  limitados em  $\mathbb{K}$ , então existe uma subsequência dada pelo algoritmo  $\{\mathbf{x}^k, \gamma^k, \lambda^k, \mathbf{B}^k, \rho^k\}_{k \in \mathbb{K}_1}$ , sendo  $\mathbb{K}_1 \subset \mathbb{K}$ , que converge para  $\mathbf{x}^*, \gamma^*, \lambda^*, \mathbf{B}^*, \rho^*$ .

Considerando que  $\mathbf{d}$  depende continuamente de  $\mathbf{x}$ ,  $\gamma$ ,  $\lambda$ ,  $\mathbf{B}$  e  $\rho$ , então

$$\{\mathbf{d}^k\}_{k \in \mathbb{K}_1} \rightarrow \mathbf{d}^* \quad (3.52)$$

Agora, considerando a seguinte subsequência  $\{t^k\}_{k \in \mathbb{K}_2}$ , sendo  $\mathbb{K}_2 \subset \mathbb{K}_1$ , que converge para  $t^*$ . Do Lema 3.5 mostra-se que  $t^* > 0$ .

Empregando a Equação (3.31), pode ser reescrita como

$$f(\mathbf{x}^k + t\mathbf{d}) \leq f(\mathbf{x}^k) + \eta t^k \mathbf{d}^{kT} \nabla f(\mathbf{x}) \quad (3.53)$$

onde assume-se que  $k \in \mathbb{K}_2$  e  $k + 1$  não necessariamente pertence a  $\mathbb{K}_2$ .

Tomando limites em ambos os lados da expressão anterior quando  $k \rightarrow \infty$ , obtém-se

$$\eta t^* \mathbf{d}^{*T} \nabla f(\mathbf{x}^*) \geq 0 \quad (3.54)$$

como  $\eta \in (0, 1)$  e  $t^* > 0$ , então

$$\mathbf{d}^{*T} \nabla f(\mathbf{x}^*) \geq 0 \quad (3.55)$$

Empregando o Lema 3.4 e tomando limites para ambos os lados quando  $k \rightarrow \infty$ , tem-se

$$\mathbf{d}^{*T} \nabla f(\mathbf{x}^*) \leq \xi \mathbf{d}_0^{*T} \nabla f(\mathbf{x}^*) \quad (3.56)$$

como  $\xi \in (0, 1)$ , então

$$\mathbf{d}^{*T} \nabla f(\mathbf{x}^*) \leq 0 \quad (3.57)$$

Das Equações (3.55) e (3.57) pode-se concluir que  $\mathbf{d}^{*T} \nabla f(\mathbf{x}^*) = \mathbf{0}$  e empregando novamente o Lema 3.4, então

$$\mathbf{d}_0^{*T} \nabla f(\mathbf{x}^*) \geq 0 \quad (3.58)$$

Do Lema 3.3 e tomando limites para ambos os lados quando  $k \rightarrow \infty$ , tem-se

$$\mathbf{d}_0^{*T} \nabla f(\mathbf{x}^*) \leq -\mathbf{d}_0^{*T} \mathbf{B}^* \mathbf{d}_0^* \quad (3.59)$$

como a matriz  $\mathbf{B}^*$  é definida positiva, então  $\mathbf{d}_0^{*T} \mathbf{B}^* \mathbf{d}_0^* = 0$ , e portanto  $\mathbf{d}_0^* = \mathbf{0}$ .

Agora, tomando o limite quando  $k \rightarrow \infty$  no sistema (3.14), obtém-se

$$\mathbf{B}^* \mathbf{d}_0^* + \nabla \mathbf{g}(\mathbf{x}^*) \boldsymbol{\gamma}_0^* + \nabla \mathcal{A}(\mathbf{x}^*) \boldsymbol{\lambda}_0^* = -\nabla f(\mathbf{x}^*) \quad (3.60a)$$

$$\boldsymbol{\Gamma}^* \nabla \mathbf{g}(\mathbf{x}^*) \mathbf{d}_0^* + \mathcal{G}(\mathbf{x}^*) \boldsymbol{\gamma}_0^* = \mathbf{0} \quad (3.60b)$$

$$[\mathbf{I}_q \otimes \boldsymbol{\Lambda}^*] \nabla \mathcal{A}(\mathbf{x}^*)^T \mathbf{d}_0^* + [\mathcal{A}(\mathbf{x}^*) \otimes \mathbf{I}_q] \boldsymbol{\lambda}_0^* = 0 \quad (3.60c)$$

e sendo  $\mathbf{d}_0^* = \mathbf{0}$ , então têm-se

$$\nabla f(\mathbf{x}^*) + \nabla \mathbf{g}(\mathbf{x}^*) \boldsymbol{\gamma}_0^* + \nabla \mathcal{A}(\mathbf{x}^*) \boldsymbol{\lambda}_0^* = \mathbf{0} \quad (3.61a)$$

$$\mathcal{G}(\mathbf{x}^*) \boldsymbol{\gamma}_0^* = \mathbf{0} \quad (3.61b)$$

$$[\mathcal{A}(\mathbf{x}^*) \otimes \mathbf{I}_q] \boldsymbol{\lambda}_0^* = 0 \quad (3.61c)$$

sendo  $\boldsymbol{\Lambda}_0^* = \text{mat}(\boldsymbol{\lambda}_0^*)$ , então a Equação (3.61c) é equivalente a  $\boldsymbol{\Lambda}_0^* \mathcal{A}(\mathbf{x}^*) = 0$ , e portanto, pode-se concluir que o ponto  $\mathbf{x}^*$  é um ponto estacionário.

### 3.3 Problema de otimização com restrições de igualdade

Considere-se agora o problema geral de programação semidefinida não linear [10, 18], definido na Equação (3.1), formulado como

$$\left\{ \begin{array}{l} \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimizar}} \quad f(\mathbf{x}) \\ \text{sujeito a:} \quad \mathbf{g}(\mathbf{x}) \leq 0; \quad \mathbf{g}(\mathbf{x}) \in \mathbb{R}^m \\ \quad \quad \quad \mathbf{h}(\mathbf{x}) = 0; \quad \mathbf{h}(\mathbf{x}) \in \mathbb{R}^p \\ \quad \quad \quad \mathcal{A}(\mathbf{x}) \preceq 0; \quad \mathcal{A}(\mathbf{x}) \in \mathcal{S}^q \end{array} \right.$$

#### 3.3.1 Notações

As seguintes notações correspondem à inclusão das restrições de igualdade do problema (3.1).

- As derivadas parciais das funções  $h_i$  em relação a  $x_i$  são agrupadas na matriz  $\nabla \mathbf{h}(\mathbf{x}) \in \mathbb{R}^{n \times p}$ , denotada por

$$\nabla \mathbf{h}(\mathbf{x}) = [\nabla h_1(\mathbf{x}), \dots, \nabla h_m(\mathbf{x})] \quad (3.62)$$

- A função Lagrangiana  $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^{q \times q} \rightarrow \mathbb{R}$  associada ao problema (3.1) é definida como

$$L(\mathbf{x}, \boldsymbol{\gamma}, \boldsymbol{\mu}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \sum_{i=1}^p \mu_i h_i(\mathbf{x}) + \text{vec}(\mathcal{A}(\mathbf{x}))^T \boldsymbol{\lambda} \quad (3.63)$$

- O gradiente da função Lagrangiana em relação a  $\mathbf{x}$  é denotado por  $\nabla_{\mathbf{x}} L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^{q^2} \rightarrow \mathbb{R}^n$  e definido como

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\gamma}, \boldsymbol{\mu}, \boldsymbol{\lambda}) = \nabla f(\mathbf{x}) + \sum_{i=1}^m \lambda_i \nabla g_i(\mathbf{x}) + \sum_{i=1}^p \mu_i \nabla h_i(\mathbf{x}) + \nabla \mathbf{A}(\mathbf{x}) \boldsymbol{\lambda} \quad (3.64)$$

- A Hessiana da função Lagrangiana, associada ao problema (3.1), é denotada por  $\mathbf{H} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^{q^2} \rightarrow \mathbb{R}^{n \times n}$  e definida como

$$\mathbf{H}(\mathbf{x}, \boldsymbol{\gamma}, \boldsymbol{\mu}, \boldsymbol{\Lambda}) = \nabla^2 f(\mathbf{x}) + \sum_{i=1}^m \lambda_i \nabla^2 g_i(\mathbf{x}) + \sum_{i=1}^p \mu_i \nabla^2 h_i(\mathbf{x}) + \nabla^2 \mathbf{A}(\mathbf{x}) \boldsymbol{\lambda} \quad (3.65)$$

### 3.3.2 Condições de otimalidade de primeira ordem

As condições de otimalidade de primeira ordem de Karush-Kuhn-Tucker para este problema são

$$\nabla f(\mathbf{x}) + \nabla \mathbf{g}(\mathbf{x}) \boldsymbol{\gamma} + \nabla \mathbf{h}(\mathbf{x}) \boldsymbol{\mu} + \nabla \mathcal{A}(\mathbf{x}) \boldsymbol{\lambda} = \mathbf{0} \quad (3.66a)$$

$$\mathcal{G}(\mathbf{x}) \boldsymbol{\gamma} = \mathbf{0} \quad (3.66b)$$

$$\mathbf{h}(\mathbf{x}) = \mathbf{0} \quad (3.66c)$$

$$\boldsymbol{\Lambda} \mathcal{A}(\mathbf{x}) = \mathbf{0} \quad (3.66d)$$

$$\mathbf{g}(\mathbf{x}) \leq \mathbf{0} \quad (3.66e)$$

$$\mathcal{A}(\mathbf{x}) \preceq \mathbf{0} \quad (3.66f)$$

$$\boldsymbol{\gamma} \geq \mathbf{0} \quad (3.66g)$$

$$\boldsymbol{\Lambda} \succeq \mathbf{0} \quad (3.66h)$$

onde  $\boldsymbol{\mu} \in \mathbb{R}^p$  é o vetor dos multiplicadores de Lagrange das restrições de igualdade. Neste caso, a Equação (3.66c) representa a condição de complementariedade das restrições de igualdade.

As seguintes definições são uma extensão das condições de otimalidade de primeira ordem de Karush-Kuhn-Tucker do problema (3.1).

**Definição 3.4 (Ponto regular).** Um ponto  $\mathbf{x}$  é um ponto regular do problema (3.1) se os vetores  $\nabla g_i(\mathbf{x})$ ,  $i \in I(\mathbf{x})$ ,  $\nabla h_i(\mathbf{x})$ ,  $i = 1, \dots, p$  e  $\nabla \mathcal{A}(\mathbf{x}) \mathbf{b}_i(\mathbf{x})$ ,  $i = 1, \dots, r$  são linearmente independentes.

**Definição 3.5 (Ponto estacionário).** Um ponto  $\mathbf{x}$  é um ponto estacionário do problema (3.1) se existem uns vetores  $\boldsymbol{\gamma} \in \mathbb{R}^m$  e  $\boldsymbol{\mu} \in \mathbb{R}^p$  e uma matriz  $\boldsymbol{\Lambda} \in \mathbb{R}^{q \times q}$  tal que as seguintes condições são verificadas

$$\begin{aligned}\nabla f(\mathbf{x}) + \nabla \mathbf{g}(\mathbf{x})\boldsymbol{\gamma} + \nabla \mathcal{A}(\mathbf{x})\boldsymbol{\lambda} &= \mathbf{0} \\ \mathcal{G}(\mathbf{x})\boldsymbol{\gamma} &= \mathbf{0} \\ \mathbf{h}(\mathbf{x}) &= \mathbf{0} \\ \boldsymbol{\Lambda}\mathcal{A}(\mathbf{x}) &= \mathbf{0} \\ \mathcal{A}(\mathbf{x}) &\preceq \mathbf{0}\end{aligned}$$

**Definição 3.6 (Ponto de Karush-Kuhn-Tucker).** Um ponto estacionário  $\mathbf{x}$  é um ponto de Karush-Kuhn-Tucker do problema (3.1) se todas as Equações (3.66a)-(3.66h) são verificadas simultaneamente.

### 3.3.3 Hipóteses sobre o problema

As seguintes hipóteses incluem as restrições de igualdade do problema do problema (3.3).

**Hipótese 3.8.** Para qualquer  $\mathbf{x} \in \text{int}(\Omega_a)$ , verificam-se simultaneamente que  $\mathbf{g}(\mathbf{x}) < \mathbf{0}$ ,  $\mathbf{h}(\mathbf{x}) < \mathbf{0}$  e  $\mathcal{A}(\mathbf{x}) \prec \mathbf{0}$ .

**Hipótese 3.9.** As funções  $f$ ,  $\mathbf{g}$ ,  $\mathbf{h}$  e  $\mathcal{A}$  são continuamente diferenciáveis em  $\Omega_a$  e suas derivadas  $\nabla f$ ,  $\nabla \mathbf{g}$ ,  $\nabla \mathbf{h}$  e  $\frac{\partial \mathcal{A}}{\partial x_p}$ , para  $p = 1, \dots, n$ , satisfazem a condição de Lipschitz.

### 3.3.4 Descrição do algoritmo

O processo de análise deste algoritmo é similar do que o procedimento descrito na seção anterior mas realizando algumas alterações necessárias pela inclusão de restrições de igualdade.

- **Matriz do sistema linear**

A matriz do sistema linear deste problema, seguindo o procedimento descrito anteriormente, está dada por

$$\mathbf{W}(\mathbf{x}, \mathbf{B}, \boldsymbol{\gamma}, \boldsymbol{\mu}, \boldsymbol{\Lambda}) = \begin{bmatrix} \mathbf{B} & \nabla \mathbf{g}(\mathbf{x}) & \nabla \mathbf{h}(\mathbf{x}) & \nabla \mathcal{A}(\mathbf{x}) \\ \boldsymbol{\Gamma} \nabla \mathbf{g}(\mathbf{x})^T & \mathcal{G}(\mathbf{x}) & \mathbf{0} & \mathbf{0} \\ \nabla \mathbf{h}(\mathbf{x})^T & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ [\mathbf{I}_q \otimes \boldsymbol{\Lambda}] \nabla \mathcal{A}(\mathbf{x})^T & \mathbf{0} & \mathbf{0} & \mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q \end{bmatrix} \quad (3.67)$$



- **Direção de busca**

Seguindo o procedimento descrito na Subseção (3.2.5), define-se o seguinte sistema linear

$$\mathbf{B}^k \mathbf{d}_0^k + \nabla \mathbf{g}(\mathbf{x}^k) \boldsymbol{\gamma}_0^{k+1} + \nabla \mathbf{h}(\mathbf{x}^k) \boldsymbol{\mu}_0^{k+1} + \nabla \mathcal{A}(\mathbf{x}^k) \boldsymbol{\lambda}_0^{k+1} = -\nabla f(\mathbf{x}^k) \quad (3.68a)$$

$$\boldsymbol{\Gamma}^k \nabla \mathbf{g}(\mathbf{x}^k) \mathbf{d}_0^k + \mathcal{G}(\mathbf{x}^k) \boldsymbol{\gamma}_0^{k+1} = \mathbf{0} \quad (3.68b)$$

$$\nabla \mathbf{h}(\mathbf{x}^k)^T \mathbf{d}_0^k = -\mathbf{h}(\mathbf{x}^k) \quad (3.68c)$$

$$[\mathbf{I}_q \otimes \boldsymbol{\Lambda}^k] \nabla \mathcal{A}(\mathbf{x}^k)^T \mathbf{d}_0^k + [\mathcal{A}(\mathbf{x}^k) \otimes \mathbf{I}_q] \boldsymbol{\lambda}_0^{k+1} = 0 \quad (3.68d)$$

A solução deste sistema fornece uma direção  $\mathbf{d}_0^k$  no espaço primal e uma nova estimativa dos multiplicadores de Lagrange  $\boldsymbol{\gamma}_0^{k+1}$ ,  $\boldsymbol{\mu}_0^{k+1}$  e  $\boldsymbol{\lambda}_0^{k+1}$ . Pode-se observar que este sistema não depende do valor atual de  $\boldsymbol{\mu}^k$ .

Sem embargo,  $\mathbf{d}_0^k$  não é necessariamente uma direção viável, pois quando alguma restrição de igualdade tende a ser ativa,  $\mathbf{d}_0^k$  torna-se tangente à região viável.

Escrevendo a Equação (3.68c) para a  $i$ -ésima restrição de igualdade, tem-se a seguinte expressão

$$\nabla h_i(\mathbf{x}^k)^T \mathbf{d}_0^k = -h_i(\mathbf{x}^k), \quad i = 1, \dots, p, \quad (3.69)$$

Se  $h_i(\mathbf{x}^k) = 0$ , para  $i = 1, \dots, p$ , então a Equação (3.69) resulta

$$\nabla h_i(\mathbf{x}^k)^T \mathbf{d}_0^k = 0 \quad (3.70)$$

Da Equação (3.70) pode-se concluir que os vetores  $\nabla h_i(\mathbf{x}^k)$  e  $\mathbf{d}_0^k$  são ortogonais e, portanto, a direção  $\mathbf{d}_0^k$  é inviável.

Neste problema, o segundo sistema linear é definido como

$$\mathbf{B}^k \mathbf{d}^k + \nabla \mathbf{g}(\mathbf{x}^k) \bar{\boldsymbol{\gamma}}^{k+1} + \nabla \mathbf{h}(\mathbf{x}^k) \bar{\boldsymbol{\mu}}^{k+1} + \nabla \mathcal{A}(\mathbf{x}^k) \bar{\boldsymbol{\lambda}}^{k+1} = -\nabla f(\mathbf{x}^k) \quad (3.71a)$$

$$\boldsymbol{\Gamma}^k \nabla \mathbf{g}(\mathbf{x}^k) \mathbf{d}^k + \mathcal{G}(\mathbf{x}^k) \bar{\boldsymbol{\gamma}}^{k+1} = -\rho^k \boldsymbol{\gamma}^k \quad (3.71b)$$

$$\nabla \mathbf{h}(\mathbf{x}^k)^T \mathbf{d}^k = -\boldsymbol{\omega}^k \quad (3.71c)$$

$$[\mathbf{I}_q \otimes \boldsymbol{\Lambda}^k] \nabla \mathcal{A}(\mathbf{x}^k)^T \mathbf{d}^k + [\mathcal{A}(\mathbf{x}^k) \otimes \mathbf{I}_q] \bar{\boldsymbol{\lambda}}^{k+1} = -\rho^k \boldsymbol{\lambda}^k \quad (3.71d)$$

onde  $\boldsymbol{\omega} = (1, \dots, 1) \in \mathbb{R}^p$  representa os fatores de deflexão associados às restrições de igualdade. Se as restrições de igualdade fossem lineares, então, considera-se  $\omega_i = 0$ . Este sistema fornece uma nova direção  $\mathbf{d}^k$  e novas estimativas dos multiplicadores de Lagrange  $\bar{\boldsymbol{\gamma}}^{k+1}$ ,  $\bar{\boldsymbol{\mu}}^{k+1}$  e  $\bar{\boldsymbol{\lambda}}^{k+1}$ .

O sistema linear auxiliar resulta

$$\mathbf{B}^k \mathbf{d}_1^k + \nabla \mathbf{g}(\mathbf{x}^k) \gamma_1^{k+1} + \nabla \mathbf{h}(\mathbf{x}^k) \boldsymbol{\mu}_1^{k+1} + \nabla \mathcal{A}(\mathbf{x}^k) \boldsymbol{\lambda}_1^{k+1} = \mathbf{0} \quad (3.72a)$$

$$\mathbf{\Gamma}^k \nabla \mathbf{g}(\mathbf{x}^k) \mathbf{d}_1^k + \mathcal{G}(\mathbf{x}^k) \gamma_1^{k+1} = -\gamma^k \quad (3.72b)$$

$$\nabla \mathbf{h}(\mathbf{x}^k)^T \mathbf{d}_1^k = -\boldsymbol{\omega}^k \quad (3.72c)$$

$$[\mathbf{I}_q \otimes \boldsymbol{\Lambda}^k] \nabla \mathcal{A}(\mathbf{x}^k)^T \mathbf{d}_1^k + [\mathcal{A}(\mathbf{x}^k) \otimes \mathbf{I}_q] \boldsymbol{\lambda}_1^{k+1} = -\boldsymbol{\lambda}^k \quad (3.72d)$$

com

$$\mathbf{d}^k = \mathbf{d}_0^k + \rho^k \mathbf{d}_1^k \quad (3.73)$$

$$\bar{\gamma}^{k+1} = \gamma_0^{k+1} + \rho^k \gamma_1^{k+1} \quad (3.74)$$

$$\bar{\boldsymbol{\mu}}^{k+1} = \boldsymbol{\mu}_0^{k+1} + \rho^k \boldsymbol{\mu}_1^{k+1} \quad (3.75)$$

$$\bar{\boldsymbol{\lambda}}^{k+1} = \boldsymbol{\lambda}_0^{k+1} + \rho^k \boldsymbol{\lambda}_1^{k+1} \quad (3.76)$$

A inclusão das restrições de igualdade produz que o valor mínimo da função objetivo seja maior ou igual do que o valor mínimo da função no problema (3.2) devido a que sempre tem que ficar ativas na solução.

Assim, emprega-se a seguinte função auxiliar chamada função potencial [4] que é definida como

$$\varphi_c(\mathbf{x}) = f(\mathbf{x}) + \sum_{i=1}^p c_i |h_i(\mathbf{x})| \quad (3.77)$$

onde  $c_i$  são constantes positivas.

Em uma iteração  $k$ ,  $c_i^k$  é tal que

$$\frac{h_i(\mathbf{x})}{|h_i(\mathbf{x})|} (c_i^k + \mu_{0i}^{k+1}) < 0, \quad i = 1, \dots, p \quad (3.78)$$

Herskovits provou em [4] que a direção  $\mathbf{d}_0$ , obtida do sistema linear (3.68), é uma direção de descida da função potencial  $\varphi_c$  desde que  $c_i > |\mu_{0i}|$ , para  $i = 1, \dots, p$ . Além disso,  $\varphi_c(\mathbf{x})$  é uma função de penalidade exata, ou seja, existe um  $c_i$  finito tal que que o mínimo da função potencial corresponde à solução do problema original (3.3).

O uso da função de penalidade exata tem a vantagem de não precisar de parâmetros de penalidade crescendo indefinidamente.

Por outro lado, a função potencial  $\varphi_c$  não é continuamente diferenciável já que não possui derivada única nos pontos onde as restrições de igualdade ficam ativas, ou seja  $h_i(\mathbf{x}) = 0$ . Isto pode ser evitado, considerando que o ponto

inicial seja interior a uma região viável definida por

$$\Delta = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) \leq \mathbf{0}, \mathcal{A}(\mathbf{x}) \preceq \mathbf{0}\} \quad (3.79)$$

e seu interior

$$\Delta^0 = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) < \mathbf{0}, \mathbf{h}(\mathbf{x}) < \mathbf{0}, \mathcal{A}(\mathbf{x}) \prec \mathbf{0}\}, \quad (3.80)$$

desta maneira, o problema torna-se muito similar ao problema (3.1).

Em cada iteração pode-se evitar que o novo ponto fique na fronteira da região viável empregando o mesmo procedimento de deflexão da direção  $\mathbf{d}_0$  descrito na seção anterior que garante um decréscimo da função  $\varphi_c$ .

Sendo  $\mathbf{d}_0$  uma direção de descida de  $\varphi_c$ , ou seja  $\mathbf{d}_0^T \nabla \varphi_c(\mathbf{x}) > 0$ , então é possível estabelecer um limite na deflexão impondo

$$\mathbf{d}^T \nabla \varphi_c(\mathbf{x}) \leq \xi \mathbf{d}_0^T \nabla \varphi_c(\mathbf{x}), \quad \xi \in (0, 1) \quad (3.81)$$

a qual é uma expressão similar à Equação (2.59) onde foi substituído o termo  $\nabla f(\mathbf{x})$  por  $\nabla \varphi_c(\mathbf{x})$ .

Assim, define-se a cota superior de  $\rho$

$$\rho \leq (\xi - 1) \frac{\mathbf{d}_0^{kT} \nabla \varphi_c(\mathbf{x}^k)}{\mathbf{d}_1^{kT} \nabla \varphi_c(\mathbf{x}^k)} \quad (3.82)$$

se  $\mathbf{d}_1^{kT} \nabla f(\mathbf{x}^k) > 0$ . Caso contrário, se  $\mathbf{d}_1^{kT} \nabla f(\mathbf{x}^k) \leq 0$ , assume-se que  $\rho \leq \varphi \|\mathbf{d}_0\|^2$ .

- **Busca linear**

Na busca linear é necessário usar técnicas para evitar a não diferenciabilidade da função potencial  $\varphi_c$ , esta busca deve garantir que o novo ponto seja interior à região viável  $\Delta$  e também obter um decréscimo de  $\varphi_c$ . Isto é possível substituindo  $f(\mathbf{x})$  por  $\varphi_c(\mathbf{x})$  na Equação (3.31).

No início do processo de otimização as restrições de igualdade são redefinidas de tal maneira que todas elas tenham o mesmo sinal, ou seja, são sempre negativas como no caso das restrições de desigualdade.

- **Atualizações**

As regras de atualização da matriz  $\mathbf{B}$ , do vetor dos multiplicadores de Lagrange e da matriz dos multiplicadores de Lagrange ficam inalteráveis devido a que o sistema linear não depende de  $\boldsymbol{\mu}$ .

As hipóteses do algoritmo para este problema são as mesmas do que as estabelecidas no algoritmo sem restrições de igualdade acrescentando a seguinte hipótese acerca do fator  $\omega$  no processo de atualização do presente algoritmo.

**Hipótese 3.10.** *Existem números positivos  $\omega_1$  e  $\omega_2$  tal que*

$$\omega_1 \leq \omega_i \leq \omega_2$$

### 3.3.5 Análise de convergência global

A análise de convergência do algoritmo FDIPA-GSDP considerando restrições de igualdade segue o esquema da demonstração realizada na seção anterior na qual substitui-se a função objetivo  $f(\mathbf{x})$  pela função potencial  $\varphi_c(\mathbf{x})$ .

Aqui, somente será provado o Lema 3.7 que garante que os sistemas (3.68), (3.71) e (3.72) têm soluções únicas.

**Lema 3.7.** *Dados qualquer  $\mathbf{x} \in \Omega_a$ , qualquer matriz  $B$  simétrica definida positiva, qualquer  $\gamma \in \mathbb{R}^m$  tal que  $\gamma > 0$  se  $g_i(\mathbf{x}) = 0$  e qualquer  $\Lambda \in \mathcal{S}_{++}^q$ , tal que  $\Lambda$  e  $\mathcal{A}(\mathbf{x})$  comutam, a matriz*

$$\mathbf{W}(\mathbf{x}, \mathbf{B}, \boldsymbol{\gamma}, \boldsymbol{\mu}, \Lambda) = \begin{bmatrix} \mathbf{B} & \nabla \mathbf{g}(\mathbf{x}) & \nabla \mathbf{h}(\mathbf{x}) & \nabla \mathcal{A}(\mathbf{x}) \\ \Gamma \nabla \mathbf{g}(\mathbf{x})^T & \mathcal{G}(\mathbf{x}) & \mathbf{0} & \mathbf{0} \\ \nabla \mathbf{h}(\mathbf{x})^T & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ [\mathbf{I}_q \otimes \Lambda] \nabla \mathcal{A}(\mathbf{x})^T & \mathbf{0} & \mathbf{0} & \mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q \end{bmatrix}$$

é não singular.

*Demonstração.* Para provar este lema basta mostrar que o seguinte sistema

$$\mathbf{B}\mathbf{v}_1 + \nabla \mathbf{g}(\mathbf{x})\mathbf{v}_2 + \nabla \mathbf{h}(\mathbf{x})\mathbf{v}_3 + \nabla \mathcal{A}(\mathbf{x})\mathbf{v}_4 = \mathbf{0} \quad (3.83a)$$

$$\Gamma \nabla \mathbf{g}(\mathbf{x})\mathbf{v}_1 + \mathcal{G}(\mathbf{x})\mathbf{v}_2 = \mathbf{0} \quad (3.83b)$$

$$\nabla \mathbf{h}(\mathbf{x})^T \mathbf{v}_1 = \mathbf{0} \quad (3.83c)$$

$$[\mathbf{I}_q \otimes \Lambda] \nabla \mathcal{A}(\mathbf{x})^T \mathbf{v}_1 + [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \mathbf{v}_4 = \mathbf{0} \quad (3.83d)$$

tem solução única  $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4]^T = [\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}]^T$ .

Pré-multiplicando a Equação (3.83a) por  $\mathbf{v}_1$  resulta

$$\mathbf{v}_1^T \mathbf{B}\mathbf{v}_1 + \sum_{i=1}^m v_{2i} \nabla g_i(\mathbf{x})^T \mathbf{v}_1 + \sum_{i=1}^p v_{3i} \nabla h_i(\mathbf{x})^T \mathbf{v}_1 + \mathbf{v}_1^T \nabla \mathcal{A}(\mathbf{x})\mathbf{v}_4 = \mathbf{0} \quad (3.84)$$

Das Equações (3.83b), (3.83c) e (3.83d), obtêm-se as seguintes expressões

$$\nabla g_i(\mathbf{x})^T \mathbf{v}_1 = -\frac{v_{2i}}{\gamma_i} g_i(\mathbf{x}), \quad (3.85)$$

$$\nabla h_i(\mathbf{x})^T \mathbf{v}_1 = \mathbf{0} \quad (3.86)$$

e

$$\mathbf{v}_1^T \nabla \mathcal{A}(\mathbf{x}) = -\mathbf{v}_4^T [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] [\mathbf{I}_q \otimes \mathbf{\Lambda}]^{-1} \quad (3.87)$$

Substituindo as Equações (3.85), (3.86) e (3.87) na Equação (3.84) resulta

$$\mathbf{v}_1^T \mathbf{B} \mathbf{v}_1 = \sum_{i=1}^m \frac{v_{2i}^2}{\gamma_i} g_i(\mathbf{x}) + \mathbf{v}_4^T [\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] [\mathbf{I}_q \otimes \mathbf{\Lambda}]^{-1} \mathbf{v}_4, \quad (3.88)$$

esta expressão é idêntica à Equação (3.38), portanto, pode-se concluir que  $\mathbf{v}_1 = \mathbf{0}$ .

Substituindo o valor de  $\mathbf{v}_1$  na Equação (3.83d), obtêm-se

$$[\mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q] \mathbf{v}_4 = \mathbf{0}, \quad (3.89)$$

e empregando a mesma condição da Equação (3.39), tem-se que  $\mathbf{v}_4 = \mathbf{0}$ .

Por outro lado, substituindo os valores de  $\mathbf{v}_1$  e  $\mathbf{v}_4$  na Equação (3.88), obtêm-se a seguinte expressão

$$\sum_{i=1}^m \frac{v_{2i}^2}{\gamma_i} g_i(\mathbf{x}) = \mathbf{0}, \quad (3.90)$$

Considerando somente as restrições inativas, isto é,  $g_i(\mathbf{x}) < 0$ , para  $i \notin I(\mathbf{x})$ , então, tem-se que  $v_{2i} = 0$ .

Da Equação (3.83a) podem-se estabelecer as seguintes expressões

$$\sum_{i=1}^m v_{2i} \nabla g_i(\mathbf{x}) = \mathbf{0}, \quad (3.91)$$

e

$$\sum_{i=1}^p v_{3i} \nabla h_i(\mathbf{x}) = \mathbf{0}, \quad (3.92)$$

e sendo os vetores  $\nabla g_i(\mathbf{x})$ , para  $i \in I(\mathbf{x})$ , e  $\nabla h_i(\mathbf{x})$ , para  $i = 1, \dots, p$ , linearmente independentes, então, tem-se que  $v_{2i} = 0$  e  $\mathbf{v}_3 = \mathbf{0}$ . Desta maneira, combinando os resultados obtidos para as restrições ativas e inativas, conclui-se que  $\mathbf{v}_2 = \mathbf{0}$ .

Portanto, pode-se concluir que a solução única do sistema é o vetor  $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4]^T = [\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}]^T$ .  $\square$

### 3.4 Algoritmo FDIPA-GSDP

Finalmente, apresenta-se o algoritmo de pontos interiores e direções viáveis para programação semidefinida generalizada (FDIPA-GSDP) que consiste no seguinte:

**Parâmetros de controle:**  $\xi \in (0, 1)$ ,  $\eta \in (0, 1)$ ,  $\varphi > 0$  e  $\nu \in (0, 1)$ .

**Dados Iniciais:**  $\mathbf{x} \in \text{int}(\Delta)$ ,  $\boldsymbol{\gamma} > \mathbf{0}$ ,  $\mathbf{B} \in \mathcal{S}_{++}^n$ ,  $\boldsymbol{\Lambda} \in \mathcal{S}_{++}^q$  comutando com  $\mathcal{A}(\mathbf{x})$ ,  $\boldsymbol{\omega} = (1, \dots, 1) \in \mathbb{R}^p$  e  $\mathbf{c} \in \mathbb{R}_+^p$ .

**Passo 1: Cálculo da direção de busca**

**Passo 1a:** Resolva o primeiro sistema linear para obter  $\mathbf{d}_0 \in \mathbb{R}^n$ ,  $\boldsymbol{\gamma}_0 \in \mathbb{R}^m$ ,  $\boldsymbol{\mu}_0 \in \mathbb{R}^p$  e  $\boldsymbol{\lambda}_0 \in \mathbb{R}^{q^2}$ .

$$\begin{bmatrix} \mathbf{B} & \nabla \mathbf{g}(\mathbf{x}) & \nabla \mathbf{h}(\mathbf{x}) & \nabla \mathcal{A}(\mathbf{x}) \\ \boldsymbol{\Gamma} \nabla \mathbf{g}(\mathbf{x})^T & \mathcal{G}(\mathbf{x}) & \mathbf{0} & \mathbf{0} \\ \nabla \mathbf{h}(\mathbf{x})^T & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ [\mathbf{I}_q \otimes \boldsymbol{\Lambda}] \nabla \mathcal{A}(\mathbf{x})^T & \mathbf{0} & \mathbf{0} & \mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q \end{bmatrix} \begin{bmatrix} \mathbf{d}_0 \\ \boldsymbol{\gamma}_0 \\ \boldsymbol{\mu}_0 \\ \boldsymbol{\lambda}_0 \end{bmatrix} = \begin{bmatrix} -\nabla f(\mathbf{x}) \\ \mathbf{0} \\ -\mathbf{h}(\mathbf{x}) \\ \mathbf{0} \end{bmatrix} \quad (3.93)$$

Se  $\mathbf{d}_0 = \mathbf{0}$ , parar.

**Passo 1b:** Resolva o segundo sistema linear para obter  $\mathbf{d}_1 \in \mathbb{R}^n$ ,  $\boldsymbol{\gamma}_1 \in \mathbb{R}^m$ ,  $\boldsymbol{\mu}_1 \in \mathbb{R}^p$  e  $\boldsymbol{\lambda}_1 \in \mathbb{R}^{q^2}$ .

$$\begin{bmatrix} \mathbf{B} & \nabla \mathbf{g}(\mathbf{x}) & \nabla \mathbf{h}(\mathbf{x}) & \nabla \mathcal{A}(\mathbf{x}) \\ \boldsymbol{\Gamma} \nabla \mathbf{g}(\mathbf{x})^T & \mathcal{G}(\mathbf{x}) & \mathbf{0} & \mathbf{0} \\ \nabla \mathbf{h}(\mathbf{x})^T & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ [\mathbf{I}_q \otimes \boldsymbol{\Lambda}] \nabla \mathcal{A}(\mathbf{x})^T & \mathbf{0} & \mathbf{0} & \mathcal{A}(\mathbf{x}) \otimes \mathbf{I}_q \end{bmatrix} \begin{bmatrix} \mathbf{d}_1 \\ \boldsymbol{\gamma}_1 \\ \boldsymbol{\mu}_1 \\ \boldsymbol{\lambda}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ -\boldsymbol{\gamma} \\ -\boldsymbol{\omega} \\ -\boldsymbol{\lambda} \end{bmatrix} \quad (3.94)$$

**Passo 1c:** Seja a seguinte função potencial

$$\varphi(\mathbf{x}) = f(\mathbf{x}) + \sum_{i=1}^p c_i |h_i(\mathbf{x})| \quad (3.95)$$

**Passo 1d:** Se  $c_i < -1.2\mu_0(i)$ , então faça  $c_i = -2\mu_0(i)$ ;  $i = 1, \dots, p$

**Passo 1e:** Se  $\mathbf{d}_1^T \nabla f(\mathbf{x}) > 0$ , calcule

$$\rho = \min \left\{ \varphi \|\mathbf{d}_0\|^2, (\xi - 1) \frac{\mathbf{d}_0^T \nabla \varphi(\mathbf{x})}{\mathbf{d}_1^T \nabla \varphi(\mathbf{x})} \right\} \quad (3.96)$$

caso contrário, calcule

$$\rho = \varphi \|\mathbf{d}_0\|^2 \quad (3.97)$$

**Passo 1f:** Calcule a direção de busca  $\mathbf{d}$

$$\mathbf{d} = \mathbf{d}_0 + \rho \mathbf{d}_1 \quad (3.98)$$

**Passo 2: Busca Linear**

Calcule o comprimento do passo  $t$ , o primeiro número da série  $\{1, \nu, \nu^2, \nu^3, \dots\}$ , tal que

$$\varphi_c(\mathbf{x} + t\mathbf{d}) \leq \varphi_c(\mathbf{x}) + t\eta \nabla \varphi(\mathbf{x})^T \mathbf{d} \quad (3.99)$$

$$g_i(\mathbf{x} + t\mathbf{d}) < 0; \quad i = 1, \dots, m \quad (3.100)$$

$$h_i(\mathbf{x} + t\mathbf{d}) < 0; \quad i = 1, \dots, p \quad (3.101)$$

e

$$\mathcal{A}(\mathbf{x} + t\mathbf{d}) \prec \mathbf{0} \quad (3.102)$$

**Passo 3: Atualizações**

**Passo 3a:** Obtenha um novo ponto, como  $\mathbf{x} := \mathbf{x} + t\mathbf{d}$ ;

**Passo 3b:** Defina novos valores para  $\mathbf{B} \in \mathcal{S}_{++}^n$ ;

**Passo 3c:** Defina novos valores para  $\gamma_i > 0, i = 1, \dots, m$ ;

**Passo 3d:** Defina novos valores para  $\mathbf{\Lambda} \in \mathcal{S}_{++}^q$  comutando com  $\mathcal{A}(\mathbf{x})$ ;

**Passo 3e:** Retorne ao Passo 1.

# Capítulo 4

## Resultados Numéricos

### 4.1 Introdução

No presente capítulo apresentam-se os resultados numéricos de sete problemas teste correspondentes a problemas algébricos e problemas de otimização estrutural amplamente utilizados na literatura com a finalidade de avaliar o comportamento do algoritmo FDIPA-GSDP o qual está estruturado em sete seções distribuídos da seguinte maneira:

- Na seção 4.2, apresentam-se os dados de controle do algoritmo: parâmetros, valores iniciais dos multiplicadores de Lagrange, critérios de parada e atualizações dos multiplicadores de Lagrange e da matriz Quase-Newton;
- Na seção 4.3, apresentam-se dois problemas algébricos da coleção de problemas teste de Hock e Schittkowski [41] que foram modificados para incluir matrizes semidefinidas.
- Nas seções 4.4, 4.5 e 4.6, apresentam-se aplicações em problemas de otimização estrutural de treliças planas e espaciais, descritos e analisados em [27], [42], [43], [44], [45], [46], [47] e [25], os quais foram modificados para incluir restrições matriciais;
- Finalmente, na seção 4.7 mostra-se o resumo dos resultados numéricos.

Nos exemplos de otimização estrutural foi empregado o programa CALFEM [48] para fazer a análise estrutural (Ver Apêndice C) e foi usado o mesmo material para todos os exemplos com as seguintes propriedades mecânicas:

- Módulo de elasticidade  $1 \times 10^7$  psi (68.95 GPa e
- peso específico  $0.1 \text{ lb/in}^3$  ( $27154 \text{ N/m}^3$ ).



O cálculo do gradiente da função objetivo e das restrições em relação às variáveis de projeto foi realizado por métodos analíticos em todos os exemplos, exceto no problema de otimização estrutural dimensional e de forma onde foi empregado o método das diferenças finitas.

A notação empregada nas Tabelas 4.11, 4.15, 4.19, 4.23 e 4.26 que define a dimensão do problema, tais como: o número de variáveis de projeto e o número de restrições, é descrita na Tabela 4.1.

Tabela 4.1: Notação empregada para definir a dimensão do problema

Notação	Descrição
$n$	Número de variáveis de projeto
$m$	Número de restrições de desigualdade
$p$	Número de restrições de igualdade
$q$	Número de filas ou colunas da restrição matricial
$nt$	Número total de restrições de igualdade e desigualdade
$nc$	Número de restrições de caixa

Por outro lado, a notação empregada nas Tabelas 4.4, 4.7 e 4.29 que representa o esforço computacional do algoritmo na solução dos problemas teste é descrita na Tabela 4.2.

Tabela 4.2: Notação empregada para definir o esforço computacional

Notação	Descrição
$NI$	Número de iterações do algoritmo
$NF$	Número de avaliações da função objetivo

No caso que as restrições de desigualdade e/ou as restrições matriciais impostas nos problemas teste não sejam verificadas em um ponto inicial deve-se empregar o seguinte programa matemático auxiliar

$$\begin{cases} \underset{\mathbf{x}, z}{\text{minimizar}} & z \\ \text{sujeito a:} & \mathbf{g}(\mathbf{x}) \leq \mathbf{0} \\ & \mathcal{A}(\mathbf{x}) - z\mathbf{I}_q \preceq \mathbf{0} \end{cases}, \quad (4.1)$$

o qual fornece um ponto inicial viável. Aqui,  $z$  representa uma variável escalar auxiliar.

Todos os exemplos foram implementados na linguagem de programação MATLAB e executados em um computador com as seguintes características técnicas:

processador Intel Core i5 de 3,6 GHz, 16 GB de memória RAM e sistema operacional Windows 7 de 64 bits.

## 4.2 Dados de controle do algoritmo

A seguir, mostram-se os dados de controle considerados na solução dos problemas teste tais como: os parâmetros de controle do algoritmo, os dados iniciais, os critérios de parada e as rotinas de atualização dos dados.

- **Parâmetros de controle**

Em todos os problemas foi definido o mesmo conjunto de parâmetros para o cálculo da deflexão e para a busca linear de Armijo:  $\xi = 0.7$  (exceto no problema de otimização estrutural topológica onde  $\xi = 0.9$ ),  $\varphi = 1$ ,  $\eta = 0.1$  e  $\nu = 0.7$ .

- **Dados iniciais**

Os dados iniciais correspondem aos valores iniciais dos multiplicadores de Lagrange, da matriz Quase-Newton, do fator de deflexão e das constantes das restrições de igualdade. Assim, têm-se:

- (i) Os valores iniciais dos multiplicadores de Lagrange das restrições vetoriais de desigualdade foram  $\gamma_i = 1$ , para  $i = 1, \dots, m$ ;
- (ii) O valor inicial da matriz dos multiplicadores de Lagrange das restrições matriciais semidefinidas foi a identidade, ou seja  $\mathbf{\Lambda} = \mathbf{I}_q$ ;
- (iii) O valor inicial da matriz Quase-Newton foi também a identidade, ou seja  $\mathbf{B} = \mathbf{I}_n$ ;
- (iv) Finalmente, os valores iniciais das constantes das restrições de igualdade foram  $c_i = 1$ , para  $i = 1, \dots, p$ .

- **Critérios de parada**

- (i) Os critérios de parada do algoritmo FDIPA-GSDP considerados em todos os problemas testes foram a norma da direção de busca  $\|\mathbf{d}_0\| < 1.0 \times 10^{-4}$  e a norma do gradiente da função lagrangiana  $\|\nabla L\| < 1.0 \times 10^{-4}$ ;
- (ii) As condições de complementariedade das restrições vetoriais e matriciais foram quantificadas em cada iteração para verificar a precisão destas (Ver Apêndice B).

- **Atualizações**

- (i) A rotina de atualização dos multiplicadores de Lagrange para as restrições de desigualdade é descrita em [5] e corresponde ao algoritmo FDIPA;
- (ii) A rotina de atualização da matriz dos multiplicadores de Lagrange das restrições matriciais semidefinidas foi descrito na Seção 1.2.1;
- (iii) Finalmente, na atualização da matriz quase-Newton empregou-se o método Broyden-Fletcher-Goldfarb-Shanno (BFGS) descrito em [38].

### 4.3 Problemas algébricos

Os problemas algébricos apresentados a continuação correspondem aos problemas 43 e 71 da coleção de problemas teste de Hock e Schittkowski [49] modificados por Yamashita [17] e Kočvara [18] para incluir também uma restrição matricial semidefinida denominados: problema 43 aumentado e problema 71 aumentado, respectivamente. Realizou-se uma comparação dos resultados obtidos nestes problemas entre os algoritmos FDIPA-GSDP e PENLAB [31].

#### 4.3.1 Exemplo 1 - Problema 43 aumentado

Este primeiro exemplo é formulado como

$$\left\{ \begin{array}{l} \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimizar}} \quad x_1^2 + x_2^2 + 2x_3^2 + x_4^2 - 5x_1 - 5x_2 - 21x_3 + 7x_4 \\ \text{sujeito a:} \quad 8 - x_1^2 - x_2^2 - x_3^2 - x_4^2 - x_1 + x_2 - x_3 + x_4 \leq 0 \\ \quad \quad \quad 10 - x_1^2 - 2x_2^2 - x_3^2 - 2x_4^2 + x_1 + x_4 \leq 0 \\ \quad \quad \quad 5 - 2x_1^2 - x_2^2 - x_3^2 - 2x_1 + x_2 + x_4 \leq 0 \\ \quad \quad \quad \begin{bmatrix} -x_2 - x_3 & 0 & 0 & 0 \\ 0 & -2x_4 & -x_1 & 0 \\ 0 & -x_1 & -2x_4 & 0 \\ 0 & 0 & 0 & -x_2 - x_3 \end{bmatrix} \preceq \mathbf{0} \end{array} \right. \quad (4.2)$$

O número de variáveis de projeto e o número de restrições são mostrados na Tabela 4.3.

Tabela 4.3: Número de variáveis e de restrições do problema 43 aumentado.

$n$	$m$	$p$	$q$	$nt$	$nc$
4	3	0	4	3	0

A Tabela 4.4 mostra os valores inicial e final da função objetivo e das variáveis de projeto para o problema 43 aumentado.

Tabela 4.4: Solução ótima do problema 43 aumentado.

Variável de projeto	Inicial	Final
$x_1$	0.0000	0.0000
$x_2$	0.0000	1.0384
$x_3$	0.0000	2.2271
$x_4$	0.0000	0.0000
$f$	0.0000	-40.9632

A Figura 4.1 mostra a evolução da redução da função objetivo para o problema 43 aumentado. Pode-se observar que a solução ótima deste problema foi obtida após 14 iterações.

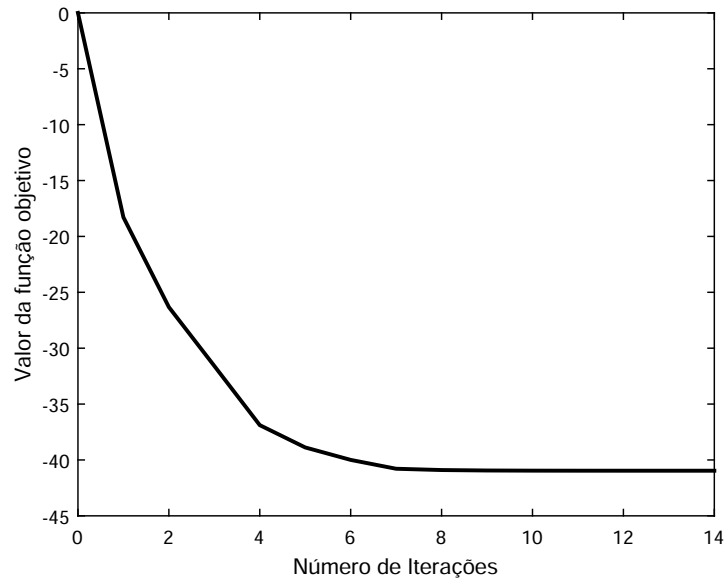


Figura 4.1: Evolução da redução da função objetivo do problema 43 aumentado.

A Tabela 4.5 mostra a comparação dos resultados ótimos obtidos entre os algoritmos PENLAB e FDIPA-GSDP.

Tabela 4.5: Comparação dos resultados obtidos.

Variável de projeto	PENLAB	FDIPA-GSDP
$x_1$	0.0000	0.0000
$x_2$	1.0384	1.0384
$x_3$	2.2271	2.2271
$x_4$	0.0000	0.0000
$f$	-40.9633	-40.9632
$NI$	8	14
$NF$	29	37

Pode-se observar que o esforço computacional e os resultados obtidos são muitos parecidos entre ambos os algoritmos.

As restrições ativas na solução ótima foram: a terceira restrição de desigualdade e os dois maiores autovalores da matriz semidefinida.

### 4.3.2 Exemplo 2 - Problema 71 aumentado

Neste segundo exemplo foi incluída uma restrição de igualdade. Este problema é formulado como

$$\left\{ \begin{array}{l} \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimizar}} \quad x_1 x_4 (x_1 + x_2 + x_3) + x_3 \\ \text{sujeito a:} \quad 25 - x_1 x_2 x_3 x_4 \leq 0 \\ \quad \quad \quad x_1^2 + x_2^2 + x_3^2 + x_4^2 - x_6 - 40 = 0 \\ \quad \quad \quad 1 \leq x_i \leq 5, \quad i = 1, \dots, 5, \quad x_i \geq 0, \quad i = 5, 6 \\ \quad \quad \quad \begin{bmatrix} -x_1 & -x_2 & 0 & 0 \\ -x_2 & -x_4 & -x_2 - x_3 & 0 \\ 0 & -x_2 - x_3 & -x_4 & -x_3 \\ 0 & 0 & -x_3 & -x_1 \end{bmatrix} \preceq \mathbf{0} \end{array} \right. \quad (4.3)$$

O número de variáveis de projeto e o número de restrições são mostradas na Tabela 4.6.

Tabela 4.6: Número de variáveis e de restrições do problema 71 aumentado.

$n$	$m$	$p$	$q$	$nt$	$nc$
6	1	1	4	2	6

A Tabela 4.7 mostra os valores inicial e final da função objetivo e das variáveis de projeto para o problema 71 aumentado.

Tabela 4.7: Solução do problema 71 aumentado.

Variável de projeto	Inicial	Final
$x_1$	4.9600	2.7586
$x_2$	1.0400	2.5278
$x_3$	1.0400	1.0000
$x_4$	4.9600	4.9999
$x_5$	0.0010	1.0164
$x_6$	0.0010	0.0000
$f$	172.2350	87.7105
$\lambda_1$	-2.4490	$-2.6700 \times 10^{-8}$

A Figura 4.2 mostra a evolução da redução da função objetivo para o Exemplo 2. Pode-se observar que a solução ótima foi obtida após 19 iterações.

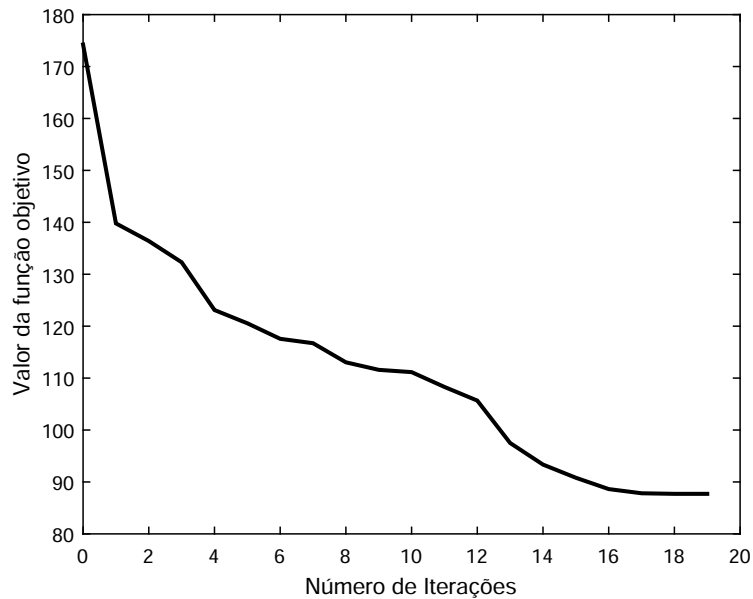


Figura 4.2: Evolução da redução da função objetivo do problema 71 aumentado.

As restrições ativas na solução ótima foram: os valores mínimos das variáveis 3 e 6, a restrição de igualdade e o maior autovalor da matriz semidefinida.

A Tabela 4.8 mostra a comparação dos resultados ótimos obtidos entre os algoritmos PENLAB e FDIPA-GSDP.

Tabela 4.8: Comparação dos resultados obtidos.

Variável de projeto	PENLAB	FDIPA-GSDP
$x_1$	2.7586	2.7586
$x_2$	2.5278	2.5278
$x_3$	1.0000	1.0000
$x_4$	5.0000	4.9999
$x_5$	9.8668	1.0164
$x_6$	0.0000	0.0000
$f$	87.7105	87.7105
$NI$	13	19
$NF$	123	38

Pode-se observar que os resultados obtidos por ambos os algoritmos são muito parecidos. Sem embargo o esforço computacional do algoritmo PENLAB foi muito maior devido a que neste algoritmo a restrição de igualdade é formulada como duas restrições de desigualdade.

## 4.4 Aplicações em problemas de otimização estrutural dimensional

Um problema de otimização estrutural dimensional para minimizar o peso de uma treliça sujeita a restrições estáticas (tensões e deslocamentos) e dinâmicas (frequências naturais) é formulado como

$$\left\{ \begin{array}{l} \underset{\mathbf{b} \in \mathbb{R}^m}{\text{minimizar}} \quad W(\mathbf{b}) = \sum_{i=1}^{ne} \rho_i b_i L_i \\ \text{sujeito a:} \quad \sigma^- \leq \sigma_i(\mathbf{b}) \leq \sigma^+, \quad i = 1, \dots, ne \\ \quad \sigma_{bi}(\mathbf{b}) \leq \sigma_i(\mathbf{b}), \quad i = 1, \dots, ne \\ \quad u^- \leq u_i(\mathbf{b}) \leq u^+, \quad i = 1, \dots, ng \\ \quad b_{\min} \leq b_i \leq b_{\max}, \quad i = 1, \dots, ne \\ \quad (\lambda^-)\mathbf{M}(\mathbf{b}) - \mathbf{K}(\mathbf{b}) \preceq \mathbf{0} \end{array} \right. \quad (4.4)$$

onde  $W$  é o peso total da estrutura,  $ne$  é o número de elementos ou número de barras da estrutura,  $ng$  é o número de graus de liberdade reduzidos,  $\rho_i$  é o peso específico da barra  $i$ ,  $b_i$  é a área da seção transversal da barra  $i$ ,  $L_i$  é o comprimento da barra  $i$ ,  $\sigma_i$  é a tensão da barra  $i$ ,  $\sigma^+$  e  $\sigma^-$  são as tensões admissíveis de tração e compressão da barra  $i$ , respectivamente,  $u_i$  é o deslocamento nodal ao longo do grau de liberdade  $i$ ,  $u^-$  e  $u^+$  são os limites inferior e superior do deslocamento nodal  $u_i$ , respectivamente,  $\mathbf{K}(\mathbf{b}) \in \mathcal{S}^{ng}$  é a matriz de rigidez reduzida,  $\mathbf{M}(\mathbf{b}) \in \mathcal{S}^{ng}$  é a matriz de massa reduzida,  $\lambda^-$  é o limite inferior da primeira frequência natural e  $b_{\min}$  e  $b_{\max}$  são os limites inferior e superior da área da seção transversal  $b_i$ , respectivamente.

A formulação SDP da restrição matricial imposta no problema (4.4) foi proposta por Achtziger e Kočvara [25] onde mudaram a formulação tradicional da restrição do mínimo autovalor do sistema

$$\mathbf{K}(\mathbf{b})\phi = \lambda\mathbf{M}(\mathbf{b})\phi \quad (4.5)$$

e provaram a seguinte equivalência entre ambas formulações dada pela seguinte expressão

$$\lambda_{\min} \geq \lambda^- \iff \mathbf{K}(\mathbf{b}) - (\lambda^-)\mathbf{M}(\mathbf{b}) \succeq \mathbf{0} \quad (4.6)$$

Uma grande dificuldade na formulação tradicional está na restrição vetorial  $\lambda_{\min} \geq \lambda^-$  já que esta função não é derivável em  $\mathbf{x}$  quando o mínimo autovalor tem multiplicidade maior que um. Nesse caso, é necessário empregar as técnicas de otimização não diferenciável para resolver o problema de otimização. Sem embargo, isto não acontece com a restrição matricial devido a que o cálculo da derivada

envolve a derivada das matrizes de rigidez e de massa em relação às variáveis de projeto convertendo-se em uma vantagem da formulação SDP em comparação com a formulação tradicional.

#### 4.4.1 Normalização das restrições

Em problemas de otimização estrutural trabalham-se com restrições com valores muito grandes ou muito pequenos portanto é importante normalizar estas restrições, tais como: deslocamentos nodais e tensões.

- **Deslocamentos nodais**

Os deslocamentos nodais são determinados através da Equação de Equilíbrio

$$\mathbf{K}(\mathbf{b})\mathbf{U}(\mathbf{b}) = \mathbf{F} \quad (4.7)$$

onde  $\mathbf{F} \in \mathbb{R}^{ng \times s}$  é a matriz dos vetores de carregamentos  $\mathbf{f}^l \in \mathbb{R}^{ng}$ , para  $l = 1, \dots, s$ ,  $\mathbf{U} \in \mathbb{R}^{ng \times s}$  é a matriz dos vetores de deslocamentos nodais  $\mathbf{u}^l \in \mathbb{R}^{ng}$ , para  $l = 1, \dots, s$ , e  $s$  é o número de casos de carregamento.

Desta maneira, as restrições de deslocamentos nodais podem ser definidas como

$$g_i(\mathbf{b}) = \frac{|u_i(\mathbf{b})|}{u_i^a} - 1 \leq 0, \quad \begin{cases} u_i^a = u_i^-, & \text{se } u_i \geq 0, \quad i = 1, \dots, ng \\ u_i^a = u_i^+, & \text{se } u_i < 0, \quad i = 1, \dots, ng \end{cases} \quad (4.8)$$

- **Tensões de tração e compressão**

As restrições de tensão são definidas como

$$g_i(\mathbf{b}) = \frac{|\sigma_i(\mathbf{b})|}{\sigma_i^a} - 1 \leq 0, \quad \begin{cases} \sigma_i^a = \sigma_i^+, & \text{se } \sigma_i \geq 0, \quad i = 1, \dots, ne \\ \sigma_i^a = \sigma_i^-, & \text{se } \sigma_i < 0, \quad i = 1, \dots, ne \end{cases} \quad (4.9)$$

- **Tensão de Flambagem**

A tensão de flambagem de Euler é formulada como

$$\sigma_{bi}(\mathbf{b}) = -\frac{k_e E_e b_e}{L_e^2} \quad (4.10)$$

onde  $k_e$  representa o coeficiente de flambagem de Euler. Assim, as restrições de flambagem são definidas como

$$g_i(\mathbf{b}) = \left| \frac{\sigma_i(\mathbf{b})}{\sigma_{bi}(\mathbf{b})} \right| - 1 \leq 0, \quad \sigma_i(\mathbf{b}) < 0, \quad i = 1, \dots, ne \quad (4.11)$$



#### 4.4.2 Análise de sensibilidade

Uma análise de sensibilidade corresponde ao cálculo dos gradientes das restrições e da função objetivo em relação às variáveis de projeto. Este procedimento foi realizado através do método de diferenciação direta.

- **Derivada da função objetivo**

A derivada da função objetivo em relação à  $j$ -ésima variável de projeto está dada por

$$\frac{\partial f}{\partial b_j} = \sum_{i \in G_j} \rho_i L_i \quad (4.12)$$

sendo  $G_j$  o grupo de união definido pela variável  $b_j$ .

- **Derivada das restrições de deslocamento nodal**

Derivando-se a Equação (4.7) em relação à  $j$ -ésima variável de projeto, obtém-se a seguinte expressão

$$\frac{\partial \mathbf{K}(\mathbf{b})}{\partial b_j} \mathbf{U}(\mathbf{b}) + \mathbf{K}(\mathbf{b}) \frac{\partial \mathbf{U}(\mathbf{x})}{\partial b_j} = \mathbf{0} \quad (4.13)$$

onde foi considerado que os carregamentos são independentes das variáveis de projeto.

Logo, tem-se

$$\frac{\partial \mathbf{U}(\mathbf{x})}{\partial b_j} = -\mathbf{K}(\mathbf{x})^{-1} \frac{\partial \mathbf{K}(\mathbf{b})}{\partial b_j} \mathbf{U}(\mathbf{b}) \quad (4.14)$$

Neste caso, a derivada da matriz de rigidez reduzida está dada por

$$\frac{\partial \mathbf{K}(\mathbf{b})}{\partial b_j} = \sum_{j=1}^n b_j \sum_{i \in G_j} \frac{E_i}{L_i} \mathbf{D}_i \quad (4.15)$$

sendo  $n$  o número de variáveis de projeto e  $\mathbf{D}_i$  a matriz dos cossenos diretores relativos à barra  $i$  definida no Apêndice C.

Para determinar a derivada da restrição matricial é necessário determinar também a derivada da matriz de Massa reduzida dada por

$$\frac{\partial \mathbf{M}(\mathbf{b})}{\partial b_j} = \sum_{j=1}^n b_j \sum_{i \in G_j} \frac{\rho_i L_i}{6} \begin{bmatrix} 2\mathbf{I}_t & \mathbf{I}_t \\ \mathbf{I}_t & 2\mathbf{I}_t \end{bmatrix} \quad (4.16)$$

onde o subíndice  $t$  pode ser 2 ou 3, se trata-se de uma treliça bidimensional ou tridimensional, respectivamente.

Portanto, a derivada da restrição de deslocamento é dada por

$$\frac{\partial g_i(\mathbf{b})}{\partial b_j} = \frac{\text{sig}(u_i(\mathbf{b}))}{u_i^a} \frac{\partial u_j(\mathbf{x})}{\partial b_j} \quad (4.17)$$

sendo  $\text{sig}(\cdot) = (\cdot)/|(\cdot)|$ .

- **Derivada das tensões de tração e compressão**

A tensão axial de uma barra  $i$  está dada por

$$\sigma_i(\mathbf{b}) = \frac{E_i}{L_i} [-1 \ 1] \mathbf{T}_i \mathbf{u}_i(\mathbf{x}) \quad (4.18)$$

sendo  $\mathbf{u}_i$  o vetor deslocamento nodal da barra  $i$  e  $\mathbf{T}_i$  a matriz de transformação relativa à barra  $i$ .

Derivando a Equação (4.18) em relação à  $j$ -ésima variável de projeto, obtém-se a seguinte expressão

$$\frac{\partial \sigma_i(\mathbf{b})}{\partial b_j} = \frac{E_i}{L_i} [-1 \ 1] \mathbf{T}_i \nabla_{b_j} \mathbf{u}_i(\mathbf{x}) \quad (4.19)$$

Portanto, a derivada da restrição de tensão está dada por

$$\frac{\partial g_i(\mathbf{b})}{\partial b_j} = \frac{\text{sig}(\sigma_i(\mathbf{b}))}{\sigma_i^a} \frac{\partial \sigma_i(\mathbf{b})}{\partial b_j} \quad (4.20)$$

- **Derivada das tensões de flambagem**

Derivando a Equação (4.10) em relação à  $j$ -ésima variável de projeto, obtém-se

$$\frac{\partial \sigma_{bi}(\mathbf{b})}{\partial b_j} = -\frac{k_e E_e}{L_e^2} \quad (4.21)$$

Portanto, a derivada da restrição de flambagem está dado por

$$\frac{\partial g_i(\mathbf{b})}{\partial b_j} = \frac{\sigma_{bi}(\mathbf{b}) \frac{\partial \sigma_i(\mathbf{b})}{\partial b_j} - \sigma_i(\mathbf{b}) \frac{\partial \sigma_{bi}(\mathbf{b})}{\partial b_j}}{\sigma_{bi}^2(\mathbf{b})} \quad (4.22)$$

### 4.4.3 Exemplo 3 - Treliça bidimensional de 10 barras

Trata-se de um exemplo de otimização de uma estrutura de treliças formada por 10 barras, cuja geometria é mostrada na Figura 4.3, onde procura-se minimizar o peso total da estrutura com restrições de tensão, de deslocamento nodal e da mínima frequência natural.

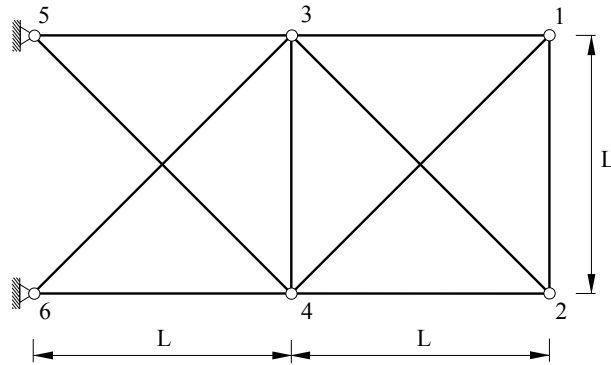


Figura 4.3: Geometria da treliça de 10 barras:  $L = 360$  in (914.4 cm).

Esta estrutura está submetida a um único caso de carregamento descrito na Tabela 4.9 e as condições de contorno estabelecem que os nós 5 e 6 estão fixos nas direções  $X$  e  $Y$ .

Tabela 4.9: Casos de carregamento. Componentes da força em Kips (KN).

Caso	Nó	$F_y$
1	2	-100 (-445)
	4	-100 (-445)

Cada uma das áreas da seção transversal das barras são consideradas variáveis independentes, gerando-se desta maneira dez variáveis de projeto.

O valor inicial de cada área foi  $30 \text{ in}^2$  ( $193.548 \text{ cm}^2$ ), sendo este um ponto inviável. Para obter um ponto inicial viável foram necessárias nove iterações auxiliares obtendo-se assim um peso inicial de  $34578.300 \text{ lb}$  ( $15684.453 \text{ Kgf}$ ). O limite inferior para todas as áreas das barras é  $0.1 \text{ in}^2$  ( $0.6452 \text{ cm}^2$ ).

A lista de variáveis de projeto, a definição dos elementos da treliça e a conectividade dos nós são descritos na Tabela 4.10.

Tabela 4.10: Definição dos elementos e conectividade dos nós.

Variável de projeto	Elemento	Conectividade dos nós	Variável de projeto	Elemento	Conectividade dos nós
1	1	5-3	6	6	1-2
2	2	3-1	7	7	5-4
3	3	6-4	8	8	6-3
4	4	4-2	9	9	3-2
5	5	3-4	10	10	4-1

As tensões admissíveis de tração e de compressão para todas as barras são  $\pm 25$  Ksi ( $\pm 172.375 \text{ MPa}$ ), respectivamente e os deslocamentos dos nós 1, 2, 3 e 4 estão limitados a  $\pm 2$  in ( $\pm 5.08 \text{ cm}$ ) na direção  $Y$ . O valor de  $\lambda^-$  é  $23700 \text{ rad}^2/\text{s}^2$  ( $\omega_1 = 24.5 \text{ Hz}$ ).

O número de variáveis de projeto e o número de restrições vetoriais e matriciais impostas neste problema são mostradas na Tabela 4.11.

Tabela 4.11: Número de variáveis e de restrições da treliça de 10 barras.

$n$	$m$	$p$	$q$	$nt$	$nc$
10	28	0	8	28	10

A Figura 4.4 mostra as configurações inicial e final da estrutura para este problema.

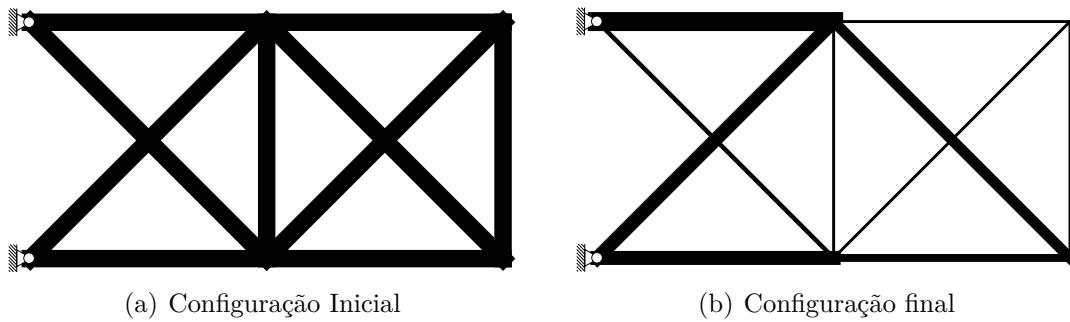


Figura 4.4: Configurações inicial e final da treliça de 10 barras.

A Tabela 4.12 mostra a solução ótima e o peso final obtidos pelo algoritmo. Pode-se observar que o peso final da estrutura foi 5111.47 lb (2318.52 Kgf) que representa uma redução de peso de 59%.

Tabela 4.12: Estrutura ótima obtida para a treliça de 10 barras.

Variável de projeto	Configuração Inicial	Configuração Final
$b_1$	30.000	32.648
$b_2$	30.000	0.100
$b_3$	30.000	24.100
$b_4$	30.000	13.631
$b_5$	30.000	0.100
$b_6$	30.000	0.281
$b_7$	30.000	7.821
$b_8$	30.000	23.720
$b_9$	30.000	18.652
$b_{10}$	30.000	0.100
Peso (lb)	12589.400	5111.470
$\omega_1$ (Hz)	15.830	24.500

A Figura 4.5 mostra a evolução da redução do peso da estrutura para este problema. Foram necessárias 37 iterações para encontrar a solução ótima e 43 avaliações da função objetivo.

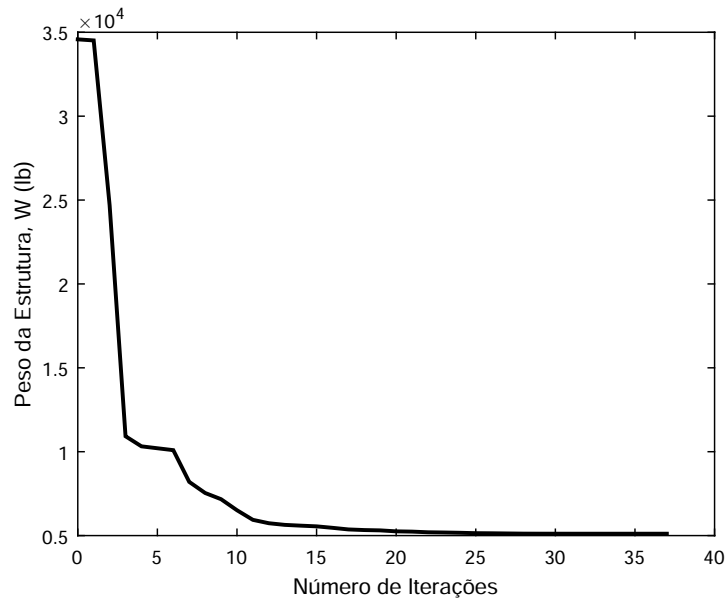


Figura 4.5: Evolução da redução do peso da treliça de 10 barras.

As restrições ativas na configuração final foram: a área mínima das barras 2, 5 e 10, o deslocamento do nó 1 na direção  $-Y$ , a tensão de tração da barra 5 e a primeira frequência natural.

#### 4.4.4 Exemplo 4 - Treliça espacial de 25 barras

Neste exemplo, trata-se de otimizar o peso de uma treliça espacial de 25 barras (torre de transmissão), cuja geometria é mostrada na Figura 4.6. As restrições impostas neste problema são as tensões de tração e de compressão, as tensões de flambagem, os deslocamentos nodais e a mínima frequência natural.

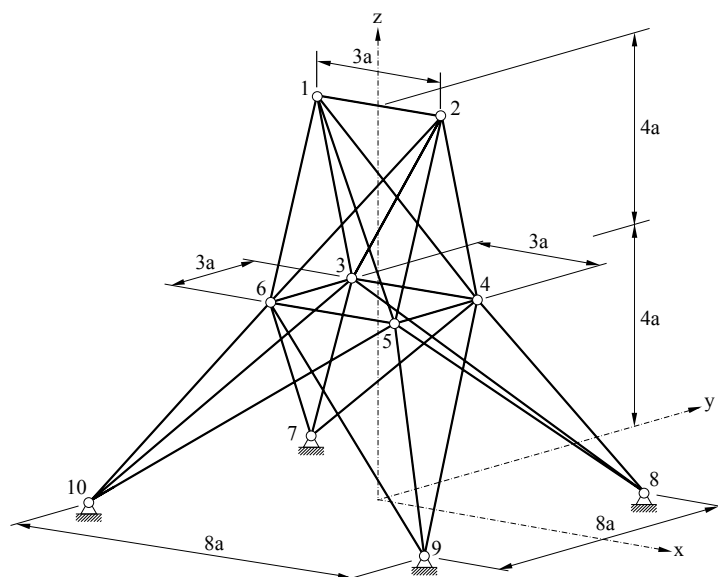


Figura 4.6: Geometria da treliça de 25 barras:  $a = 25$  in (63.5 cm).

A treliça está submetida a dois casos de carregamento descritos na Tabela 4.13 e as condições de borde estabelecem que os nós 7, 8, 9 e 10 estão fixos.

Tabela 4.13: Casos de carregamento. Componentes da força em Kips (KN).

Caso	Nó	$F_x$	$F_y$	$F_z$
1	1	1 (4.45)	10 (44.5)	-5 (-22.25)
	2	0	10 (44.5)	-5 (-22.25)
	3	0.5 (2.225)	0	0
	6	0.5 (2.225)	0	0
2	1	0	20 (89)	-5 (-22.25)
	2	0	-20 (-89)	-5 (-22.25)

Com a finalidade de manter a simetria, as áreas das barras são distribuídas em oito grupos da seguinte maneira:  $(G_1) b_1$ ,  $(G_2) b_2 - b_5$ ,  $(G_3) b_6 - b_9$ ,  $(G_4) b_{10} - b_{11}$ ,  $(G_5) b_{12} - b_{13}$ ,  $(G_6) b_{14} - b_{17}$ ,  $(G_7) b_{18} - b_{21}$  e  $(G_8) b_{22} - b_{25}$ , gerando-se desta maneira oito variáveis de projeto.

O valor inicial de todas as variáveis de projeto foi  $3 \text{ in}^2$  ( $19.356 \text{ cm}^2$ ), sendo este um ponto inviável. Para obter um ponto inicial viável foram necessárias seis iterações auxiliares obtendo-se assim um peso inicial de  $4134.120 \text{ lb}$  ( $1875.205 \text{ Kgf}$ ).

A lista das variáveis de projeto, a definição dos elementos da treliça e a conectividade dos nós são descritos na Tabela 4.14.

Tabela 4.14: Definição dos elementos e conectividade dos nós.

Variável de projeto	Elemento	Conectividade dos nós
1	1	1-2
2	2-5	1-4, 2-3, 1-5, 2-6
3	6-9	2-4, 2-5, 1-3, 1-6
4	10,11	3-6, 4-5
5	12,13	3-4, 5-6
6	14-17	3-10, 6-7, 4-9, 5-8
7	18-21	4-7, 3-8, 5-10, 6-9
8	22-25	6-10, 3-7, 4-8, 5-9

As tensões admissíveis de tensão e de compressão são  $\pm 40 \text{ Ksi}$  ( $\pm 275.800 \text{ MPa}$ ), respectivamente, para todas as barras e os deslocamentos dos nós 1, 2, 3 e 4 estão limitados a  $\pm 0.35 \text{ in}$  ( $\pm 0.889 \text{ cm}$ ) em todas as direções. O limite inferior para todas as áreas das barras é  $0.01 \text{ in}^2$  ( $0.06452 \text{ cm}^2$ ). O valor de  $\lambda^-$  é  $272000 \text{ rad}^2/\text{s}^2$  ( $\omega_1 = 83 \text{ Hz}$ ) e o valor do coeficiente de flambagem é  $39.274$ .

O número de variáveis de projeto e o número de restrições vetoriais e matriciais impostas neste problema são mostradas na Tabela 4.15.

Tabela 4.15: Número de variáveis e de restrições da treliça de 25 barras.

$n$	$m$	$p$	$q$	$nt$	$nc$
8	174	0	18	174	8

A Figura 4.7 mostra as configurações inicial e final da estrutura para este problema.

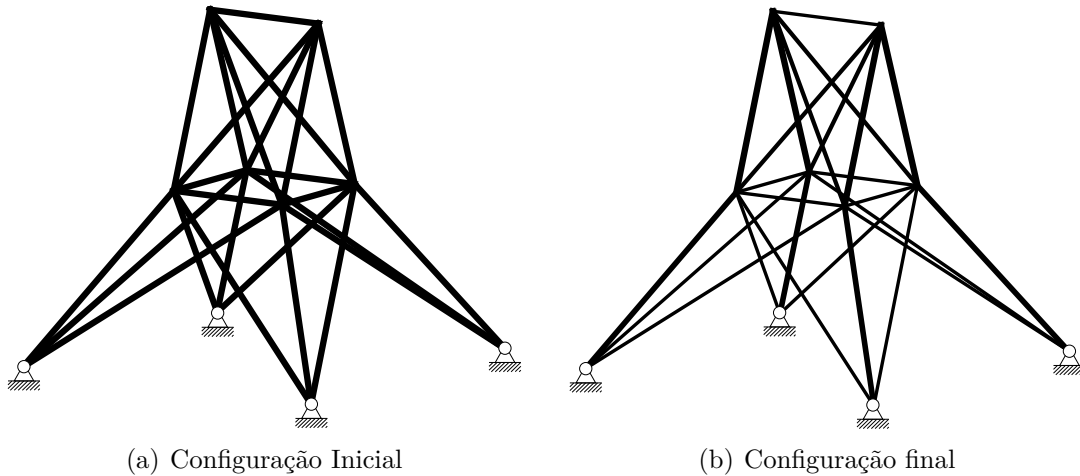


Figura 4.7: Configurações inicial e final da treliça de 25 barras.

A Tabela 4.16 mostra a solução ótima e o peso final obtidos pelo algoritmo. Pode-se observar que o peso final foi 630.192 lb (285.850 Kgf) que representa uma redução de peso de 36%.

Tabela 4.16: Estrutura ótima da treliça de 25 barras.

Variável de projeto	Configuração Inicial	Configuração Final
$b_1$	3.000	0.010
$b_2$	3.000	2.139
$b_3$	3.000	2.264
$b_4$	3.000	0.550
$b_5$	3.000	0.038
$b_6$	3.000	0.787
$b_7$	3.000	2.097
$b_8$	3.000	3.821
Peso (lb)	992.162	630.192
$\omega_1$ (Hz)	68.990	83.000

As restrições ativas na configuração final foram: a área mínima da barra 1, a tensão de flambagem da barra 12 correspondente ao primeiro caso de carregamento, os deslocamentos dos nós 1 e 2 nas direções  $+Y$  e  $-Y$ , respectivamente, correspondentes ao segundo caso de carregamento e a primeira frequência natural.

A Figura 4.8 mostra a evolução da redução do peso para este problema. Foram necessárias 31 iterações para encontrar a solução ótima e 43 avaliações da função objetivo.

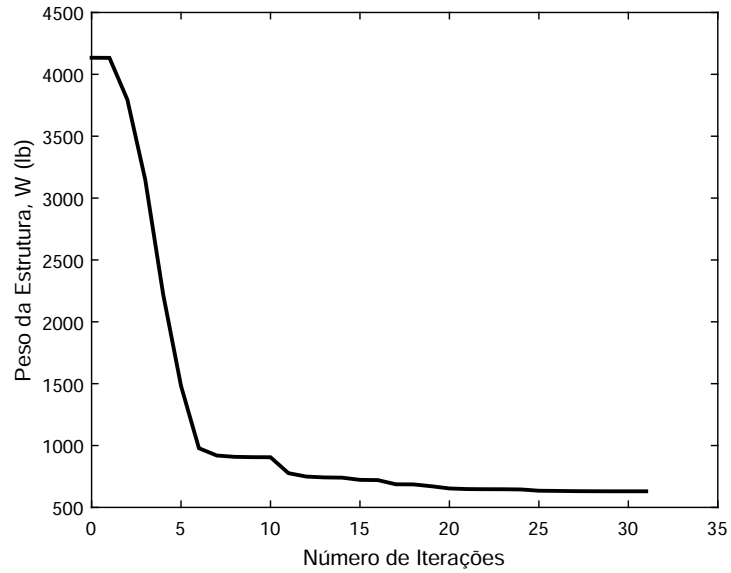


Figura 4.8: Evolução da redução do peso da treliça de 25 barras.

#### 4.4.5 Exemplo 5 - Treliça espacial de 72 barras

Neste exemplo, trata-se de otimizar o peso de uma treliça espacial de 72 barras, cuja geometria é mostrada na Figura 4.9.

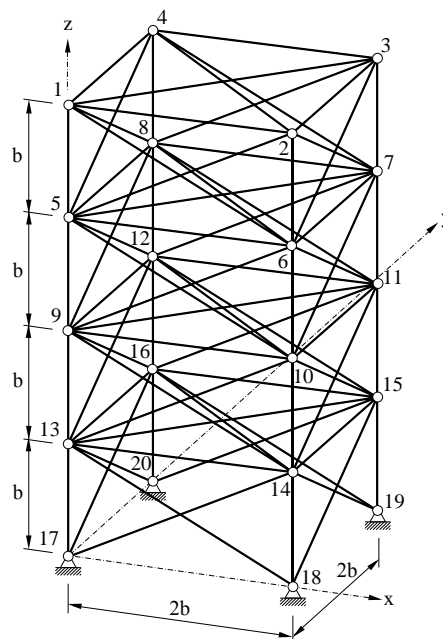


Figura 4.9: Geometria da treliça de 72 barras:  $b = 60$  in (152.4 cm).



As restrições impostas neste problema são de tensão, de deslocamento nodal e da mínima frequência natural. Aqui não foram consideradas as restrições de flambagem.

A treliça está submetida a dois casos de carregamento que são descritos na Tabela 4.17 e as condições de contorno estabelecem que os nós 17, 18, 19 e 20 estão fixos nas direções  $X$ ,  $Y$  e  $Z$ .

Tabela 4.17: Casos de carregamento. Componentes da força em Kips (KN).

Caso	Nó	$F_x$	$F_y$	$F_z$
1	1	5 (22.25)	5 (22.25)	-5 (-22.25)
2	1	0	0	-5 (-22.25)
	2	0	0	-5 (-22.25)
	3	0	0	-5 (-22.25)
	4	0	0	-5 (-22.25)

Com a finalidade de manter a simetria, as áreas das barras são distribuídas em 16 grupos da seguinte maneira: ( $G_1$ )  $b_1 - b_4$ , ( $G_2$ )  $b_5 - b_{12}$ , ( $G_3$ )  $b_{13} - b_{16}$ , ( $G_4$ )  $b_{17} - b_{18}$ , ( $G_5$ )  $b_{19} - b_{22}$ , ( $G_6$ )  $b_{23} - b_{30}$ , ( $G_7$ )  $b_{31} - b_{34}$ , ( $G_8$ )  $b_{35} - b_{36}$ , ( $G_9$ )  $b_{37} - b_{40}$ , ( $G_{10}$ )  $b_{41} - b_{48}$ , ( $G_{11}$ )  $b_{49} - b_{52}$ , ( $G_{12}$ )  $b_{53} - b_{54}$ , ( $G_{13}$ )  $b_{55} - b_{58}$ , ( $G_{14}$ )  $b_{59} - b_{66}$ , ( $G_{15}$ )  $b_{67} - b_{70}$ , ( $G_{16}$ )  $b_{71} - b_{72}$ , gerando-se desta maneira 16 variáveis de projeto.

A lista das variáveis de projeto, a definição dos elementos da treliça e a conectividade dos nós são descritos na Tabela 4.18.

Tabela 4.18: Definição dos elementos e conectividade dos nós.

Variável de projeto	Elemento	Conectividade dos nós
1	1-4	1-5, 2-6, 3-7, 4-8
2	5-12	2-5, 1-6, 3-6, 2-7, 4-7, 3-8, 1-8, 4-5
3	13-16	1-2, 2-3, 3-4, 4-1
4	17,18	1-3, 2-4
5	19-22	5-9, 6-10, 7-11, 8-12
6	23-30	6-9, 5-10, 7-10, 6-11, 8-1, 7-12, 5-12, 8-9
7	31-34	5-6, 6-7, 7-8, 8-5
8	35,36	5-7, 6-8
9	37-40	9-13, 10-14, 11-15, 12-16
10	41-48	10-13, 9-14, 11-14, 10-15, 12-15, 11-16, 9-16, 12-13
11	49-52	9-10, 10-11, 11-12, 12-9
12	53,54	9-11, 10-12
13	55-58	13-17, 14-18, 15-19, 16-20
14	59-66	14-17, 13-18, 15-18, 14-19, 16-19, 15-20, 13-20, 16-17
15	67-70	13-14, 14-15, 15-16, 16-13
16	71,72	13-15, 14-16

O valor inicial de todas as variáveis de projeto foi  $1 \text{ in}^2$  ( $6.452 \text{ cm}^2$ ), sendo este um ponto inviável. Para obter um ponto inicial viável foram necessárias 11 iterações auxiliares obtendo-se assim um peso inicial de  $2770.510 \text{ lb}$  ( $1256.682 \text{ Kgf}$ ).

As tensões admissíveis de tensão e de compressão são  $\pm 25 \text{ Ksi}$  ( $\pm 172.375 \text{ MPa}$ ), respectivamente e os deslocamentos dos nós 1, 2, 3 e 4 estão limitados a  $\pm 0.25 \text{ in}$  ( $\pm 0.635 \text{ cm}$ ) em todas as direções. O limite inferior para todas as áreas das seções transversais é  $0.1 \text{ in}^2$  ( $0.6452 \text{ cm}^2$ ) e o valor de  $\lambda^-$  é  $83536 \text{ rad}^2/\text{s}^2$  ( $\omega_1 = 46 \text{ Hz}$ ).

O número de variáveis de projeto e o número de restrições vetoriais e matriciais impostas neste problema são mostradas na Tabela 4.19.

Tabela 4.19: Número de variáveis e de restrições da treliça de 72 barras.

$n$	$m$	$p$	$q$	$nt$	$nc$
16	336	0	48	336	16

A Figura 4.10 mostra as configurações inicial e final da estrutura para este problema.

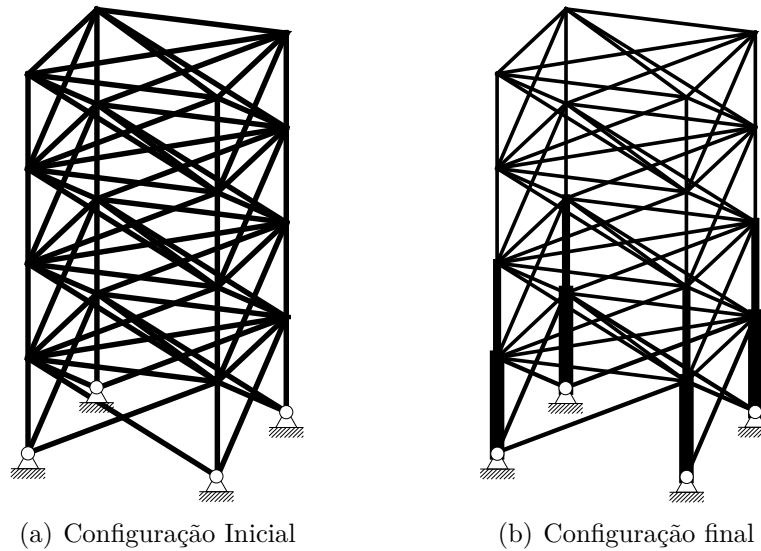


Figura 4.10: Configurações inicial e final da treliça de 72 barras.

A Tabela 4.20 mostram a solução ótima e o peso final obtidos pelo algoritmo. Pode-se observar que o peso final da estrutura foi  $419.019 \text{ lb}$  ( $190,064 \text{ Kgf}$ ) que representa uma redução de peso de 51%.

As restrições ativas na configuração final foram: as áreas mínimas dos grupos de barras 7, 8, 11, 12, 15 e 16, os deslocamentos do nó 1 nas direções  $+X$  e  $+Y$  correspondentes ao primeiro caso de carregamento, as tensões de compressão das barras 1, 2, 3 e 4 correspondentes ao segundo caso de carregamento e as duas primeiras frequências naturais. Neste exemplo ficaram ativas as duas primeiras frequências naturais devido à simetria da geometria da treliça.

Tabela 4.20: Estrutura ótima da treliça de 72 barras.

Variável de projeto	Configuração Inicial	Configuração Final
$b_1$	1.000	0.166
$b_2$	1.000	0.387
$b_3$	1.000	0.235
$b_4$	1.000	0.340
$b_5$	1.000	0.581
$b_6$	1.000	0.500
$b_7$	1.000	0.100
$b_8$	1.000	0.100
$b_9$	1.000	1.493
$b_{10}$	1.000	0.675
$b_{11}$	1.000	0.100
$b_{12}$	1.000	0.100
$b_{13}$	1.000	2.668
$b_{14}$	1.000	0.802
$b_{15}$	1.000	0.100
$b_{16}$	1.000	0.100
Peso (lb)	853.100	419.019
$\omega_1$ (Hz)	26.070	46.000

A Figura 4.11 mostra a evolução da redução do peso para este problema. Foram necessárias 37 iterações para encontrar a solução ótima e 39 avaliações da função objetivo.

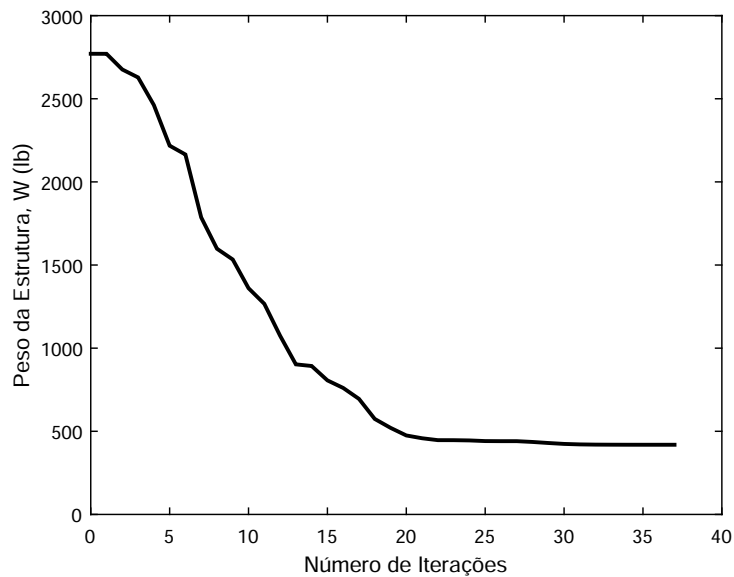


Figura 4.11: Evolução da redução do peso da treliça de 72 barras.

## 4.5 Aplicações em problemas de otimização estrutural dimensional e geométrica

Um problema de otimização estrutural dimensional e geométrica para minimizar o peso de uma treliça sujeita a restrições estáticas (tensões) e dinâmicas (frequências naturais) pode ser formulado da seguinte maneira

$$\left\{ \begin{array}{l} \underset{\mathbf{b} \in \mathbb{R}^{ne}}{\text{minimizar}} \quad W(\mathbf{b}, \mathbf{X}) = \sum_{i=1}^{ne} \rho_i b_i L_i(\mathbf{X}) \\ \text{sujeito a:} \quad \sigma_i^- \leq \sigma_i(\mathbf{b}, \mathbf{X}) \leq \sigma_i^+, \quad i = 1, \dots, ne \\ \quad \sigma_{ib}(\mathbf{b}, \mathbf{X}) \leq \sigma_i(\mathbf{b}, \mathbf{X}), \quad i = 1, \dots, ne \\ \quad (\lambda^-) \mathbf{M}(\mathbf{b}, \mathbf{X}) - \mathbf{K}(\mathbf{b}, \mathbf{X}) \preceq \mathbf{0} \\ \quad b_{\min} \leq b_i \leq b_{\max}, \quad i = 1, \dots, ne \\ \quad X_j^- \leq X_j \leq X_j^+, \quad j = 1, \dots, k \end{array} \right. \quad (4.23)$$

onde  $X_j^-$  e  $X_j^+$  representam os limites inferior e superior da coordenada  $X_j$  do nó  $j$ , respectivamente e o comprimento da barra  $i$  está dado por

$$L_i(\mathbf{X}) = \sqrt{(X_2 - X_1)_i^2 + (Y_2 - Y_1)_i^2 + (Z_2 - Z_1)_i^2} \quad (4.24)$$

Neste problema, as variáveis de projeto são as áreas das barras e também as coordenadas nodais e, portanto, a função objetivo e as restrições de desigualdade são funções não lineares das coordenadas nodais assim como também as restrições matriciais.

### 4.5.1 Exemplo 6 - Treliça espacial de 25 barras

Neste exemplo, trata-se de uma treliça espacial de 25 barras onde a geometria, a definição dos elementos, os casos de carregamento, as condições de contorno, as tensões admissíveis e os valores iniciais das áreas das barras são os mesmos do que o Exemplo 4. Sem embargo não foram consideradas as restrições nos deslocamentos nodais.

Com a finalidade de manter novamente a simetria da estrutura foram consideradas cinco variáveis geométricas adicionais correspondentes às coordenadas nodais  $X_4, Y_4, Z_4, X_8$  e  $Y_8$  distribuídas da seguinte maneira:  $X_4 = X_5 = -X_3 = -X_6, Y_3 = Y_4 = -Y_5 = -Y_6, Z_3 = Z_4 = Z_5 = Z_6, X_8 = X_9 = -X_7 = -X_{10}$  e  $Y_7 = Y_8 = -Y_9 = -Y_{10}$ , gerando-se desta maneira um total de 13 variáveis de projeto (oito variáveis dimensionais e cinco geométricas).

Os valores iniciais das coordenadas nodais são dadas na Tabela 4.21.

Tabela 4.21: Coordenadas nodais iniciais em in (cm).

Nó	$X$	$Y$	$Z$
1	-37.5 (-95.25)	0	200 (508)
2	37.5 (95.25)	0	200 (508)
3	-37.5 (-95.25)	37.5 (95.25)	100 (254)
4	37.5 (95.25)	37.5 (95.25)	100 (254)
5	37.5 (95.25)	-37.5 (-95.25)	100 (254)
6	-37.5 (95.25)	-37.5 (-95.25)	100 (254)
7	-100 (-254)	100 (254)	0
8	100 (254)	100 (254)	0
9	100 (254)	-100 (-254)	0
10	-100 (-254)	-100 (-254)	0

O limite inferior para todas as áreas das barras é  $0.01 \text{ in}^2$  ( $0.6452 \text{ cm}^2$ ) e o valor de  $\lambda^-$  é  $252662 \text{ rad}^2/\text{s}^2$  ( $\omega_1 = 80 \text{ Hz}$ ).

Os limites inferior e superior das variáveis geométricas, correspondentes às coordenadas nodais, são dados na Tabela 4.22.

Tabela 4.22: Limites das variáveis geométricas da treliça de 25 barras em in (cm).

Variável de projeto	Limite Inferior	Limite Superior
$X_4$	20 (50.8)	60 (152.4)
$Y_4$	40 (101.6)	80 (203.2)
$Z_4$	90 (228.6)	130 (330.2)
$X_8$	40 (101.6)	80 (203.2)
$Y_8$	100 (254.0)	140 (355.6)

O número de variáveis de projeto e o número de restrições vetoriais e matriciais impostas neste problema são mostradas na Tabela 4.23.

Tabela 4.23: Número de variáveis e de restrições da treliça de 25 barras.

$n$	$m$	$p$	$q$	$nt$	$nc$
13	150	0	18	150	18

A Figura 4.12 mostra as configurações inicial e final da estrutura para este problema.

A Tabela 4.24 mostra a solução ótima e o peso final obtidos pelo algoritmo. Pode-se observar que o peso final da estrutura foi 173.248 lb (78.584 Kgf) que representa uma redução de peso de 83%.

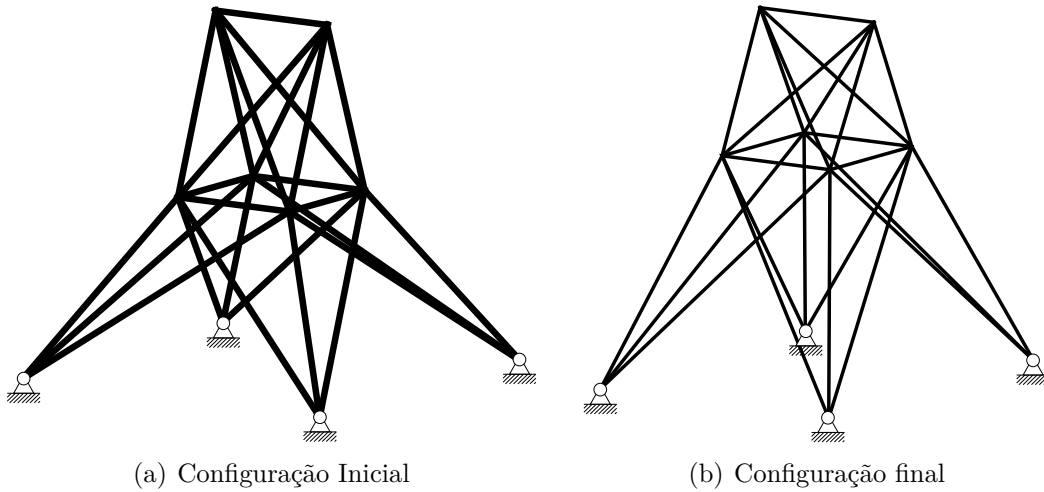


Figura 4.12: Configurações inicial e final da treliça de 25 barras.

Tabela 4.24: Estrutura ótima obtida para a treliça de 25 barras.

Variável de projeto	Configuração Inicial	Configuração Final
$b_1$	3.000	0.010
$b_2$	3.000	0.515
$b_3$	3.000	0.580
$b_4$	3.000	0.065
$b_5$	3.000	0.163
$b_6$	3.000	0.239
$b_7$	3.000	0.933
$b_8$	3.000	0.740
$X_4$	37.500	35.502
$Y_4$	37.500	40.000
$Z_4$	100.000	124.274
$X_8$	100.000	75.166
$Y_8$	100.000	100.000
Peso (lb)	992.162	173.248
$\omega_1$ (Hz)	68.990	80.000

As restrições ativas na configuração final foram as seguintes: a área mínima da barra 1, as tensões de flambagem das barras 13, 17 e 24 correspondentes ao primeiro caso de carregamento, as tensões de flambagem das barras 2, 5, 7, 8, 19 e 20 correspondentes ao segundo caso de carregamento, assim como a primeira frequência natural.

A Figura 4.13 mostra a evolução da redução do peso para este problema. Foram necessárias 44 iterações para encontrar a solução ótima e 49 avaliações da função objetivo.

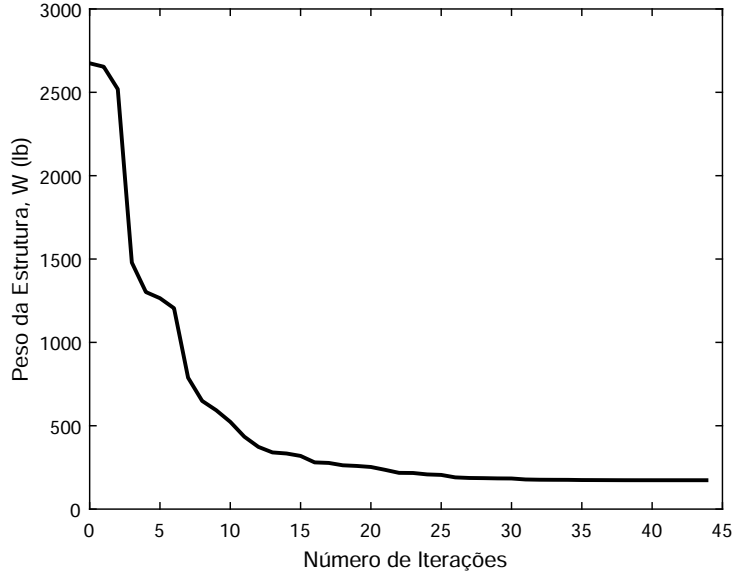


Figura 4.13: Evolução da redução do peso da treliça de 25 barras.

## 4.6 Aplicações em problemas de otimização estrutural topológica

Um problema de otimização estrutural topológica na formulação SDP para minimizar o peso de uma treliça sujeita a restrições na complacência estrutural e frequências naturais pode ser formulada da seguinte maneira

$$\left\{ \begin{array}{l} \text{minimizar}_{\mathbf{b} \in \mathbb{R}^{ne}} \quad W(\mathbf{b}) = \sum_{i=1}^{ne} \rho_i b_i L_i \\ \text{sujeito a:} \quad \begin{bmatrix} -\kappa^+ & \mathbf{f}^{lT} \\ \mathbf{f}^l & -\mathbf{K}(\mathbf{b}) \end{bmatrix} \preceq \mathbf{0}, \quad l = 1, \dots, s \\ \quad \quad \quad (\lambda^-) \mathbf{M}(\mathbf{b}) - \mathbf{K}(\mathbf{b}) \preceq \mathbf{0} \\ \quad \quad \quad b_i \geq 0, \quad i = 1, \dots, ne \end{array} \right. \quad (4.25)$$

sendo  $\kappa^+$  o limite superior da complacência estrutural.

Esta formulação foi proposta por Achtzinger e Kocvara [25] onde mudaram as restrições de igualdade, dadas pela equação de equilíbrio, e as restrições de desigualdade, dadas pela complacência estrutural, por uma matriz semidefinida e estabeleceram a seguinte equivalência

$$\left( \sum_{i=1}^m b_i \mathbf{K}_i \right) \mathbf{u}^l(\mathbf{b}) = \mathbf{f}^l \quad \Leftrightarrow \quad \begin{bmatrix} \kappa^+ & -\mathbf{f}^l \\ -\mathbf{f}^{lT} & \mathbf{K}(\mathbf{b}) \end{bmatrix} \succeq \mathbf{0} \quad (4.26)$$

$$\mathbf{f}^{lT} \mathbf{u}^l(\mathbf{b}) \leq \kappa^+$$

Neste problema as variáveis de projeto são as áreas das barras sem embargo algumas delas podem ser nulas na solução.

#### 4.6.1 Exemplo 7 - Treliça bidimensional de 26 barras

Neste exemplo, trata-se de otimizar o peso de uma treliça bidimensional de 26 barras, cuja geometria é mostrada na 4.14.

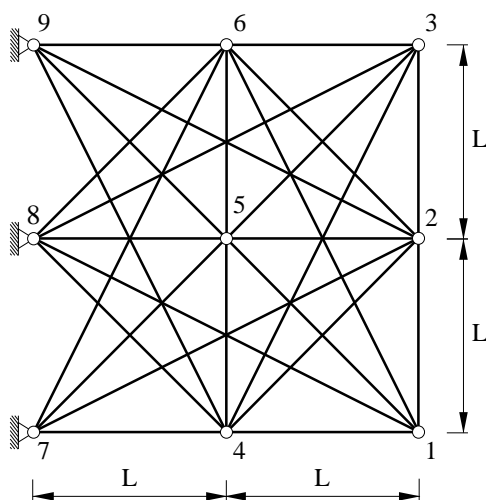


Figura 4.14: Geometria da treliça de 26 barras:  $L = 360$  in (914.4 cm).

A treliça está submetida a um único caso de carregamento descrito na Tabela 4.25 e as condições de contorno estabelecem que os nós 7, 8 e 9 estão fixos nas direções  $X$  e  $Y$ .

Tabela 4.25: Casos de carregamento. Componentes da força em Kips (KN).

Caso	Nó	$F_x$
1	2	-100 (-445)

Cada uma das áreas da seção transversal das barras são consideradas variáveis independentes, gerando-se desta maneira 26 variáveis de projeto. O valor inicial de cada área foi  $1$  in<sup>2</sup> (6.452 cm<sup>2</sup>), sendo este um ponto viável, obtendo-se assim um peso inicial de 1411.280 lb (640.146 Kgf).

O número de variáveis de projeto e o número de restrições vetoriais e matriciais impostas neste problema são mostradas na Tabela 4.26.

Tabela 4.26: Número de variáveis e de restrições da treliça de 26 barras.

$n$	$m$	$p$	$q$	$nt$	$nc$
26	0	0	25	0	26



A lista das variáveis de projeto, a definição dos elementos da treliça e a conectividade dos nós são descritos na Tabela 4.27.

Tabela 4.27: Definição dos elementos e conectividade dos nós.

Variável de projeto	Elemento	Conectividade dos nós	Variável de projeto	Elemento	Conectividade dos nós
1	1	7-4	14	14	1-5
2	2	4-1	15	15	8-6
3	3	4-5	16	16	5-9
4	4	1-2	17	17	5-3
5	5	8-5	18	18	2-6
6	6	5-2	19	19	7-2
7	7	5-6	20	20	7-6
8	8	2-3	21	21	4-9
9	9	9-6	22	22	4-3
10	10	6-3	23	23	1-8
11	11	7-5	24	24	1-6
12	12	4-8	25	25	8-3
13	13	4-2	26	26	2-9

Os valores de  $\kappa^+$  e  $\lambda^-$  são  $3.1 \times 10^5$  lbf-pulg (35025.28 N-m) e  $3948 \text{ rad}^2/\text{s}^2$  ( $\omega_1 = 10 \text{ Hz}$ ), respectivamente.

A Figura 4.15 mostra as configurações inicial e final da estrutura para este problema.

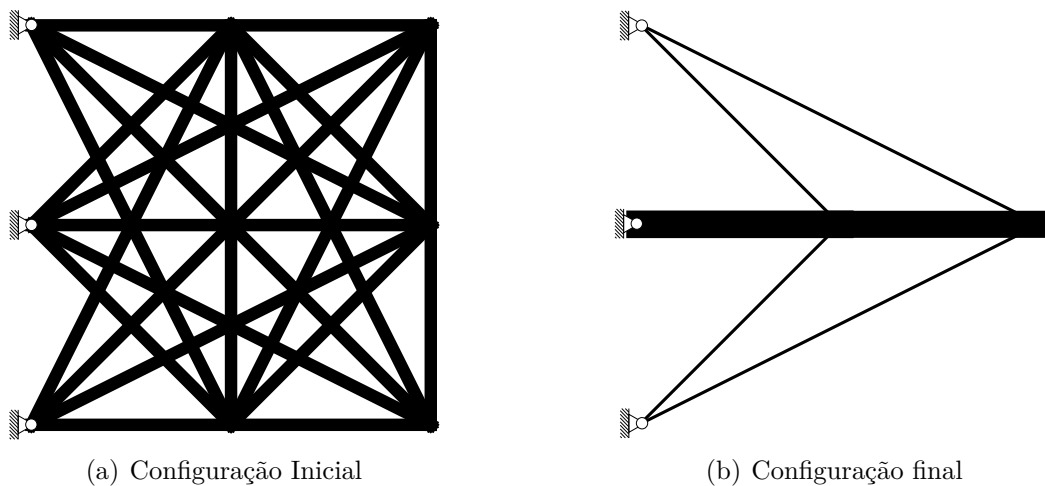


Figura 4.15: Configurações inicial e final da treliça de 26 barras.

A Tabela 4.28 mostra a solução ótima e o peso final obtidos pelo algoritmo. Pode-se observar que o peso final da estrutura foi 174.834 lb (79.303 Kgf) que representa uma redução de peso de 88%.

Tabela 4.28: Estrutura ótima obtida para a treliça de 26 barras.

Variável de projeto	Configuração Inicial	Configuração Final
$b_5$	1.000	2.218
$b_6$	1.000	2.178
$b_{11}$	1.000	0.040
$b_{16}$	1.000	0.040
$b_{19}$	1.000	0.078
$b_{26}$	1.000	0.078
Peso (lb)	1411.280	174.834
$\omega_1$ (Hz)	18.381	10.000

A Figura 4.16 mostra a evolução da redução do peso para este problema. Foram necessárias 29 iterações para encontrar a solução ótima e 31 avaliações da função objetivo.

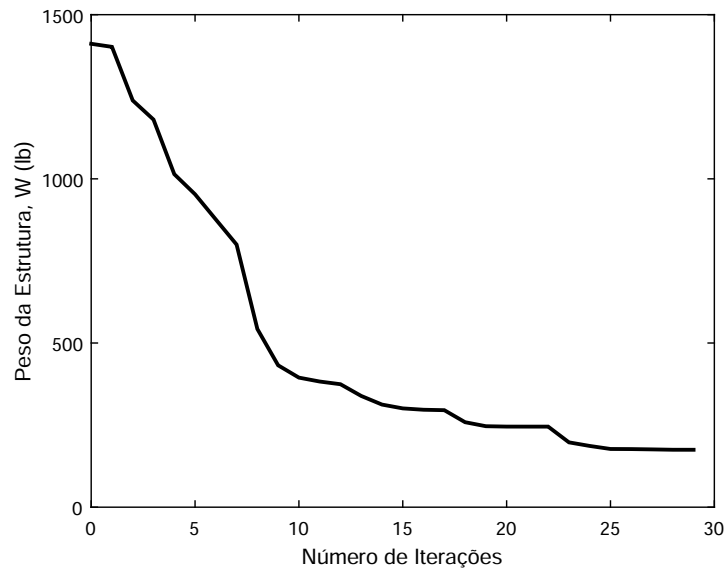


Figura 4.16: Evolução da redução do peso da treliça de 26 barras.

As restrições ativas na configuração final de treliça de 26 barras foram: as áreas mínima de todas as barras com exceção das barras 5, 6, 11, 16, 19 e 26, a complacência estrutural e a mínima frequência natural.

## 4.7 Resumo dos resultados

A seguir, mostram-se o resumo dos resultados obtidos com o algoritmo FDIPA-GSDP, onde pode-se observar a dimensão dos problemas e o esforço computacional dos algoritmo na solução dos problemas teste tais como: o número de iterações do algoritmo e o número de avaliações da função objetivo. Além disso,

são mostrados os valores finais das condições de complementariedade e o número de restrições ativas na solução.

### 4.7.1 Esforço computacional

A Tabela 4.29 mostra o resumo dos resultados numéricos obtidos pelo algoritmo FDIPA-GSDP.

Tabela 4.29: Esforço computacional dos resultados numéricos obtidos com o algoritmo FDIPA-GSDP.

Exemplo	Dimensão do problema					Esforço computacional		
	$n$	$m$	$p$	$q$	$nc$	$NI$	$NF$	tempo (s)
Problema 43 aumentando	4	3	0	4	0	14	37	0.231
Problema 71 aumentando	6	1	1	4	6	19	38	0.281
Treliça plana de 10 barras	10	28	0	8	10	37	43	1.152
Treliça espacial de 25 barras	8	174	0	18	8	31	43	2.049
Treliça espacial de 72 barras	16	336	0	48	16	37	39	9.463
Treliça espacial de 25 barras	13	150	0	18	18	44	53	8.623
Treliça plana de 26 barras	26	0	0	25	26	29	31	0.719

Pode-se observar que em todos os problemas teste o número de iterações para encontrar a solução ótima assim como o número de avaliações da função objetivo foram muito pequenos (Ver Apêndice B).

### 4.7.2 Valores finais das condições de complementariedade

A Tabela 4.30 mostra os valores finais de  $\|\mathbf{d}_0\|$  e  $\|\nabla L\|$  empregados como critérios de parada do algoritmo. Mostra também os valores finais das condições de complementariedade das restrições  $\|\mathcal{G}\gamma_0\|$ ,  $|e_h|$  e  $|tr(\mathbf{\Lambda}_0\mathcal{A})|$  na solução dos problemas teste obtidos com o algoritmo FDIPA-GSDP.

Tabela 4.30: Valores finais de  $\|\mathbf{d}_0\|$ ,  $\|\nabla L\|$  e das condições de complementariedade  $\|\mathcal{G}\gamma_0\|$ ,  $|e_h|$  e  $|tr(\mathbf{\Lambda}_0\mathcal{A})|$  para o algoritmo FDIPA-GSDP.

Exemplo	$\ \mathbf{d}_0\ $	$\ \nabla L\ $	$\ \mathcal{G}\gamma_0\ $	$ e_h $	$ tr(\mathbf{\Lambda}_0\mathcal{A}) $
Problema 43 aumentando	$6.83 \times 10^{-5}$	$7.38 \times 10^{-4}$	$1.49 \times 10^{-5}$	-	$3.50 \times 10^{-5}$
Problema 71 aumentando	$3.09 \times 10^{-5}$	$2.24 \times 10^{-4}$	$8.10 \times 10^{-12}$	$2.43 \times 10^{-5}$	$4.39 \times 10^{-7}$
Treliça plana de 10 barras	$3.42 \times 10^{-5}$	$2.56 \times 10^{-3}$	$5.87 \times 10^{-6}$	-	$8.32 \times 10^{-7}$
Treliça espacial de 25 barras	$1.52 \times 10^{-5}$	$9.38 \times 10^{-4}$	$9.58 \times 10^{-8}$	-	$4.40 \times 10^{-6}$
Treliça espacial de 72 barras	$1.58 \times 10^{-5}$	$2.23 \times 10^{-3}$	$5.85 \times 10^{-8}$	-	$1.66 \times 10^{-6}$
Treliça espacial de 25 barras	$3.10 \times 10^{-5}$	$5.12 \times 10^{-4}$	$4.83 \times 10^{-7}$	-	$2.15 \times 10^{-6}$
Treliça plana de 26 barras	$3.96 \times 10^{-4}$	$4.59 \times 10^{-5}$	$2.87 \times 10^{-3}$	-	-

Pode-se observar que as condições de complementariedade foram verificadas com uma precisão muito boa em todos os problemas teste.

### 4.7.3 Número de restrições ativas

A Tabelas 4.31 mostra o número de restrições estáticas e dinâmicas que ficaram ativas na solução dos problemas obtidos com o algoritmo FDIPA-GSDP.

Tabela 4.31: Número de restrições ativas na solução dos problemas teste com o algoritmo FDIPA-GSDP.

Exemplo	Deslocamento	Tensão	Flambagem	Autovalor	Complacência
Problema 43 aumentando	n/c	n/c	n/c	2	n/c
Problema 71 aumentando	n/c	n/c	n/c	1	n/c
Treliça plana de 10 barras	1	1	n/c	1	n/c
Treliça espacial de 25 barras	4	0	1	1	n/c
Treliça espacial de 72 barras	2	4	n/c	2	n/c
Treliça espacial de 25 barras	n/c	0	9	1	n/c
Treliça plana de 26 barras	n/c	n/c	n/c	1	1

Nota: n/c = não considerada

# Capítulo 5

## Conclusões e sugestões

### 5.1 Conclusões

Neste trabalho foi desenvolvido o algoritmo de pontos interiores e direções viáveis para programação semidefinida generalizada (FDIPA-GSDP) para lidar com problemas de otimização sujeitos a restrições vetoriais de igualdade e desigualdade como também a restrições matriciais semidefinidas.

Este novo algoritmo integra as técnicas empregadas nos algoritmos de pontos interiores e direções viáveis FDIPA e FDIPA-SDP-NS. Dado um ponto inicial no interior de uma região viável definida pelas restrições o algoritmo gera uma sequência de pontos no interior desta região que converge para a solução local do problema. Para isto, em cada iteração é calculada uma direção de busca viável resolvendo dois sistemas lineares com a mesma matriz seguida de uma busca linear ao longo desta direção para definir o novo ponto da sequência para garantir convergência global do algoritmo. Os resultados teóricos apresentados no Capítulo 3 asseguram esta convergência global.

O algoritmo FDIPA-GSDP é muito simples de implementar e permite a busca de um ponto inicial viável quando não são verificadas algumas das restrições vetoriais ou matriciais empregando-se para isso um programa matemático auxiliar. Esta característica do algoritmo é muito importante já que em muitas aplicações não é uma tarefa simples encontrar um ponto inicial que seja viável e não todos os algoritmos existentes na literatura oferecem esta possibilidade.

Os resultados numéricos dos problemas teste apresentados no Capítulo 4 mostram a eficiência e robustez do algoritmo FDIPA-GSDP comparados com os resultados obtidos com o algoritmo PENLAB. Em todos os problemas o esforço computacional foi muito pequeno tomando em conta as poucas iterações necessárias para atingir a solução ótima assim como também o número pequeno de avaliações da função objetivo. Em cada iteração foram avaliadas também as condições

de complementariedade das restrições obtendo-se uma boa precisão em todos os problemas. Em estes problemas foi utilizado o mesmo conjunto de parâmetros.

Finalmente, pode-se mencionar que este trabalho resultou no artigo intitulado “A feasible direction algorithm for general nonlinear semidefinite programming” o qual foi submetido em novembro de 2015 e aprovado para publicação em agosto de 2016 na revista *Structural and Multidisciplinary Optimization* estando disponível pelo momento na versão digital (DOI 10.1007/s00158-016-1564-5).

## 5.2 Sugestões para trabalhos futuros

Como parte dos estudos futuros, destacam-se os seguintes tópicos:

- Estudar possíveis regras de atualização da matriz Quase-Newton  $B$  com a finalidade de melhorar a velocidade de convergência do algoritmo FDIPA-GSDP.
- Explorar mais problemas que possam ser resolvidos com o algoritmo FDIPA-GSDP.
- Realizar uma busca linear ao longo de um arco viável integrando as técnicas do algoritmo de pontos interiores e arcos viáveis (FAIPA) [50].
- Desenvolver um novo algoritmo onde a restrição matricial semidefinida do problema (3.1) é substituída por um cono secante formado por os  $p$  maiores autovalores desta matriz o que permitiria uma redução da dimensão dos sistemas lineares internos do algoritmo FDIPA-GSDP. Desta maneira, a nova reformulação pode ser escrita como

$$\left\{ \begin{array}{l} \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimizar}} \quad f(\mathbf{x}) \\ \text{sujeito a:} \quad \mathbf{g}(\mathbf{x}) \leq 0, \\ \quad \quad \quad \mathbf{c}(\mathbf{x}) \leq 0, \\ \quad \quad \quad \mathbf{h}(\mathbf{x}) = 0 \end{array} \right. \quad (5.1)$$

sendo  $\mathbf{c} : \mathbb{R}^n \rightarrow \mathbb{R}^r$  a função das restrições de desigualdade correspondentes aos autovalores da restrição matricial  $\mathcal{A}(\mathbf{x})$ .

Sem embargo, uma grande dificuldade nesta nova formulação do problema, como já foi mencionado no Capítulo 4, está na ocorrência de autovalores com multiplicidade algebraica maior que um, sendo necessário empregar as técnicas de otimização não diferenciável para poder resolver este novo problema.

Existem muitos trabalhos que proporcionam métodos para o cálculo de derivadas para autovalores com multiplicidade algebraica maior que um tais

como: Ojalvo [51], Mills-Curram [52], Dailey [53], Pedersen [54] e Seyranian et al. [55]. Para o caso de autovalores distintos podem-se utilizar os procedimentos de cálculo de derivadas descritos nos trabalhos de Fox e Kappor [56], Olhoff e Lund [57] e Felix [58].

# Referências Bibliográficas

- [1] HERSKOVITS, J., SOARES, C. M. M., GOULART, E., et al., “Mathematical programming models and algorithms for engineering design optimization”, *Computer Methods in Applied Mechanics and Engineering*, v. 194, pp. 3244–3268, 2005.
- [2] OÑATE, E., *Cálculo de estructuras por el Método de Elementos Finitos. Análisis estático Lineal*. 1st ed. Centro Internacional de Métodos en Ingeniería: España, 1992.
- [3] ELOY, L., *Método dos Elementos Finitos em Análise de Estruturas*. Elsevier Editora, 2011.
- [4] HERSKOVITS, J., “Feasible Direction Interior-Point Technique for Nonlinear Optimization”, *JOTA - Journal of Optimization Theory and Applications*, v. 99, pp. 121–146, 1998.
- [5] HERSKOVITS, J., “A View on Nonlinear Optimization”, In: HERSKOVITS, J. (ed), *Advances in Structural Optimization*, 1st ed., v. 25, chap. 3, pp. 71–116, *Solid Mechanics and its Applications*, Springer Science+Business Media: Dordrecht, 1995.
- [6] SHAPIRO, A., “First and Second Order Analysis of Nonlinear Semidefinite Programs”, *Mathematical Programming*, v. 77, n. 1, pp. 301–320, 1997.
- [7] SHAPIRO, A., “On uniqueness of Lagrange multipliers in optimization problems subject to cone constraints”, *SIAM Journal on Optimization*, v. 7, n. 2, pp. 508–518, 1997.
- [8] AROZTEGUI, M., *Técnicas de programação semidefinida e aplicações em otimização de material*, Tese de D.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 2010.
- [9] KLERK, E. D., *Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications*. v. 65. *Applied Optimization*, Kluber Academic Publishers: Dordrecht, 2002.



- [10] CORREA, R., RAMÍREZ, H., “A Global Algorithm for Nonlinear Semidefinite Programming”, *SIAM Journal on Optimization*, v. 15, n. 1, pp. 303–318, 2004.
- [11] ALIZADEH, F., “Interior Point Methods in Semidefinite Programming with Applications to Combinatorial Optimization”, *SIAM Journal on Optimization*, v. 5, n. 1, pp. 13–51, 1995.
- [12] STINGL, M., *On the solution of nonlinear semidefinite programs by augmented Lagrangian methods*, Phd thesis, University of Erlangen-Nürnberg, España, 2005.
- [13] NESTEROV, Y., NEMIROVSKII, A., *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM, 1994.
- [14] ANTONIOU, A., LU, W., *Practical Optimization: Algorithms and Engineering Applications*. Springer Science+Business Media, LLC, 2007.
- [15] HELMBERG, C., RENDL, F., VANDERBEI, R., et al., “An Interior-Point Method for Semidefinite Programming”, *SIAM Journal on Optimization*, v. 6, n. 2, pp. 342–361, 1996.
- [16] BABYNIN, M., ZHADAN, V., “A primal interior point method for the linear semidefinite programming problem”, *Computational Mathematics and Mathematical Physics*, v. 48, n. 10, pp. 1746–1767, 2008.
- [17] YAMASHITA, H., YABE, H., HARADA, K., “A primal-dual interior point method for Nonlinear Semidefinite Programming”, *Mathematical Programming*, v. 132, n. 1–2, pp. 1–30, 2012.
- [18] KOČVARA, M., STINGL, M., “PENNON: A Code for Convex Nonlinear and Semidefinite Programming”, *Optimization Methods and Software*, v. 18, pp. 317–333, 2003.
- [19] JARRE, F., “An interior Method for Nonlinear Semidefinite Programs”, *Optimization and Engineering*, v. 1, pp. 347–372, 2000.
- [20] FARES, B., NOLL, D., APKARIAN, P., “Robust Control via Sequential Semidefinite Programming”, *SIAM Journal on Optimization*, v. 40, n. 6, pp. 1791–1820, 2002.
- [21] HUANG, X., YANG, X., TEO, K., “Augmented Lagrangian and Nonlinear Semidefinite Programs”, In: GIANNESSE, F., MAUGERI, A. (eds), *Variational Analysis and Applications*, v. 79, chap. 32, pp. 513–529,

*Nonconvex Optimization and Its Applications*, Springer Science+Business Media, Inc.: SingaporeBoston, 2005.

- [22] KANZOW, C., NAGEL, C., KATO, H., et al., “Sucessive Linearization Methods for Nonlinear Semidefinite Programs”, *Computational Optimization and Applications*, v. 31, n. 3, pp. 251–273, 2005.
- [23] LOURENÇO, B., FUKUDA, E., FUKUSHIMA, M., “Optimality conditions for nonlinear semidefinite”, *Mathematical Programming*, pp. 1–24, 2016.
- [24] BAZÁN, E., *Otimização estrutural mediante programação semidefinida*, Dissertação de M.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 2012.
- [25] ACHTZINGER, W., KOCVARA, M., “Structural Topology Optimization with Eigenvalues”, *SIAM Journal on Optimization*, v. 18, pp. 1129–1164, 2007.
- [26] BEN-TAL, A., NEMIROVSKI, A., “Robust Truss Topology Design via Semidefinite Programming”, *SIAM Journal on Optimization*, v. 7, pp. 991–1016, 1997.
- [27] HAFTKA, R., GÜRDAL, Z., *Elements of Structural Optimization*. 3rd ed., v. 11. *Solid Mechanics and its Applications*, Kluwer Academic Publishers: Dordrecht, 1992.
- [28] DEDEA, T., BEKIROGLUB, S., AYVAZC, Y., “Weight minimization of trusses with genetic algorithm”, *Applied Soft Computing*, v. 11, pp. 2565–2575, 2011.
- [29] DEB, K., GULATI, S., “Design of truss-structures for minimum weight using genetic algorithms”, *Finite Elements in Analysis and Design*, v. 37, pp. 447–465, 2001.
- [30] ACHTZINGER, W., KOCVARA, M., “PENNON: Software for Linear and Nonlinear Matrix Inequalities”, In: ANJOS, M., LASSERRE, J. (eds), *Handbook on Semidefinite, Conic and Polynomial Optimization*, 1st ed., v. 166, chap. 26, pp. 755–791, *International Series in Operations Research & Management Science*, Springer Science+Business Media: New York, 2012.
- [31] FIALA, J., KOCVARA, M., STINGL, M., “PENLAB: a matlab solver for nonlinear semidefinite optimization”, *ArXiv e-prints*, 2013.
- [32] OHSAKI, M., FUJISAWA, K., KATO, K., et al., “Semi-definite programming for topology optimization of trusses under multiple

- eigenvalue constraints”, *Computer Methods in Applied Mechanics and Engineering*, v. 180, pp. 203–217, 1999.
- [33] CHONG, E., ZAK, S., *An Introduction to Optimization*. John Wiley & Sons: USA, 2001.
- [34] IZMAILOV, A., SOLODOV, M., *Otimização - Volume 1: Condições de Otimalidade, Análise Convexa e de Dualidade*. IMPA, 2009.
- [35] AROZTEGUI, M., HERSKOVITS, J., ROCHE, J., et al., “A feasible direction interior point algorithm for nonlinear semidefinite programming”, *Structural and Multidisciplinary Optimization*, v. 50, pp. 1019–1035, 2014.
- [36] HARTWING, R., “A note on the partial ordering of positive semidefinite matrices”, *Linear and Multilinear Algebra*, v. 6, pp. 223–226, 2007.
- [37] HORN, R. A., JOHNSON, C. R., *Matrix Analysis*. Cambridge University Press: Cambridge, 1985.
- [38] LUENBERGER, D. G., YE, Y., *Linear and Nonlinear Programming*. 4th ed., v. 228. *International Series in Operations Research & Management Science*, Springer Science+Business Media: Cham, 2016.
- [39] PISSANETZKY, A., *Sparse Matrix Technology*. Academic Press: London, 1984.
- [40] ROCHE, J., HERSKOVITS, J., BAZÁN, E., et al., “A feasible direction algorithm for general nonlinear semidefinite programming”, *Structural and Multidisciplinary Optimization*, pp. 1–19, 2016.
- [41] HOCK, W., SCHITTKOWSKI, K., *Test examples for Nonlinear Programming Codes, Lecture Notes in Economics and Mathematical System*. v. 187. Springer-Verlag: Berlin, Germany, 1981.
- [42] RAO, S., *Engineering Optimization. Theory and Practice*. 4th ed. John Wiley & Sons, 2009.
- [43] FARSHI, B., ALINIA-ZIAZI, A., “Sizing optimization of truss structures by method of centers and force formulation”, *International Journal of Solids and Structures*, v. 47, pp. 2508–2524, 2010.
- [44] SALAJEGHEH, E., VANDERPLAATS, G., “Optimum design of trusses with discrete sizing and shape variables”, *Structural Optimization*, v. 6, pp. 79–85, 1993.

- [45] LEMONGE, A., *Aplicação de Algoritmos Genéticos em Otimização Estrutural*, Tese de D.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 1999.
- [46] LEE, K., HAN, S., GEEM, Z., “Discrete and Discrete-Continuous Configuration Optimization Methods for Truss Structures Using the Harmony Search Algorithm”, *International Journal of Optimization in Civil Engineering*, v. 1, pp. 107–126, 2011.
- [47] EBADI, M., RASHTCHI, V., BEHRAVESH, A., “Sizing and Geometry Optimization of Pin Connected Structures Via Real Coded Genetic Algorithm”, *Global Journal of Computer Science and Technology*, v. 11, n. 9, pp. 25–30, 2011.
- [48] AUSTRELL, P., DAHLBOM, O., LINDEMANN, J., et al., *CALFEM: A Finite Element Toolbox 3.4*, Division of Structural Mechanics, Lund University, <http://www.byggmek.lth.se/Calfem>, 2004.
- [49] HOCK, W., SCHITTKOWSKI, K., “Test examples for nonlinear programming codes”, *Journal of Optimization Theory and Applications*, v. 30, n. 1, pp. 127–129, 1980.
- [50] HERSKOVITS, J., GOULART, E., AROZTEGUI, M., “Large Scale Structural Optimization”, In: ARORA, J. (ed), *Optimization of Structural and Mechanical Systems*, chap. 2, pp. 35–57, World Scientific Publishing Co. Pte. Ltd.: Singapore, 2006.
- [51] OJALVO, I., “Efficient Computation of Modal Sensitivities for System with Repeated Frequencies”, *AIAA Journal*, v. 26, n. 3, pp. 361–366, 1988.
- [52] MILLS-CURRAN, W., “Calculation of eigenvector derivatives for structures with repeated eigenvalues”, *AIAA Journal*, v. 26, n. 7, pp. 867–871, 1988.
- [53] DAILEY, R., “Eigenvector Derivatives with Repeated Eigenvalues”, *AIAA Journal*, v. 27, n. 4, pp. 486–491, 1989.
- [54] PEDERSEN, N., NIELSEN, A., “Optimization of practical trusses with constraints on eigenfrequencies, displacement, stresses, and buckling”, *Structural and Multidisciplinary Optimization*, v. 25, pp. 436–445, 2003.
- [55] SEYRANIAN, A., LUND, E., OLHOFF, N., “Multiple eigenvalues in structural optimization problems”, *Structural Optimization*, v. 8, pp. 207–227, 1994.
- [56] FOX, R., KAPOOR, M., “Rates of Change of Eigenvalues and Eigenvectors”, *AIAA Journal*, v. 6, n. 12, pp. 2426–2429, 1968.

- [57] OLHOFF, N., LUND, E., “Finite Element Based Engineering Design Sensitivity Analysis and Optimization”, In: HERSKOVITS, J. (ed), *Advances in Structural Optimization*, 1st ed., v. 25, chap. 1, pp. 1–45, *Solid Mechanics and its Applications*, Springer Science+Business Media: Dordrecht, 1995.
- [58] FELIX, J., *Shape Optimization of Trusses Subject to Strength, Displacement, and Frequency Constraints*, Master’s thesis, Naval Postgraduate School, Monterey, California, 1981.
- [59] OHSAKI, M., KANNO, Y., “Semidefinite Programming for Structural Optimization”, In: ARORA, J. (ed), *Optimization of Structural and Mechanical Systems*, chap. 19, pp. 541–567, World Scientific Publishing Co. Pte. Ltd.: Singapore, 2007.
- [60] CHANDRUPATLA, T., BELEGUNDU, A., *Introducción al estudio del Elemento Finito en Ingeniería*. 2nd ed. Prentice Hall, 1999.
- [61] KIM, N., SANKAR, B., *Introduction to Finite Element Analysis and Design*. John Wiley & Sons, 2009.

# Apêndice A

## Definições matemáticas

### A.1 Matrizes

**Definição A.1 (Matriz Identidade).** Uma matriz identidade é uma matriz diagonal onde seus elementos são todos iguais a um, isto é

$$\mathbf{I}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \quad (\text{A.1})$$

**Definição A.2 (Matriz Simétrica).** Uma matriz quadrada  $\mathbf{A} \in \mathbb{R}^{n \times n}$  é chamada simétrica se a matriz e sua transposta são iguais, isto é

$$\mathbf{A}^T = \mathbf{A} \quad (\text{A.2})$$

**Definição A.3 (Matriz Definida Positiva).** Uma matriz  $\mathbf{M} \in \mathbb{R}^{n \times n}$  é definida positiva quando

$$\mathbf{v}^T \mathbf{M} \mathbf{v} > 0 \quad (\text{A.3})$$

para qualquer  $\mathbf{v} \neq \mathbf{0} \in \mathbb{R}^n$ .

**Teorema A.1 (Propriedades de uma matriz definida (ou semidefinida) positiva).** *Seja  $\mathbf{X} \in \mathcal{S}^n$ . Os seguintes enunciados são equivalentes:*

- (i)  $\mathbf{X} \in \mathcal{S}_+^n$  ou  $\mathbf{X} \succcurlyeq \mathbf{0}$  ( $\mathbf{X}$  é semidefinida positiva);
- (ii)  $\mathbf{z}^T \mathbf{X} \mathbf{z} \geq 0, \forall \mathbf{z} \in \mathbb{R}^n$ ;
- (iii)  $\lambda_{\min}(\mathbf{X}) \geq 0$ , onde  $\lambda_{\min}$  é o menor autovalor da matriz  $\mathbf{X}$ ;
- (iv) Todos os menores principais do  $\mathbf{X}$  são não negativos;

(v)  $\mathbf{X} = \mathbf{L}\mathbf{L}^T$  para algum  $\mathbf{L} \in \mathbb{R}^{n \times n}$  é chamada *descomposição de Cholesky* da matriz  $\mathbf{X}$ .

*Demonstração.* Ver [37].

**Definição A.4 (Traço de uma matriz).** Para uma matriz quadrada  $\mathbf{A} \in \mathbb{R}^{q \times q}$ , define-se o traço da matriz  $\mathbf{A}$  como a soma dos elementos da diagonal de  $\mathbf{A}$ , isto é

$$\text{tr}(\mathbf{A}) = \sum_{i=1}^n a_{ii} \quad (\text{A.4})$$

**Lema A.1.** *Sejam  $\mathbf{A} \in \mathbb{R}^{q \times q}$  e  $\mathbf{B} \in \mathbb{R}^{q \times q}$ . Os seguintes enunciados são equivalentes:*

$$(i) \text{tr}(\mathbf{A}) = \sum_{i=1}^q \lambda_i(\mathbf{A})$$

$$(ii) \text{tr}(\mathbf{A}) = \text{tr}(\mathbf{A}^T)$$

$$(iii) \text{tr}(\mathbf{A}\mathbf{B}) = \text{tr}(\mathbf{B}\mathbf{A})$$

$$(iv) \text{tr}(\mathbf{A}\mathbf{B}^T) = \text{vec}(\mathbf{A})^T \text{vec}(\mathbf{B})$$

*Demonstração.* Ver [9].

**Definição A.5 (Norma de uma matriz).** Define-se a norma Frobenius em  $\mathbb{R}^{m \times n}$  da matriz  $\mathbf{A}$ , denotada por  $\|\mathbf{A}\|_F$  como:

$$\|\mathbf{A}\|_F = \left( \sum_{i=1}^m \sum_{j=1}^n (a_{ij})^2 \right)^{1/2} \quad (\text{A.5})$$

onde  $\mathbf{A} \in \mathbb{R}^{m \times n}$ . A norma de Frobenius em  $\mathbb{R}^{n \times n}$  define-se através do produto interno como:

$$\|\mathbf{A}\| = \sqrt{\langle \mathbf{A}, \mathbf{A} \rangle} = \sqrt{\text{tr}(\mathbf{A}\mathbf{A}^T)} \quad (\text{A.6})$$

a qual satisfaz as seguintes condições:

(i)  $\|\mathbf{A}\| > 0$  se  $\mathbf{A} \neq \mathbf{O}$ , e  $\|\mathbf{O}\| = 0$ , onde  $\mathbf{O}$  é a matriz com todos seus elementos nulos;

(ii)  $\|c\mathbf{A}\| = |c| \|\mathbf{A}\|$ , para todo  $c \in \mathbb{R}$ ;

(iii)  $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$ ;

(iv)  $\|\mathbf{A}\mathbf{B}\| \leq \|\mathbf{A}\| \|\mathbf{B}\|$ .

*Demonstração.* Ver [33].

## A.2 Problema de autovalores

Um problema de autovalores é definido como

$$\mathbf{A}\phi = \lambda\phi \quad (\text{A.7})$$

onde  $\mathbf{A} \in \mathbb{R}^{n \times n}$ .

A Equação (A.7) pode ser reescrita como

$$[\mathbf{A} - \lambda\mathbf{I}_n]\phi = \mathbf{0} \quad (\text{A.8})$$

Uma solução não nula para  $\phi$  poderá ser obtida quando a matriz  $[\mathbf{A} - \lambda\mathbf{I}_n]$  seja singular já que  $\phi \neq \mathbf{0}$ , ou seja se

$$\det(\mathbf{A} - \lambda\mathbf{I}_n) = 0 \quad (\text{A.9})$$

A Equação (A.9) é conhecida como a equação característica de  $\mathbf{A}$  e a expressão  $\det(\mathbf{A} - \lambda\mathbf{I}_n)$  representa o polinômio característico de  $\mathbf{A}$ .

Podem-se encontrar as  $n$  raízes do polinômio característico que representam os autovalores de  $\mathbf{A}$  denotados por  $\lambda_1, \lambda_2, \dots, \lambda_n$ . A cada autovalor  $\lambda_i$  corresponde-lhe um autovetor  $\phi_i$ . Definem-se  $\lambda_{\min} = \lambda_1$  o autovalor mínimo e  $\lambda_{\max} = \lambda_n$  o autovalor máximo.

## A.3 Transformação linear

Uma transformação (ou mapeamento)  $T$  é uma aplicação que satisfaz as seguintes condições:

- (i)  $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ , para todo  $\mathbf{u}, \mathbf{v}$  no domínio de  $T$ ;
- (ii)  $T(c\mathbf{u}) = cT(\mathbf{u})$ , para todo  $\mathbf{u}$  e todo escalar  $c$ .

*Demonstração.* Ver [37].

## A.4 Complacência estrutural

Define-se a complacência estrutural  $W$  [59] como o trabalho externo equivalente à energia de deformação:

$$W = \mathbf{p}^T \mathbf{u} = \mathbf{u}^T \mathbf{K} \mathbf{u} = 2 \left( \mathbf{p}^T - \frac{1}{2} \mathbf{u}^T \mathbf{K} \mathbf{u} \right) = \mathbf{p}^T \mathbf{K}^{-1} \mathbf{p} \quad (\text{A.10})$$



## A.5 Complemento de Schur

**Teorema A.2.** *Seja a matriz:*

$$\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{pmatrix} \quad (\text{A.11})$$

onde  $\mathbf{A}$  é uma matriz definida positiva e  $\mathbf{C}$  é uma matriz simétrica, então a matriz:

$$\mathbf{D} = \mathbf{C} - \mathbf{B}^T \mathbf{A}^{-1} \mathbf{B} \quad (\text{A.12})$$

é chamada o complemento de Schur de  $\mathbf{A}$  em  $\mathbf{M}$ . Os seguintes enunciados são equivalentes:

(i)  $\mathbf{M}$  é definida (ou semidefinida) positiva.

(ii)  $\mathbf{D}$  é definida (ou semidefinida) positiva.

*Demonstração.* Ver [9].

# Apêndice B

## Histórico das condições de complementariedade dos problemas teste resolvidos com o algoritmo FDIPA-GSDP

As Tabelas B.2, B.3, B.4, B.5, B.6, B.7 e B.8 mostram a evolução da redução da função objetivo e o histórico das condições de complementariedade das restrições para os exemplos numéricos do Capítulo 4, onde foi empregada a seguinte notação descrita na seguinte tabela

Tabela B.1: Notação empregada nas tabelas dos resultados numéricos

Notação	Descrição
$iter$	Iteração $k$
$f$	Valor da função objetivo avaliada no ponto $\mathbf{x}^k$
$\ \mathbf{d}_0\ $	Norma da direção de busca avaliada no ponto $\mathbf{x}^k$
$\ \nabla L\ $	Norma do gradiente da função Lagrangiana avaliada no ponto $\mathbf{x}^k$
$\ \mathcal{G}\gamma_0\ $	Norma da condição de complementariedade das restrições de desigualdade avaliada no ponto $\mathbf{x}^k$
$ tr(\mathbf{\Lambda}_0\mathcal{A}) $	Valor absoluto do traço da condição de complementariedade das restrições matriciais avaliada no ponto $\mathbf{x}^k$
$ e_h $	Máximo erro nas restrições de igualdade avaliada no ponto $\mathbf{x}^k$
$t$	Comprimento do passo $t^k$ da busca linear de Armijo
$itlin$	Número de iterações na busca linear de Armijo

## B.1 Problema 43 aumentado

Tabela B.2: Evolução da redução da função objetivo para o problema 43 aumentado.

<i>iter</i>	<i>f</i>	$\ \mathbf{d}_0\ $	$\ \nabla L\ $	$\ \mathcal{G}\gamma_0\ $	$ e_h $	$ tr(\mathbf{\Lambda}_0\mathcal{A}) $	t	<i>itlin</i>
0	-0.002							
1	-18.288	10.762	10.762	1.799	0	22.725	0.020	12
2	-26.324	4.541	12.467	1.792	0	15.916	0.206	2
3	-31.589	2.604	3.862	1.527	0	11.026	0.283	2
4	-36.885	0.884	1.426	0.458	0	4.674	0.727	1
5	-38.881	0.848	1.479	0.048	0	1.852	0.700	2
6	-39.995	0.420	3.132	0.087	0	0.863	0.700	2
7	-40.788	0.199	1.014	0.068	0	0.118	1.000	1
8	-40.903	0.036	0.329	0.020	0	0.041	0.700	2
9	-40.943	0.020	0.216	0.006	0	0.014	0.700	2
10	-40.957	0.008	0.086	0.002	0	0.004	0.700	2
11	-40.961	0.003	0.027	$5.500\times 10^{-4}$	0	0.002	0.700	2
12	-40.963	$7.579\times 10^{-4}$	0.008	$1.650\times 10^{-4}$	0	$3.883\times 10^{-4}$	0.700	2
13	-40.963	$2.275\times 10^{-4}$	0.002	$4.951\times 10^{-5}$	0	$1.166\times 10^{-4}$	0.700	2
14	-40.963	$6.827\times 10^{-5}$	$7.379\times 10^{-4}$	$1.485\times 10^{-5}$	0	$3.497\times 10^{-5}$	0.700	2

## B.2 Problema 71 aumentado

Tabela B.3: Evolução da redução da função objetivo para o problema 71 aumentado.

<i>iter</i>	<i>f</i>	$\ \mathbf{d}_0\ $	$\ \nabla L\ $	$\ \mathcal{G}\gamma_0\ $	$ e_h $	$ tr(\mathbf{\Lambda}_0\mathcal{A}) $	t	<i>itlin</i>
0	174.235							
1	139.791	0.803	0.803	43.660	11.365	1.842	1.000	1
2	136.399	8.624	56.292	14.759	1.778	0.364	0.027	4
3	132.309	17.576	86.877	24.675	1.844	16.911	0.015	3
4	123.110	2.371	34.293	10.887	1.787	16.569	0.204	1
5	120.557	4.647	44.991	12.293	1.663	14.155	0.030	1
6	117.564	2.252	47.332	7.723	1.638	1.016	0.250	1
7	116.722	4.282	48.145	8.686	1.389	1.600	0.029	1
8	113.039	0.508	9.0127	1.642	1.358	0.285	0.818	1
9	111.589	1.153	8.031	0.659	0.474	0.118	0.122	1
10	111.152	4.239	15.913	0.081	0.451	0.409	0.012	1
11	108.300	2.022	31.125	0.337	0.534	0.521	0.168	6
12	105.668	3.700	55.574	0.087	0.531	1.663	0.075	8
13	97.502	0.687	66.521	0.021	0.305	0.427	0.700	2
14	93.347	0.759	27.538	0.010	0.320	0.600	0.233	1
15	90.810	0.197	5.566	0.002	0.281	0.426	1.000	1
16	88.629	0.104	2.756	$3.421\times 10^{-4}$	0.041	0.130	1.000	1
17	87.822	0.030	0.535	$1.705\times 10^{-5}$	0.014	0.015	0.947	1
18	87.711	0.004	0.993	$5.092\times 10^{-8}$	0.002	$4.459\times 10^{-5}$	1.000	1
19	87.711	$3.089\times 10^{-5}$	$2.239\times 10^{-4}$	$8.102\times 10^{-12}$	$2.433\times 10^{-5}$	$4.388\times 10^{-7}$	0.990	1

## B.3 Treliça de 10 barras

Tabela B.4: Evolução da redução do peso da treliça de 10 barras.

$iter$	$f$	$\ \mathbf{d}_0\ $	$\ \nabla L\ $	$\ \mathcal{G}\gamma_0\ $	$ tr(\mathbf{\Lambda}_0\mathcal{A}) $	$t$	$itlin$
0	34578.300						
1	34514.700	0.631	0.631	0.001	$3.437\times 10^4$	1.000	1
2	24691.700	94.239	21.355	$6.032\times 10^{-5}$	$1.999\times 10^4$	1.000	1
3	10920.900	133.843	18.783	$4.437\times 10^{-6}$	$9.834\times 10^4$	1.000	1
4	10319.600	63.326	9.594	$5.518\times 10^{-6}$	$9.975\times 10^4$	0.088	1
5	10205.300	77.621	7.598	$4.998\times 10^{-5}$	$1.003\times 10^4$	0.014	1
6	10092.500	93.183	6.320	0.003	$1.001\times 10^4$	0.016	1
7	8201.540	79.023	2.783	1.419	$6.043\times 10^3$	0.331	2
8	7539.330	13.105	2.961	216.422	$1.482\times 10^3$	1.000	1
9	7169.090	5.156	14.355	147.286	$2.298\times 10^3$	1.000	1
10	6510.980	10.538	35.334	209.122	$1.127\times 10^3$	1.000	1
11	5937.810	12.532	25.787	310.958	438.921	0.777	1
12	5736.560	10.139	21.796	357.433	362.393	0.272	1
13	5638.510	9.693	7.6814	355.492	389.914	0.131	1
14	5592.120	7.945	14.099	345.559	393.930	0.077	1
15	5550.570	6.530	33.061	333.610	280.862	0.088	1
16	5462.580	5.623	91.697	274.437	120.607	0.257	1
17	5366.520	4.965	125.572	199.906	80.075	0.363	1
18	5329.600	3.717	67.618	186.913	57.249	0.167	1
19	5310.220	3.013	56.832	174.175	42.450	0.109	1
20	5252.430	3.018	51.394	134.804	24.442	0.343	1
21	5236.260	2.241	34.049	124.475	19.490	0.128	1
22	5197.040	1.885	42.440	93.229	9.273	0.361	1
23	5184.070	1.631	22.821	84.236	7.362	0.153	1
24	5166.950	1.586	20.400	69.973	3.734	0.233	1
25	5143.840	1.570	16.244	50.768	1.973	0.374	1
26	5137.490	0.911	9.732	33.336	2.094	0.187	1
27	5126.670	0.880	47.427	24.888	1.805	0.327	3
28	5116.130	0.951	85.580	8.770	0.170	0.704	1
29	5113.370	0.220	59.671	2.095	0.121	0.700	2
30	5112.280	0.234	32.146	0.304	0.037	0.764	1
31	5111.770	0.088	14.922	0.101	0.071	1.000	1
32	5111.590	0.114	3.190	0.042	0.012	1.000	1
33	5111.520	0.059	2.408	0.007	0.010	0.700	2
34	5111.480	0.022	4.802	0.005	0.005	1.000	1
35	5111.470	$8.040\times 10^{-4}$	0.075	0.002	$2.438\times 10^{-5}$	1.000	1
36	5111.470	$1.657\times 10^{-4}$	0.012	$1.380\times 10^{-4}$	$5.596\times 10^{-6}$	1.000	1
37	5111.470	$3.419\times 10^{-5}$	0.003	$5.873\times 10^{-6}$	$8.315\times 10^{-7}$	1.000	1

## B.4 Treliça de 25 barras

Tabela B.5: Evolução da redução do peso da treliça de 25 barras.

$iter$	$f$	$\ \mathbf{d}_0\ $	$\ \nabla L\ $	$\ \mathcal{G}\gamma_0\ $	$ tr(\mathbf{\Lambda}_0\mathcal{A}) $	$t$	$itlin$
0	4134.120						
1	4133.170	0.008	0.008	$1.226\times 10^{-4}$	4133	1.000	1
2	3795.260	2.707	0.549	$1.713\times 10^{-6}$	3781	1.000	1
3	3148.090	5.084	0.428	$7.175\times 10^{-8}$	3140	1.000	1
4	2223.110	7.109	1.059	$6.198\times 10^{-9}$	2216	1.000	1
5	1481.400	8.050	1.366	$1.246\times 10^{-9}$	1472	0.700	2
6	978.338	8.962	4.001	$3.551\times 10^{-10}$	948.717	0.467	2
7	919.277	10.211	2.625	$7.468\times 10^{-10}$	905.204	0.060	1
8	908.597	9.946	0.722	$9.819\times 10^{-9}$	905.696	0.011	1
9	906.099	9.578	0.084	$6.633\times 10^{-7}$	905.544	0.003	1
10	905.455	9.467	0.040	$1.827\times 10^{-4}$	904.199	$8.034\times 10^{-4}$	1
11	776.864	5.895	0.717	4.355	179.415	0.581	1
12	749.099	0.789	12.611	8.391	165.121	0.704	1
13	742.379	0.804	9.852	9.291	139.139	0.180	1
14	740.448	1.658	12.963	9.747	118.019	0.035	1
15	722.619	2.412	38.120	12.039	93.303	0.261	1
16	720.611	1.906	3.587	12.194	86.405	0.032	1
17	686.815	2.175	27.500	14.951	46.217	0.559	1
18	685.803	1.392	7.117	15.087	42.680	0.024	1
19	670.846	1.334	7.609	13.939	27.104	0.363	1
20	653.187	0.886	6.565	11.052	12.027	0.575	1
21	648.134	0.405	6.154	9.313	8.205	0.277	2
22	647.069	0.311	6.993	8.791	7.345	0.072	3
23	646.744	0.255	4.547	8.802	6.422	0.023	2
24	645.013	0.176	2.258	8.153	3.393	0.144	2
25	634.783	0.141	6.739	2.995	0.324	1.000	1
26	633.254	0.100	0.109	2.020	0.102	0.462	3
27	631.448	0.027	1.333	0.828	0.052	0.700	2
28	630.640	0.011	0.599	0.294	0.061	0.688	2
29	630.203	0.004	0.220	0.006	0.023	1.000	1
30	630.192	$1.181\times 10^{-4}$	0.008	$1.581\times 10^{-4}$	$1.846\times 10^{-4}$	0.976	1
31	630.192	$1.516\times 10^{-5}$	$9.380\times 10^{-4}$	$9.578\times 10^{-8}$	$4.401\times 10^{-6}$	1.000	1

## B.5 Treliça de 72 barras

Tabela B.6: Evolução da redução do peso da treliça de 72 barras.

<i>iter</i>	<i>f</i>	$\ \mathbf{d}_0\ $	$\ \nabla L\ $	$\ \mathcal{G}\gamma_0\ $	$ tr(\mathbf{\Lambda}_0\mathbf{A}) $	<i>t</i>	<i>itlin</i>
0	2770.51						
1	2770.26	0.001	0.001	$1.710\times 10^{-4}$	$2.770\times 10^4$	1.000	1
2	2675.97	0.992	0.206	$8.518\times 10^{-6}$	$2.673\times 10^4$	0.488	1
3	2628.31	2.192	0.220	$8.165\times 10^{-6}$	$2.626\times 10^4$	0.112	1
4	2462.11	4.302	0.427	$1.733\times 10^{-4}$	$2.443\times 10^4$	0.234	1
5	2217.77	7.479	1.494	0.007	$2.180\times 10^4$	0.224	1
6	2164.7	11.247	3.814	0.009	$2.108\times 10^4$	0.041	1
7	1786.47	13.462	4.872	0.010	$1.707\times 10^4$	0.257	1
8	1598.01	13.755	3.228	1.446	$1.522\times 10^4$	0.151	1
9	1533.29	12.643	24.395	3.439	$1.458\times 10^4$	0.058	1
10	1360.68	11.930	24.904	7.627	$1.279\times 10^4$	0.162	1
11	1266.74	10.508	31.722	10.388	$1.181\times 10^4$	0.099	1
12	1074.9	9.639	34.698	16.032	972.512	0.222	1
13	902.574	7.749	26.086	23.888	785.997	0.242	1
14	892.732	5.871	79.193	31.554	692.367	0.018	1
15	806.254	5.109	176.815	29.020	464.549	0.211	1
16	760.707	3.241	49.136	29.674	590.168	0.149	1
17	694.976	1.824	36.275	18.614	565.096	0.409	1
18	574.295	2.085	49.000	11.680	415.453	0.629	1
19	520.79	1.436	25.552	12.387	289.715	0.406	1
20	474.979	0.532	12.947	5.270	264.437	1.000	1
21	458.154	0.177	7.390	2.129	253.298	1.000	1
22	446.739	0.151	6.437	1.048	161.180	0.884	1
23	446.573	0.066	3.432	1.338	72.115	0.027	1
24	445.069	0.048	1.950	1.112	39.240	0.318	1
25	441.173	0.061	4.206	0.603	26.793	1.000	1
26	440.642	0.088	6.012	0.618	29.239	0.105	1
27	440.64	0.182	12.219	0.623	33.501	$2.116\times 10^{-4}$	1
28	435.858	0.080	7.895	0.491	6.775	1.000	1
29	429.626	0.151	9.302	0.111	5.289	1.000	1
30	424.075	0.135	18.407	0.050	2.209	1.000	1
31	421.185	0.118	15.966	0.072	1.111	1.000	1
32	419.843	0.060	4.828	0.080	0.360	1.000	1
33	419.372	0.052	4.560	0.033	0.035	0.700	2
34	419.108	0.012	4.946	0.022	0.016	1.000	1
35	419.023	0.003	1.265	$6.965\times 10^{-4}$	0.002	1.000	1
36	419.019	$3.051\times 10^{-4}$	0.148	$9.053\times 10^{-6}$	$1.636\times 10^{-5}$	1.000	1
37	419.019	$1.578\times 10^{-5}$	0.002	$5.851\times 10^{-8}$	$1.656\times 10^{-6}$	1.000	1

## B.6 Treliça de 25 barras

Tabela B.7: Evolução da redução do peso da treliça de 25 barras.

<i>iter</i>	<i>f</i>	$\ \mathbf{d}_0\ $	$\ \nabla L\ $	$\ \mathcal{G}\gamma_0\ $	$ \text{tr}(\Lambda_0 \mathcal{A}) $	<i>t</i>	<i>itlin</i>
0	2674.240						
1	2653.480	0.597	0.597	0.007	2651.110	1.000	1
2	2519.140	58.384	40.204	0.062	1871.420	0.090	1
3	1478.910	58.530	30.893	0.002	1136.130	0.801	1
4	1301.950	21.151	19.687	0.006	1093.680	0.419	1
5	1265.050	22.013	33.697	0.014	928.845	0.056	1
6	1204.960	19.618	43.472	0.011	738.820	0.078	1
7	788.206	13.126	38.122	6.449	510.965	0.442	1
8	649.004	19.338	15.754	7.574	590.921	0.167	3
9	592.978	49.130	27.301	7.061	522.463	0.100	2
10	523.389	6.262	13.954	3.543	425.005	0.283	1
11	434.344	4.994	20.248	13.835	252.004	1.000	1
12	373.145	4.542	37.287	7.736	285.312	1.000	1
13	339.986	5.009	65.055	7.276	231.484	0.241	1
14	333.566	6.902	84.755	6.877	205.316	0.040	1
15	318.985	8.566	92.867	7.151	178.842	0.085	1
16	280.178	8.182	82.253	7.078	137.422	0.238	1
17	276.954	5.078	33.947	6.059	69.387	0.044	1
18	262.782	0.841	14.971	2.082	126.753	1.000	1
19	258.864	8.931	11.749	2.824	106.561	0.101	1
20	252.702	4.521	17.824	3.612	78.887	0.132	1
21	235.678	4.356	13.324	4.368	50.017	0.377	1
22	217.391	4.187	19.594	3.934	21.763	0.607	1
23	216.607	7.230	8.321	4.337	19.203	0.032	1
24	207.993	5.150	13.596	3.148	11.070	0.516	1
25	205.027	10.394	12.309	2.711	9.088	0.202	1
26	190.227	20.417	53.211	1.886	4.895	0.858	1
27	186.698	16.611	25.258	1.649	5.649	0.218	1
28	185.742	10.010	35.547	1.564	7.159	0.096	1
29	184.201	6.592	24.026	1.357	4.954	0.206	1
30	183.637	9.169	13.190	1.460	4.786	0.066	1
31	177.724	8.412	11.563	0.783	1.576	0.801	1
32	176.184	4.245	15.877	0.590	0.931	0.426	1
33	175.736	1.565	19.005	0.517	0.868	0.200	1
34	175.510	1.147	21.914	0.472	0.784	0.121	1
35	174.283	2.360	19.401	0.289	0.177	0.695	2
36	174.021	0.934	13.192	0.224	0.152	0.324	1
37	173.683	0.429	4.827	0.130	0.078	0.566	1
38	173.357	0.231	1.219	0.032	0.004	1.000	1
39	173.276	0.036	1.330	0.008	$3.555 \times 10^{-4}$	1.000	1
40	173.255	0.008	0.066	0.002	0.001	1.000	1
41	173.251	0.001	0.006	$9.470 \times 10^{-4}$	$5.363 \times 10^{-4}$	0.553	1
42	173.249	$9.220 \times 10^{-4}$	0.011	$1.622 \times 10^{-4}$	$4.574 \times 10^{-4}$	0.849	1
43	173.248	$1.929 \times 10^{-4}$	0.001	$1.306 \times 10^{-5}$	$7.892 \times 10^{-5}$	0.906	1
44	173.248	$3.099 \times 10^{-5}$	$5.118 \times 10^{-4}$	$4.831 \times 10^{-7}$	$2.152 \times 10^{-6}$	0.975	1

## B.7 Treliça de 26 barras

Tabela B.8: Evolução da redução do peso da treliça de 26 barras.

<i>iter</i>	<i>f</i>	$\ \mathbf{d}_0\ $	$\ \nabla L\ $	$ tr(\mathbf{\Lambda}_0 \mathcal{A}) $	<i>t</i>	<i>itlin</i>
0	1411.280					
1	1401.970	0.087	0.087	1230.390	1.000	1
2	1238.670	5.542	1.564	1070.280	0.237	1
3	1180.470	8.197	1.862	1007.690	0.065	1
4	1014.210	4.795	2.678	790.895	0.206	1
5	952.986	4.289	1.673	723.685	0.091	1
6	876.211	7.524	3.163	675.211	0.111	1
7	799.779	2.564	1.477	580.822	0.143	1
8	542.663	2.789	1.546	326.292	0.488	1
9	432.093	1.806	0.556	220.434	0.331	1
10	394.594	1.226	0.248	185.960	0.161	1
11	382.684	1.046	0.114	175.086	0.060	1
12	374.493	0.986	0.069	167.589	0.044	1
13	338.779	0.945	0.073	136.447	0.198	1
14	312.502	0.774	0.066	113.075	0.177	1
15	300.869	0.654	0.026	102.801	0.093	1
16	296.738	0.599	0.013	99.0114	0.036	1
17	295.436	0.577	0.012	97.6651	0.012	1
18	258.809	0.566	0.013	66.816	0.337	1
19	246.625	0.402	0.013	56.503	0.162	1
20	245.446	0.342	0.007	55.359	0.018	1
21	245.373	0.336	0.004	55.166	0.001	1
22	245.358	0.335	0.003	55.089	$2.487 \times 10^{-4}$	1
23	197.439	0.335	0.003	17.529	0.758	1
24	186.505	0.107	0.004	8.755	0.521	2
25	177.268	0.054	0.001	1.765	0.826	1
26	176.958	0.012	$7.726 \times 10^{-4}$	1.517	0.129	1
27	176.010	0.011	$1.951 \times 10^{-4}$	0.841	0.453	1
28	174.902	0.006	$1.402 \times 10^{-4}$	0.054	0.946	1
29	174.834	$3.963 \times 10^{-4}$	$4.586 \times 10^{-5}$	0.003	0.948	1



# Apêndice C

## Análise estrutural de treliças com CALFEM

Uma treliça tem membros sujeitos a duas forças de tração ou de compressão. Os carregamentos são aplicadas somente nos nós e todos os membros estão conectados nos extremos por meio de articulações sem fricção.

### C.1 Coordenadas locais e globais

Aqui, são definidos dois tipos coordenadas: as coordenadas locais  $s$  e as coordenadas globais  $x$  e  $y$ . A Figura C.1 mostra um elemento barra bidimensional [60, 61]. O ângulo  $\phi$  define a rotação do eixo da barra do sistema de coordenadas local em relação ao sistema de coordenadas global.

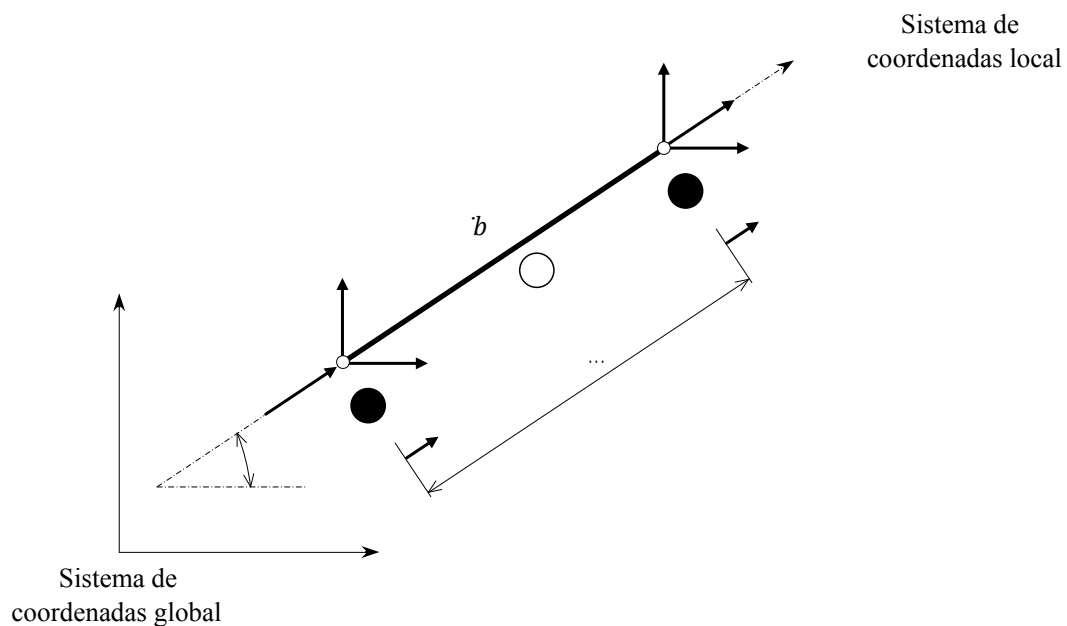


Figura C.1: Definição do elemento barra bidimensional [2, 3].

## C.2 Vetores de deslocamentos e de carregamentos

Os vetores de deslocamentos e de carregamentos nos sistemas de coordenadas local e global são definidos como

$$\bar{\mathbf{u}}^e = \begin{bmatrix} u_{1s}^e & u_{2s}^e \end{bmatrix}^T \quad (\text{C.1})$$

$$\bar{\mathbf{f}}^e = \begin{bmatrix} f_{1s}^e & f_{2s}^e \end{bmatrix}^T \quad (\text{C.2})$$

$$\mathbf{u}^e = \begin{bmatrix} u_1^e & v_1^e & u_2^e & v_2^e \end{bmatrix}^T \quad (\text{C.3})$$

$$\mathbf{f}^e = \begin{bmatrix} f_{1x}^e & f_{1y}^e & f_{2x}^e & f_{2y}^e \end{bmatrix}^T \quad (\text{C.4})$$

Considere-se as seguintes expressões

$$l^e = \cos \phi^e = \frac{x_2^e - x_1^e}{L^e} \quad (\text{C.5})$$

e

$$m^e = \sin \phi^e = \frac{y_2^e - y_1^e}{L^e} \quad (\text{C.6})$$

Define-se a matriz de transformação  $\mathbf{T}^e$  como

$$\mathbf{T}^e = \begin{bmatrix} l^e & m^e & 0 & 0 \\ 0 & 0 & l^e & m^e \end{bmatrix} \quad (\text{C.7})$$

onde  $L$  corresponde ao comprimento do elemento barra determinado por

$$L^e = \sqrt{(x_2^e - x_1^e)^2 + (y_2^e - y_1^e)^2} \quad (\text{C.8})$$

Podem-se relacionar os vetores de deslocamentos e de carregamentos nos sistemas de coordenadas local e global mediante a matriz de transformação, assim têm-se as seguintes expressões

$$\bar{\mathbf{u}}^e = \mathbf{T}^e \mathbf{u}^e \quad (\text{C.9})$$

e

$$\mathbf{T}^{eT} \bar{\mathbf{f}}^e = \mathbf{f}^e \quad (\text{C.10})$$

## C.3 Matrizes de rigidez e de massa

As matrizes de rigidez e massa referidas ao sistema de coordenadas local são definidas como

$$\bar{\mathbf{K}}^e = \frac{b^e E^e}{L^e} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad (\text{C.11})$$

e

$$\bar{\mathbf{M}}^e = \frac{\rho^e b^e L^e}{6} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad (\text{C.12})$$

Da mesma maneira, podem-se relacionar as matrizes de rigidez e massa nos sistemas de coordenadas local e global mediante a matriz de transformação, assim têm-se

$$\mathbf{K}^e = \mathbf{T}^{eT} \bar{\mathbf{K}}^e \mathbf{T}^e \quad (\text{C.13})$$

e

$$\mathbf{M}^e = \mathbf{T}^{eT} \bar{\mathbf{M}}^e \mathbf{T}^e \quad (\text{C.14})$$

Substituindo a Equação (C.7) nas Equações (C.13) e (C.14) temos que

$$\mathbf{K}^e = \frac{b^e E^e \mathbf{D}^e}{L^e} \quad (\text{C.15})$$

e

$$\mathbf{M}^e = \frac{\rho^e A^e L^e}{6} \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix} \quad (\text{C.16})$$

onde  $\mathbf{D}^e$  é a matriz dos cossenos diretores em relação à barra  $e$  tal que

$$\mathbf{D}^e = \begin{bmatrix} l^{2e} & l^e m^e & -l^{2e} & -l^e m^e \\ l^e m^e & m^{2e} & -l^e m^e & -m^{2e} \\ -l^{2e} & -l^e m^e & l^{2e} & l^e m^e \\ -l^e m^e & -m^{2e} & l^e m^e & m^{2e} \end{bmatrix} \quad (\text{C.17})$$

## C.4 Tensão axial

Finalmente, a tensão axial de uma barra  $e$  é determinada pela seguinte expressão

$$\sigma^e = \frac{E^e}{L^e} [-1 \ 1] \mathbf{T}^e \mathbf{u}^e \quad (\text{C.18})$$