

Isabel Cristina da Silva Martins Ribeiro

Optimização Global e Aplicações em Engenharia Estrutural



Departamento de Engenharia Civil
Faculdade de Engenharia da Universidade do Porto
2004

Isabel Cristina da Silva Martins Ribeiro

Optimização Global e Aplicações em Engenharia Estrutural



Universidade do Porto

Faculdade de Engenharia

FEUP

*Tese submetida à Faculdade de Engenharia da
Universidade do Porto para obtenção do grau de Doutor
em Ciências de Engenharia*

Departamento de Engenharia Civil
Faculdade de Engenharia da Universidade do Porto

2004



A toda a minha família, em particular à
memória do meu Pai que, sempre pre-
sente, me continua a guiar e a ajudar.

Agradecimentos

Antes de mais, quero expressar o meu mais profundo agradecimento à Professora Ana Maria Faustino e ao Professor Joaquim João Júdice, que com prontidão aceitaram o trabalho de orientação desta dissertação e cujos vastos conhecimentos muito me ajudaram. Apoiaram nos momentos mais difíceis e souberam estimular, aconselhar, criticar e orientar de forma segura e irrepreensível este trabalho até à sua conclusão. O mínimo que eu posso dizer é que foram grandes amigos.

Gostaria ainda de agradecer ao Professor Afonso Serra Neves pela colaboração e pelos profundos conhecimentos nas aplicações às formas estruturais que muito enriqueceram este trabalho.

Ao Professor Hanif Sherali uma palavra de apreço pela sua disponibilidade e pertinentes sugestões durante todo este projecto.

Agradeço do fundo do meu coração ao meu marido, Ricardo, e à minha filha, Carina, por terem compreendido o tanto tempo que não pude estar com eles e pelo constante amor e apoio que sempre me dedicaram.

Um obrigada especial à minha restante família que se ajustou às minhas necessidades.

Durante a realização deste trabalho, tive a sorte de receber sugestões, conselhos e comentários de vários colegas. Agradeço-lhes o tempo, a dedicação e a sabedoria, pois com eles muito aprendi.

A minha mais profunda estima também para todos que de qualquer modo me ajudaram a tornar este trabalho possível, e em especial para os elementos da Secção de Matemática do Departamento de Engenharia Civil da FEUP.

Resumo

O primeiro grande objectivo desta tese consiste no desenvolvimento, implementação e análise de convergência de algoritmos para a determinação de pontos estacionários e de mínimos globais para Problemas de Programação Matemática com Restrições de Equilíbrio ou de Complementaridade (MPEC). Um algoritmo de restrições activas modificado para a obtenção de um ponto estacionário de um MPEC é desenvolvido, sendo analisada a sua convergência e estudado o seu comportamento na prática. Em relação a processos de optimização global para o MPEC é desenvolvido e analisado um algoritmo de ramificação e limitação, que incorpora cortes disjuntivos para a determinação de limites inferiores e o método de restrições activas modificado para o cálculo de limites superiores. Uma nova versão de um método sequencial complementar baseada na metodologia de restrições activas é também analisada.

Como segundo objectivo deste trabalho, são estudadas algumas aplicações do MPEC. Assim é introduzido um modelo para a selecção de formas estruturais em espaços bidimensionais. Usando a técnica de reformulação-linearização é possível formular esse modelo como um programa linear inteiro misto. É ainda estabelecida uma reformulação desse problema como um MPEC e analisada a sua solução através dessa abordagem. O Problema de Ajuste de Funções Dobradiças e o Problema Complementar de Valores Próprios (EiCP) são duas outras aplicações do MPEC consideradas nesta tese. É investigada a resolução desses MPECs através das técnicas locais e globais anteriormente desenvolvidas. Finalmente é desenvolvido e analisado um algoritmo de ramificação e limitação para o EiCP baseado na sua reformulação MPEC e na técnica de reformulação-linearização.

Os algoritmos locais e globais desenvolvidos nesta tese são extensivamente testados de modo a poder atestar as suas eficiências na prática e a validade das abordagens MPEC para os problemas de aplicação discutidos nesta tese.

PALAVRAS CHAVE: Programação Não Linear, Complementaridade, Optimização Global, Optimização Estrutural, Problemas de Valores Próprios.

Abstract

The first goal of this thesis is the development, implementation and convergence analysis of algorithms for finding stationary points and global minima for Mathematical Programming Problems with Equilibrium or Complementarity Constraints (MPEC). A modified active-set method for computing a stationary point of an MPEC is developed. A convergence analysis of the algorithm is presented and its performance in practice is investigated. A branch-and-bound method is proposed for the computation of a global minimum of the MPEC. This procedure incorporates disjunctive cuts for computing of lower-bounds and the modified active-set method for finding upper-bounds. A new version of a sequential complementarity algorithm based on the active-set methodology is also discussed.

The study of some important applications of the MPEC is the second objective of this work. A new model for the selection of structural shapes in a bidimensional space is introduced. By using a reformulation-linearization technique, it is possible to formulate this model as a mixed integer linear program. A reformulation of this last problem as a MPEC is also discussed and the solution of the model by using this new formulation is analysed. A Hinge-Fitting Problem and a Complementarity Eigenvalue Problem (EiCP) are also studied as important applications of the MPEC. It is investigated the solution of these problems by using the local and global techniques for the MPEC introduced before in this thesis. Finally a new branch-and-bound method is introduced and analysed for the solution of the EiCP. This new algorithm is based on the MPEC formulation of the EiCP and on the reformulation-linearization technique.

The local and global optimization algorithms are fully tested in order to highlight their efficiency in practice and the validity of the MPEC formulations for the applications discussed in this thesis.

KEYWORDS: Nonlinear Programming, Complementarity, Global Optimization, Structural Optimization, Eigenvalue Problems.

Résumé

Le premier grand but de cette thèse consiste dans le développement, l'implémentation et l'analyse de convergence d'algorithmes pour la détermination de points stationnaires et de minimales globales pour Problèmes de Programmation Mathématique avec Restrictions d'Équilibre ou de Complémentarité (MPEC). Un algorithme avec restrictions actives modifié pour l'obtention d'un point stationnaire d'un MPEC est développé et son analyse de convergence et son comportement dans la pratique sont étudiés. En ce qui concerne le processus d'optimisation global pour le MPEC, un algorithme de ramification et limitation, qu'incorpore coupes disjonctives pour la détermination de limites inférieures et la méthode des restrictions actives modifié pour le calcul des limites supérieures est développé et analysé. Une nouvelle version d'une méthode séquentielle complémentaire basée dans la méthodologie de restrictions actives est aussi analysée.

L'étude de quelques importantes applications du MPEC est le deuxième but de ce travail. On introduit donc un nouveau modèle pour la sélection des formes structurelles dans des espaces bidimensionnels. La formulation de ce modèle comme un programme linéaire entier mixte est effectuée avec la technique de reformulation-linéarisation. Il est encore établie une reformulation de ce problème comme un MPEC et analysée sa solution par cette approche. Le problème d'ajustement de fonctions charnières et le problème complémentaire de valeurs propres (EiCP) sont autres deux applications du MPEC considérés dans cette thèse. On a enquêté la résolution de ceux MPECs par des techniques locales et globales précédemment développées. Finalement, il est encore développé et analysé un algorithme de ramification et de limitation pour l'EiCP basé dans sa reformulation MPEC et dans la technique de reformulation-linéarisation. Les algorithmes locaux et globaux développés dans cette thèse sont extensivement testés d'une façon que permet certifier leur efficacité dans la pratique et la validité des approches MPEC pour les problèmes d'application discutées dans cette thèse.

MOTS-CLÉS: Programmation Non-linéaire, Complémentarité, Optimisation Globale, Optimisation Structurale, Problèmes de Valeurs Propres.

Notações

Espaços:

$\mathbb{R}$	:	Recta real.
$\mathbb{R}^n$	:	Espaço vectorial euclideo real de dimensão n .
$\mathbb{R}_+^n$	:	Octante não negativo do espaço vectorial euclidiano $\mathbb{R}^n$.
$\mathbb{R}^{m \times n}$	:	Espaço das matrizes reais de ordem $m \times n$.

Conjuntos:

$\mathbb{N}$	:	Conjunto dos números naturais.
$[a, b]$	:	Intervalo fechado.
$ S $	:	Número de elementos de um conjunto finito S .
$\emptyset$	:	Conjunto vazio.

Escalares:

$ t $	:	Valor absoluto do número real t .
λ^k	:	Número real λ na iteração k .

Vectores:

x	:	Vector coluna.
x^T	:	Vector linha.
x_i	:	Componente i do vector x .
x_F	:	Subvector do vector x com componentes x_i , $i \in F$.
e	:	Vector com componentes iguais a um.
x^k	:	Vector x na iteração k .
$x \perp y$	:	x e y perpendiculares.
$\ x\ $	:	Norma do vector x .

Funções:

- $\nabla f(x)$: Vector gradiente de f em x .
 $C^2 (C^1)$: Conjunto das funções duas (uma) vezes continuamente diferenciáveis.

Matrizes:

- I_n : Matriz identidade de ordem n .
 e^j : j -ésima coluna da matriz identidade.
 $0_{m \times n}$: Matriz nula de dimensão $m \times n$.
 A^{-1} : Matriz inversa de A .
 A^T : Matriz transposta de A .
 a_{ij} : Elemento da linha i e coluna j da matriz A .
 A_{IJ} : Submatriz da matriz A constituída pelas linhas $i \in I$ e colunas $j \in J$.
 $A_{i.}$: Linha i da matriz A .
 $A_{.j}$: Coluna j da matriz A .
 A^k : Matriz A na iteração k .
 $\|A\|$: Norma da matriz A .
 $diag(a)$: Matriz diagonal cujos elementos diagonais são iguais às componentes do vector a .
 $det(A)$: Determinante da matriz A .

Classes de Matrizes:

- PSD : Positiva Semi-Definida.
 NSD : Negativa Semi-Definida.
 IND : Indefinida.
 PD : Positiva Definida.
 $RSuf$: Suficiente por Linhas.
 Suf : Suficiente.
 P : Matrizes com menores principais positivos.
 P_0 : Matrizes com menores principais não negativos.

Problemas:

LCP	:	Problema Linear Complementar.
mLCP	:	Problema Linear Complementar Misto.
GLCP	:	Problema Linear Complementar Generalizado.
GLCP _{JL}	:	Problema Linear Complementar Generalizado com variáveis complementares w_j , $j \in J$ e z_i , $i \in L$ fixas em zero.
NCP	:	Problema Não Linear Complementar.
CP	:	Problema Complementar.
VI	:	Desigualdade Variacional.
PNL	:	Programa Não linear.
QP	:	Programa Quadrático.
MPEC	:	Programa Matemático com Restrições de Equilíbrio (ou de Complementaridade).
EiCP	:	Problema Complementar de Valores Próprios.

Algoritmos:

ASET	:	Algoritmo de Restrições Activas.
MRG	:	Algoritmo de Gradiente Reduzido Modificado.
EMRG	:	Método Enumerativo baseado no Algoritmo de Gradiente Reduzido Modificado.
EASET	:	Método Enumerativo baseado no Algoritmo de Restrições Activas.
SLCP	:	Método Sequencial Complementar.

Índice

INTRODUÇÃO	1
1 PROBLEMAS LINEARES COMPLEMENTARES	5
1.1 Definições dos Problemas Lineares Complementares	5
1.2 Reduções a Programas Quadráticos	9
1.3 Métodos Directos para o Problema Linear Complementar	13
1.4 Métodos Iterativos para o Problema Linear Complementar	19
2 ALGORITMOS DE PROGRAMAÇÃO QUADRÁTICA E ENUMERATIVOS PARA PROBLEMAS LINEARES COMPLEMENTARES	23
2.1 Algoritmo de Gradiente Reduzido Modificado	24
2.2 Algoritmo de Restrições Activas	25
2.3 Código MINOS	31
2.4 Método Enumerativo	34
2.5 Experiência Computacional	38
3 PROBLEMA DE PROGRAMAÇÃO MATEMÁTICA COM RESTRIÇÕES DE COMPLEMENTARIDADE	47
3.1 Definição do Programa com Restrições de Complementaridade	48

3.2	Redução de um Problema de Dois Níveis a um Programa com Restrições de Complementaridade	50
3.3	Redução de um Problema Bilinear a um Programa com Restrições de Complementaridade	51
3.4	Redução de um Problema Linear Complementar a um Programa com Restrições de Complementaridade	54
3.5	Algoritmos de Programação Não Linear	55
3.6	Algoritmo de Restrições Activas Modificado	58
3.7	Convergência do Método de Restrições Activas Modificado	62
3.8	Caso Especial da Função Objectivo Quadrática	66
3.9	Implementação e Casos Degenerados	68
3.10	Experiência Computacional	72
4	ALGORITMOS PARA OPTIMIZAÇÃO GLOBAL DE PROBLEMAS DE PROGRAMAÇÃO MATEMÁTICA COM RESTRIÇÕES DE COMPLEMENTARIDADE	85
4.1	Método de Ramificação e Limitação	86
4.2	Determinação de Limites Inferiores	90
4.3	Determinação de Limites Superiores	98
4.4	Algoritmo Sequencial Complementar	104
4.5	Experiência Computacional	106
5	PROBLEMA DE AJUSTE DE FUNÇÕES DOBRADIÇAS	111
5.1	Definição do Problema	111
5.2	O Problema do Ajuste Euclideano de Funções Dobradiças	114
5.3	O Problema do Ajuste Linear de Funções Dobradiças	115

5.4	Experiência Computacional	118
6	PROBLEMA COMPLEMENTAR DE VALORES PRÓPRIOS	123
6.1	Definição do Problema Complementar de Valores Próprios	123
6.2	O Problema Complementar Simétrico de Valores Próprios	127
6.3	O Problema Complementar Não Simétrico de Valores Próprios	129
6.4	Uma Técnica de Reformulação e Linearização para o Problema Complementar de Valores Próprios	134
6.5	Escalonamento no Problema Complementar de Valores Próprios	140
6.6	Experiência Computacional	141
7	OPTIMIZAÇÃO DE FORMAS ESTRUTURAIS	147
7.1	Introdução à Optimização Estrutural	148
7.2	Um Modelo de Optimização Topológica	151
7.3	Um Exemplo de Aplicação do Modelo	156
7.4	Estruturas Cinematicamente Instáveis	160
7.5	Formulação Linear Inteira Mista	165
7.6	Formulação de Programação Matemática com Restrições de Complementaridade	172
7.7	Experiência Computacional	173
	CONCLUSÃO	183
	Referências	187

INTRODUÇÃO

A otimização global fornece estratégias e procedimentos numéricos para analisar e resolver problemas de otimização não linear na presença de múltiplos óptimos locais. Esta área tem suscitado uma intensa investigação por parte da comunidade científica, nomeadamente devido às suas inúmeras aplicações em várias áreas da ciência, economia e engenharia.

O Problema de Otimização com Restrições de Complementaridade ou de Equilíbrio (MPEC) é um problema de otimização global no qual as restrições incluem problemas complementares ou de desigualdades variacionais, que estão relacionados com fenómenos de equilíbrio que aparecem nas aplicações de engenharia e economia. Estes problemas tiveram origem nos anos sessenta num trabalho desenvolvido por K. Kirchgässner [74] e, desde então, têm sido objecto de intensa investigação. Vários algoritmos têm sido propostos para a determinação de mínimos locais e globais para os problemas.

O grande objectivo deste trabalho consiste no desenvolvimento, análise de convergência, implementação e teste de algoritmos para a resolução local e global de Problemas de Programação Matemática com Restrições de Equilíbrio. Assim é discutido em detalhe um algoritmo de restrições activas modificado para esse fim e é estabelecida a sua convergência para um ponto estacionário do MPEC. Este algoritmo é semelhante a um processo introduzido por Scholtes [108] e consiste essencialmente em utilizar uma técnica de restrições activas no conjunto de soluções do problema complementar que constitui o conjunto admissível do MPEC.

A obtenção de um mínimo global para o MPEC é um problema bastante complexo. O método Sequencial Complementar introduzido em [64] é um dos algoritmos conhecidos para esse fim. No trabalho desta tese é estudado esse processo e é desenvolvido, implementado e testado um método de ramificação e limitação que incorpora cortes disjuntivos para a determinação

de limites inferiores e o método de restrições activas modificado para a obtenção de limites superiores.

A resolução de Problemas de Optimização com Restrições de Equilíbrio necessita da existência de algoritmos de complementaridade. A Complementaridade Linear surgiu nos anos sessenta como um processo de solução do jogo de duas matrizes e do problema de programação convexa. Desde então, tem suscitado uma intensa investigação por parte da comunidade científica, nomeadamente devido às suas inúmeras aplicações em várias áreas da ciência, economia e engenharia [26, 92, 13, 51, 94, 98]. Dependendo das propriedades da matriz associada, o Problema Linear Complementar (LCP) pode ser resolvido em tempo polinomial ou ser um problema NP-difícil.

Um grande número de métodos directos e iterativos tem sido proposto para a resolução do LCP. O livro de Cottle et al [26] é uma boa referência para uma descrição dos processos que têm sido desenvolvidos para resolver o LCP. A maior parte destes métodos são apenas úteis para LCPs monótonos, isto é, problemas em que a matriz M é positiva semidefinida [120]. Uma outra forma de resolver o LCP consiste em explorar a sua redução a problemas de optimização não convexa. Os métodos enumerativos [62] são os únicos algoritmos que resolvem problemas lineares complementares sem imporem restrições aos dados. Nesta tese é apresentada uma nova implementação de um método enumerativo que é baseada num método de restrições activas.

A Engenharia Estrutural é seguramente uma das áreas de maior aplicação da optimização global e tem como objectivo principal a análise do comportamento de estruturas e das suas componentes. Esse estudo enfatiza os campos de actividades dos engenheiros estruturais e o papel dos modelos, métodos de cálculo e procedimentos computacionais. Com a tendência de se projectar estruturas mais arrojadas tem-se tornado cada vez mais necessário o desenvolvimento técnico-científico na área da engenharia estrutural que permita obter projectos óptimos em medidas importantes para o projectista.

Um dos objectivos deste trabalho de doutoramento consiste no desenvolvimento e na resolução de um modelo para a selecção de formas estruturais desenvolvidas em espaços bidimensionais. Como, de um modo geral, a regulamentação de segurança aplicável às estruturas metálicas na área da Engenharia Civil [1, 2] não admite a utilização de estruturas hipostáticas ou cinematicamente instáveis foram desenvolvidas duas técnicas alternativas para evitar esse tipo de estruturas.

As formulações matemáticas desse modelo conduzem a programas não lineares inteiros mistos com uma função objectivo bilinear e restrições bilineares e lineares. Reformulações desses problemas como Problemas de Programação Linear Inteira Mista, recorrendo à técnica de reformulação-linearização (RLT) e como Problemas de Programação Matemática com Restrições de Complementaridade ou de Equilíbrio (MPEC) são discutidas de modo a obter uma solução óptima para o modelo.

O Problema de Ajuste de Funções Dobradiças, que aparece na teoria de aproximação [15, 99] e o Problema Complementar de Valores Próprios (EiCP), que teve origem no estudo dos estados de equilíbrio estático de sistemas mecânicos com atrito unilateral [24], são duas outras aplicações a serem investigadas nesta tese. Esses problemas são reduzidos a MPECs e resolvidos usando os algoritmos locais e globais desenvolvidos neste trabalho.

Um algoritmo de ramificação e limitação para a resolução do EiCP é também introduzido nesta tese. O método explora a técnica de reformulação-linearização de modo a determinar limites inferiores e superiores. A análise de convergência, implementação e teste desse algoritmo são ainda apresentados neste trabalho.

Na organização desta tese, no capítulo 1 são introduzidos os Problemas Lineares Complementares que são objecto de estudo nos capítulos seguintes. A sua redução a programas quadráticos é estudada na secção 1.2. Uma breve discussão dos métodos directos e iterativos e das suas limitações é apresentada nas secções 1.3 e 1.4, respectivamente.

No capítulo 2 é discutido um Método Enumerativo para o LCP. Esse processo pode incorporar um método de Gradiente Reduzido Modificado ou um algoritmo de Restrições Activas, que são discutidos em detalhe nas secções 2.1 e 2.2, respectivamente. O código de restrições activas MINOS usado na implementação do segundo algoritmo é descrito de uma forma sucinta na secção 2.3. O método enumerativo e resultados computacionais com esse algoritmo na resolução de LCPs NP-difíceis são apresentados nas duas últimas secções.

O capítulo 3 é dedicado aos Problemas de Programação Matemática com Restrições de Complementaridade ou de Equilíbrio (MPEC). A definição do MPEC e as reduções de Programas de Dois Níveis, Programas Bilineares e do Problema Linear Complementar a MPECs são apresentadas nas secções posteriores. Na secção 3.5 são realçadas algumas técnicas locais, mais referenciadas na literatura, para a determinação de um ponto estacionário de um

MPEC. Na secção 3.6 é discutido em detalhe um algoritmo de restrições activas modificado desenvolvido neste trabalho para esse fim. Finalmente alguns resultados computacionais obtidos na resolução de MPECs com esse algoritmo são incluídos na última secção.

No capítulo 4 são discutidos algoritmos para a determinação de um mínimo global para o MPEC. Nas secções 4.1, 4.2 e 4.3 é apresentado um método de ramificação e limitação que incorpora cortes disjuntivos para a determinação de limites inferiores e o método de restrições activas modificado para a obtenção de limites superiores. O método Sequencial Complementar introduzido em [64] é descrito na secção 4.4. Alguma experiência computacional com estes algoritmos na resolução de MPECs é apresentada na última secção deste capítulo.

No capítulo 5 é apresentado, de uma forma muito sucinta, o Problema de Ajuste de Funções Dobradiças. As suas reformulações como um MPEC com uma função quadrática convexa e um MPEC com uma função linear são apresentadas na secções 5.2 e 5.3, respectivamente. A secção 5.4 inclui resultados computacionais com os métodos de optimização global para o MPEC discutidos no capítulo 4 para a resolução destes problemas.

No capítulo 6 é estudado o Problema Complementar de Valores Próprios (EiCP). O caso simétrico é apresentado de uma forma muito sucinta na secção 6.2. A forma não simétrica do problema é discutida nas restantes secções. Em particular, a redução do EiCP a um Problema Não Linear Complementar e a um MPEC são apresentadas na secção 6.3. Um algoritmo de optimização global de Ramificação e Limitação é introduzido na secção 6.4. A secção 6.5 é dedicada ao escalonamento do problema. Experiência computacional com os vários algoritmos é incluída na última secção.

No capítulo 7 é apresentado um modelo para a selecção de formas estruturais desenvolvidas em espaços bidimensionais. Na secção 7.1 é apresentado um breve resumo dos trabalhos desenvolvidos nesta área. Na secção 7.2 é introduzida uma formulação detalhada do modelo e um exemplo de aplicação é apresentado na secção 7.3. A definição das estruturas cinematicamente instáveis e as metodologias desenvolvidas para contornar essa dificuldade são apresentadas na secção 7.4. As duas formulações do modelo como um Programa Linear Inteiro Misto e um MPEC são descritas nas secções 7.5 e 7.6, respectivamente. Alguma experiência computacional é apresentada na última secção deste capítulo.

As conclusões mais importantes do trabalho de investigação realizado e algumas áreas de investigação futura são apresentadas na última parte desta tese de doutoramento.

Capítulo 1

PROBLEMAS LINEARES COMPLEMENTARES

Neste capítulo são apresentadas as definições do Problema Linear Complementar (LCP) e de algumas das suas extensões, assim como alguns resultados de existência e unicidade relacionados com a classe da matriz do problema. Os métodos directos e iterativos mais importantes para a resolução do LCP são também sumariamente discutidos.

1.1 Definições dos Problemas Lineares Complementares

Os Problemas Complementares distinguem-se dos restantes problemas por não possuírem função objectivo a optimizar e por conterem um conjunto de restrições de complementaridade, que exigem que o produto de duas variáveis não negativas seja nulo. Essas restrições apareceram pela primeira vez nas condições de optimalidade para programas não lineares com variáveis contínuas envolvendo restrições de desigualdade [10, 82]. No entanto têm-se mostrado de importância fundamental em formulações de problemas de equilíbrio e de outros modelos de economia, engenharia e ciência [26, 92, 35]. O problema complementar mais simples e amplamente estudado é o Problema Linear Complementar assim definido:

Definição 1.1 Dado o vector $q \in \mathbb{R}^n$ e a matriz $M \in \mathbb{R}^{n \times n}$, o **Problema Linear Complementar (LCP)** consiste em encontrar (ou concluir que não existem) vectores $z \in \mathbb{R}^n$ e $w \in \mathbb{R}^n$ tais que

$$w = q + Mz, \quad z \geq 0, \quad w \geq 0 \quad (1.1)$$

$$z^T w = 0 \quad (1.2)$$

Este problema tem sido várias vezes descrito como um problema fundamental, pois traduz as condições necessárias de optimalidade de primeira ordem para um Programa Quadrático com desigualdades e variáveis não negativas. O LCP é normalmente representado por $\text{LCP}(q, M)$, uma vez que o vector q e a matriz M caracterizam-no univocamente.

O nome deste problema é bastante sugestivo e é consequência de ser constituído por restrições lineares (1.1) e por uma condição não linear de complementaridade (1.2). Esta última condição exige que pelo menos uma das variáveis de cada par (z_i, w_i) , $i = 1, \dots, n$ seja igual a zero na solução do problema. O par (z_i, w_i) é designado por o *i-ésimo par de variáveis complementares* do problema e as variáveis z_i e w_i dizem-se *complementares* uma da outra.

Uma pequena generalização do LCP é o Problema Linear Complementar Misto assim definido:

Definição 1.2 Dados os vectores $q \in \mathbb{R}^n$, $p \in \mathbb{R}^t$ e as matrizes $M \in \mathbb{R}^{n \times n}$, $N \in \mathbb{R}^{n \times m}$, $R \in \mathbb{R}^{t \times n}$ e $S \in \mathbb{R}^{t \times m}$, o **Problema Linear Complementar Misto (mLCP)** consiste em encontrar (ou concluir que não existem) vectores $z \in \mathbb{R}^n$, $w \in \mathbb{R}^n$ e $y \in \mathbb{R}^m$ tais que

$$\begin{aligned} w &= q + Mz + Ny \\ 0 &= p + Rz + Sy \\ z &\geq 0, w \geq 0 \\ z^T w &= 0 \end{aligned} \quad (1.3)$$

As condições necessárias de optimalidade de primeira ordem para um programa quadrático envolvendo restrições de igualdade e desigualdade são desta forma [10]. Se em (1.3), $t = m$ e S é não singular, então y pode ser eliminado do problema usando $y = -S^{-1}(p + Rz)$ e o problema transforma-se num $\text{LCP}(q - NS^{-1}p, M - NS^{-1}R)$.

Em muitas aplicações de economia e engenharia existem situações em que as variáveis estão relacionadas entre si por restrições lineares, mas a condição de complementaridade apenas ocorre entre algumas dessas variáveis. Assim aparece o denominado Problema Linear Complementar Generalizado [122] cuja definição é apresentada a seguir.

Definição 1.3 Dado o vector $q \in \mathbb{R}^l$ e as matrizes $E \in \mathbb{R}^{l \times n}$, $D \in \mathbb{R}^{l \times n}$ e $C \in \mathbb{R}^{l \times m}$, o **Problema Linear Complementar Generalizado (GLCP)** consiste em encontrar (ou concluir que não existem) vectores $w \in \mathbb{R}^n$, $z \in \mathbb{R}^n$ e $y \in \mathbb{R}^m$ tais que

$$\begin{aligned} Ew &= q + Dz + Cy \\ w &\geq 0, z \geq 0, y \in K_y \\ w^T z &= 0 \end{aligned} \tag{1.4}$$

com K_y um poliedro convexo de $\mathbb{R}^m$ da forma

$$K_y = \{y \in \mathbb{R}^m : Ay = b, y \geq 0\}$$

onde $A \in \mathbb{R}^{p \times m}$ e $b \in \mathbb{R}^p$.

Notar que um mLCP é um caso particular do GLCP em que $K_y = \mathbb{R}^m$.

Ao longo deste trabalho, será referenciado o caso particular do GLCP, que consiste em considerar $E = I_n$ e D uma matriz quadrada de ordem n . Este problema será denotado por mGLCP.

A existência das restrições suplementares e das variáveis y que não satisfazem a condição de complementaridade torna, de um modo geral, o GLCP mais difícil de resolver do que o LCP. Não obstante, este problema tem sido amplamente estudado por muitos autores [12, 44, 64, 67, 85] com o objectivo de resolver alguns problemas de optimização global.

Apresentaremos em seguida algumas generalizações não lineares do LCP, que irão ser utilizadas nos próximos capítulos. Uma das mais importantes dessas extensões é o Problema de Desigualdades Variacionais que, na sua forma geral, é assim definido:

Definição 1.4 Dado um subconjunto fechado K do espaço euclidiano de dimensão n , $\mathbb{R}^n$, e uma função $F : K \rightarrow \mathbb{R}^n$ contínua em $\mathbb{R}^n$, um **Problema de Desigualdades Variacionais**, designado por $VI(K, F)$, consiste em encontrar um vector $x \in K$ tal que

$$(y - x)^T F(x) \geq 0, \quad \forall y \in K. \tag{1.5}$$

O conjunto de soluções para este problema é denotado por $SOL(K, F)$.

A interpretação geométrica da definição de $\text{VI}(K, F)$ estabelece que x pertence a $\text{SOL}(K, F)$ se e só se $F(x)$ forma um ângulo não obtuso com os vectores da forma $y - x$ para todo o y em K . Se K é um cone, isto é, se a implicação $x \in K \Rightarrow \lambda x \in K$ é verdadeira para todo $\lambda \geq 0$, então [35] o VI é equivalente ao Problema Complementar definido a seguir.

Definição 1.5 *Dado um cone K e uma função $F : K \rightarrow \mathbb{R}^n$, o Problema Complementar, designado por $\text{CP}(K, F)$, consiste em encontrar um vector $x \in \mathbb{R}^n$ que satisfaz as seguintes condições*

$$K \ni x \perp F(x) \in K^* \quad (1.6)$$

onde a notação $\perp$ significa perpendicular e K^* é o cone dual de K definido como

$$K^* = \{d \in \mathbb{R}^n : v^T d \geq 0 \quad \forall v \in K\} \quad (1.7)$$

isto é, K^* contém os vectores que fazem um ângulo não obtuso com todos os vectores de K (Figura 1.1).

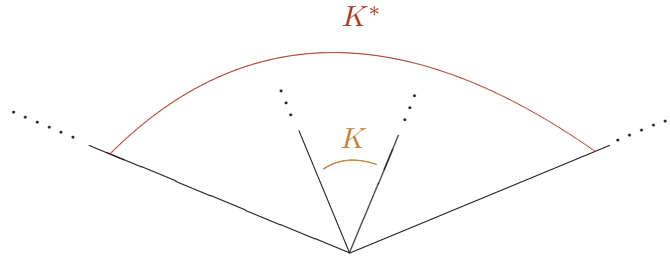


Figura 1.1: Cone e seu dual

Explorando o significado de $\perp$ na condição (1.6), o $\text{CP}(K, F)$ pode ser reescrito na seguinte forma:

$$x \in K, \quad F(x) \in K^* \quad \text{e} \quad x^T F(x) = 0 \quad (1.8)$$

Um dos casos especiais de Problemas de Complementaridade é o $\text{CP}(\mathbb{R}_+^n, F)$, onde $\mathbb{R}_+^n$ é o octante não negativo de $\mathbb{R}^n$. Neste caso o CP é normalmente conhecido por Problema Não Linear Complementar e é designado por $\text{NCP}(F)$. Como o cone dual de $\mathbb{R}_+^n$ é o próprio $\mathbb{R}_+^n$, então $\text{NCP}(F)$ é definido da seguinte forma:

Definição 1.6 *Dada uma função $F : \mathbb{R}_+^n \rightarrow \mathbb{R}^n$, o Problema Não Linear Complementar ($NCP(F)$) consiste em encontrar um vector $x \in \mathbb{R}^n$ que satisfaz*

$$0 \leq x \perp F(x) \geq 0 \quad (1.9)$$

Convém salientar que, se $F(x)$ é uma função afim $q + Mx$, então o $NCP(F)$ reduz-se ao $LCP(q, M)$. Na próxima secção iremos enunciar e demonstrar alguns resultados importantes que serão fundamentais para o desenvolvimento de alguns dos algoritmos que têm vindo a ser referenciados na literatura para a resolução de Problemas Complementares.

1.2 Reduções a Programas Quadráticos

Tal como foi referido anteriormente, dada uma matriz quadrada M de ordem n e um vector $q \in \mathbb{R}^n$, o Problema Linear Complementar, notado por LCP ou $LCP(q, M)$, procura vectores $z \in \mathbb{R}^n$ e $w \in \mathbb{R}^n$ tais que

$$w = q + Mz, \quad z \geq 0, \quad w \geq 0 \quad (1.10)$$

$$z^T w = 0 \quad (1.11)$$

Portanto o LCP contém restrições lineares (1.10), que constituem o chamado conjunto admissível

$$K = \{(z, w) : w = q + Mz, z \geq 0, w \geq 0\} \quad (1.12)$$

e a condição não linear de complementaridade (1.11). A solução (z, w) é *admissível* se pertencer ao conjunto admissível K . Por outro lado é designada de *complementar* se satisfaz

$$\begin{aligned} w &= q + Mz \\ z_i w_i &= 0 \quad i = 1, 2, \dots, n \end{aligned}$$

Devido às condições de não negatividade das variáveis z_i e w_i , (z, w) é uma solução do LCP se e só se for admissível e complementar. Além disso, o LCP não tem solução se é inadmissível ($K = \emptyset$) ou se é admissível mas não tem solução complementar em K , isto é, $z^T w > 0$ para todo $(z, w) \in K$. A existência e unicidade da solução do LCP é usualmente

apresentada em termos da classe da matriz M [26, 92]. As chamadas matrizes Positivas Semi-Definidas (PSD) e P são provavelmente as classes que têm um papel mais importante na teoria e resolução de problemas de complementaridade [26, 92]. As definições dessas classes de matrizes são apresentadas a seguir.

Definição 1.7 *Seja M uma matriz quadrada de ordem n .*

- $M \in PSD \Leftrightarrow x^T M x \geq 0$ para todo $x \in \mathbb{R}^n$
- $M \in P \Leftrightarrow$ todos os menores principais de M são positivos

O seguinte teorema estabelece a existência e a unicidade da solução do LCP quando M pertence a estas duas classes [26, 92].

Teorema 1.1

- (i) $M \in P$ se e só se o $LCP(q, M)$ tem uma única solução para cada $q \in \mathbb{R}^n$.
- (ii) Se $M \in PSD$, então o $LCP(q, M)$ tem pelo menos uma solução para cada q tal que $K \neq \emptyset$, isto é, o $LCP(q, M)$ tem sempre solução se for admissível.
- (iii) Se $M \in PSD$, então o $LCP(q, M)$ pode ser resolvido em tempo polinomial.

Alguns destes resultados podem ser generalizados a outras classes de matrizes [26, 92]. Neste trabalho, para além das já mencionadas, iremos apenas considerar as seguintes classes de matrizes:

Definição 1.8

- $M \in P_0 \Leftrightarrow$ todos os menores principais de M são não negativos
 - $M \in PD \Leftrightarrow x^T M x > 0$ para todo $x \in \mathbb{R}^n \setminus \{0\}$
 - $M \in NSD \Leftrightarrow x^T M x \leq 0$ para todo $x \in \mathbb{R}^n$
 - $M \in IND \Leftrightarrow M \notin PSD$ e $M \notin NSD$
 - $M \in RSuf \Leftrightarrow [z_i(M^T z)_i \leq 0 \text{ para todo } i \in \{1, \dots, n\}]$
- $\Downarrow$
- $[z_i(M^T z)_i = 0 \text{ para todo } i \in \{1, \dots, n\}]$
- $M \in Suf \Leftrightarrow M \in RSuf$ e $M^T \in RSuf$

As relações entre estas classes de matrizes são apresentadas no seguinte diagrama, onde cada seta significa inclusão estrita

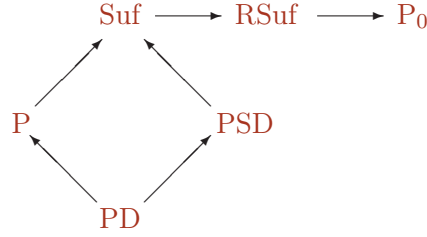


Figura 1.2: Diagrama de Inclusão das Classes de Matrizes

Podem ainda estabelecer-se as seguintes propriedades [26, 92]:

- $PD = P$ e $PSD = P_0$ para matrizes simétricas.
- Se $M \in RSuf$, então todo o LCP admissível tem solução.

Da definição do LCP, é imediato concluir que este problema é equivalente ao programa quadrático

$$\begin{aligned}
 QP \quad & \text{Minimizar} \quad g_1(z, w) = z^T w \\
 & \text{sujeito a} \quad w - Mz = q \\
 & \quad \quad \quad z \geq 0, \quad w \geq 0
 \end{aligned}$$

com $q \in \mathbb{R}^n$, $M \in \mathbb{R}^{n \times n}$, $z \in \mathbb{R}^n$ e $w \in \mathbb{R}^n$. Uma vez que, a função objectivo do QP é limitada inferiormente no conjunto K (1.12), existem três casos possíveis:

- $K = \emptyset$ e o LCP é inadmissível.
- $g_1(\bar{z}, \bar{w}) = \min_{(z, w) \in K} g_1(z, w) = 0$ e $(\bar{z}, \bar{w})$ é solução do LCP.
- $\min_{(z, w) \in K} g_1(z, w) > 0$ e o LCP é admissível mas não tem solução.

Deste modo o LCP tem solução se e só se existe uma solução óptima do programa quadrático com o valor da função objectivo igual a zero. Se substituirmos na definição do QP anterior w pelo seu valor, então o LCP é também equivalente ao seguinte programa quadrático

$$\begin{aligned}
 QP_{LCP} \quad & \text{Minimizar} \quad g_2(z) = q^T z + \frac{1}{2} z^T (M + M^T) z \\
 & \text{sujeito a} \quad q + Mz \geq 0, \quad z \geq 0
 \end{aligned}$$

Como

$$M \in PSD(NSD) \Leftrightarrow (M + M^T) \in PSD(NSD)$$

e g_2 é convexa (côncava) em $\mathbb{R}^n$ se e só se $M + M^T$ é PSD (NSD), então podemos definir LCPs convexos, côncavos e não convexos a partir das seguintes equivalências:

- LCP convexo $\Leftrightarrow M \in PSD$
- LCP côncavo $\Leftrightarrow M \in NSD$
- LCP não convexo $\Leftrightarrow M \in IND$

Devido à sua relação com o VI, um LCP convexo também se diz monótono [35]. Além disso tem-se:

$$\text{LCP estritamente convexo (monótono)} \Leftrightarrow M \in PD$$

É ainda de notar que, de acordo com o teorema 1.1, o caso (c) anterior não pode acontecer para LCPs monótonos e que apenas o caso (b) é válido para LCPs estritamente monótonos. Como foi mencionado na secção anterior, neste trabalho vamos considerar um caso particular do GLCP(1.4) que consiste em encontrar (ou concluir que não existem) vectores $z \in \mathbb{R}^n$, $w \in \mathbb{R}^n$ e $y \in \mathbb{R}^m$ tais que

$$\begin{aligned} GLCPm \quad & w = q + Mz + Ny \\ & 0 = -b + Ay \\ & z^T w = 0 \\ & z, w, y \geq 0 \end{aligned} \tag{1.13}$$

com $q \in \mathbb{R}^n$, $b \in \mathbb{R}^p$ e as matrizes $M \in \mathbb{R}^{n \times n}$, $N \in \mathbb{R}^{n \times m}$ e $A \in \mathbb{R}^{p \times m}$. Neste caso as matrizes D e E do GLCP (1.4) satisfazem

$$DE^T = M \in \mathbb{R}^{n \times n}.$$

De acordo com um resultado de Ye [122], o GLCPm pode ser resolvido em tempo polinomial desde que $M \in PSD$. Tal como no LCP podemos associar ao GLCPm (1.13) o seguinte programa quadrático

$$\begin{aligned}
& \text{QP}_{GLCPm} \\
& \text{Minimizar} \quad g_3(z, y) = \frac{1}{2} \begin{bmatrix} z \\ y \end{bmatrix}^T \begin{bmatrix} M + M^T & N \\ N^T & 0 \end{bmatrix} \begin{bmatrix} z \\ y \end{bmatrix} + \begin{bmatrix} q \\ 0 \end{bmatrix}^T \begin{bmatrix} z \\ y \end{bmatrix} \\
& \text{sujeito a} \quad Mz + Ny \geq -q \\
& \quad \quad \quad Ay = b \\
& \quad \quad \quad z, y \geq 0
\end{aligned}$$

com as mesmas conclusões (i), (ii) e (iii) apresentadas para o LCP.

Um ponto $(z, y, \alpha, \beta, \mu, \gamma)$ é designado de Ponto Estacionário ou de Karush-Kuhn-Tucker (KKT) do programa QP_{GLCPm} se (z, y) for admissível para o QP_{GLCPm} , e além disso, satisfizer

$$q + (M + M^T)z + Ny = M^T \alpha + \beta \quad (1.14)$$

$$\begin{aligned}
N^T z &= N^T \alpha + A^T \mu + \gamma \\
\alpha^T (Mz + Ny + q) &= 0
\end{aligned} \quad (1.15)$$

$$\alpha \geq 0$$

$$\beta^T z = 0$$

$$\beta \geq 0$$

$$\gamma^T y = 0$$

$$\gamma \geq 0$$

O seguinte teorema [69] fornece uma condição suficiente para que um ponto estacionário do programa QP_{GLCPm} seja solução do GLCP.

Teorema 1.2 *Seja M uma matriz suficiente por linhas ($M \in RSuf$) de ordem n .*

Se $(\bar{z}, \bar{y}, \bar{\alpha}, \bar{\beta}, \bar{\mu}, \bar{\gamma})$ é um ponto estacionário do programa quadrático QP_{GLCPm} então $(\bar{z}, \bar{y})$ é uma solução do GLCP.

Como o LCP é um caso particular do GLCP, então o teorema anterior é também válido para o LCP. Esse resultado aparece em [26], onde também é demonstrado que se $M \in P$, então para cada $q \in \mathbb{R}^n$, o ponto estacionário de QP existe sempre e é o seu único mínimo global.

1.3 Métodos Directos para o Problema Linear Complementar

Estes processos procuram obter uma solução exacta do LCP num número finito de iterações utilizando conceitos semelhantes aos do método simplex para a programação linear. Uma

breve discussão sobre alguns conceitos essenciais para a compreensão deste tipo de algoritmos é apresentada a seguir. Consideremos o LCP e escrevamos as suas restrições lineares na forma

$$Ax = q \quad (1.16)$$

$$x \geq 0 \quad (1.17)$$

com $A = [I_n \mid -M] \in \mathbb{R}^{n \times 2n}$ e $x = (w^T, z^T)^T$. Uma solução x do LCP diz-se *Básica* se for possível encontrar uma partição da matriz A em duas submatrizes B e N tais que:

- $B = \{A_{.j_1}, A_{.j_2}, \dots, A_{.j_n}\}$ é não singular, denominada *Base* associada a x ,
- as componentes de x associadas à submatriz N são iguais a zero.

Deste modo, numa solução básica x podem distinguir-se dois tipos de variáveis

- *Variáveis Básicas* (x_P) : cujos valores são obtidos resolvendo o sistema

$$Bx_P = q, \text{ com } P = \{j_1, j_2, \dots, j_n\}.$$
- *Variáveis Não Básicas* (x_T) : cujos valores são nulos, ou seja,

$$x_i = 0, \quad i \in T = \{1, 2, \dots, 2n\} \setminus P.$$

Dada uma solução básica x , o sistema (1.16) pode ser reescrito na seguinte forma

$$x_P = B^{-1}q + B^{-1}Nx_T$$

ou ainda, por

$$t = b + Cu, \quad (1.18)$$

com $t = x_P$, $b = B^{-1}q$, $C = B^{-1}N$ e $u = x_T$. Uma *Operação Pivotal Simples* consiste em transformar uma solução básica x numa outra solução básica $\bar{x}$ por troca de uma variável não básica com uma variável básica, de acordo com os seguintes passos:

1. Resolver a equação de ordem p associada à variável básica t_p do sistema (1.18) em relação à variável não básica u_m ;
2. Substituir nas outras equações, essa variável u_m pela expressão obtida.

Se escrevermos (1.18) na seguinte forma tabular

<i>Variáveis</i> <i>básicas</i>	<i>Valor</i>	<i>Variáveis não básicas</i>				
		...	u_m	...	u_j	...
$\vdots$	$\vdots$		$\vdots$		$\vdots$	
t_p	b_p	...	c_{pm}	...	c_{pj}	...
$\vdots$	$\vdots$		$\vdots$		$\vdots$	
t_i	b_i	...	c_{im}	...	c_{ij}	...
$\vdots$	$\vdots$		$\vdots$		$\vdots$	

e se efectuarmos uma operação pivotal que troque as variáveis t_p e u_m , então obtemos a seguinte forma para a nova solução básica:

<i>Variáveis</i> <i>básicas</i>	<i>Valor</i>	<i>Variáveis não básicas</i>				
		...	t_p	...	u_j	...
$\vdots$	$\vdots$		$\vdots$		$\vdots$	
u_m	$-b_p/c_{pm}$	...	$1/c_{pm}$	...	$-c_{pj}/c_{pm}$	...
$\vdots$	$\vdots$		$\vdots$		$\vdots$	
t_i	$b_i - b_p c_{im}/c_{pm}$	...	c_{im}/c_{pm}	...	$c_{ij} - c_{im} c_{pj}/c_{pm}$	...
$\vdots$	$\vdots$		$\vdots$		$\vdots$	

com, $i \in \{1, \dots, n\} \setminus \{p\}$ e $j \in \{1, \dots, n\} \setminus \{m\}$. Convém salientar que, para que esta operação pivotal simples seja possível, c_{pm} tem que ser diferente de zero. Este elemento designa-se por *pivot* da operação. Se $p = m$ então o pivot pertence à diagonal da matriz C e a operação pivotal simples diz-se *Principal*.

Esta operação pode ser efectuada usando pivots com mais do que um elemento. Neste caso, seja F um subconjunto de $\{1, \dots, n\}$ e $G = \{1, \dots, n\} \setminus F$. O sistema (1.18) pode escrever-se na seguinte forma

<i>Variáveis</i> <i>Básicas</i>	<i>Valor</i>	<i>Variáveis não básicas</i>	
		u_F	u_G
t_F	b_F	C_{FF}	C_{FG}
t_G	b_G	C_{GF}	C_{GG}

Se a submatriz principal C_{FF} é não singular, é possível efectuar uma *Operação Pivotal Principal por Blocos* com *Pivot* C_{FF} , que consiste em transformar o sistema anterior no seguinte sistema equivalente

<i>Variáveis básicas</i>	<i>Valor</i>	<i>Variáveis não básicas</i>	
		t_F	u_G
u_F	$-C_{FF}^{-1}b_F$	C_{FF}^{-1}	$-C_{FF}^{-1}C_{FG}$
t_G	$b_G - C_{GF}C_{FF}^{-1}b_F$	$C_{GF}C_{FF}^{-1}$	$C_{GG} - C_{GF}C_{FF}^{-1}C_{FG}$

Convém salientar que uma operação pivotal principal por blocos é equivalente a $|F|$ operações pivotais simples, com $|F|$ o número de elementos de F .

Os métodos directos para o LCP são métodos pivotais que, como o próprio nome indica, se baseiam em operações pivotais. Seguidamente iremos apresentar descrições sucintas dos chamados métodos de Lemke e pivotais principais. Sugerimos [26] para uma descrição mais detalhada destes algoritmos e de outros processos semelhantes.

ALGORITMO DE LEMKE

O método de Lemke [83] é de todos os métodos directos aquele que se pode aplicar a uma classe mais vasta de matrizes [26, 92]. Neste algoritmo são introduzidos uma variável $z_0 \geq 0$ e um vector $d \geq 0$ de modo a obter um LCP alargado da forma

$$\begin{aligned} w &= q + Mz + dz_0 \\ z^T w &= 0 \\ z &\geq 0, \quad w \geq 0, \quad z_0 \geq 0 \end{aligned} \tag{1.19}$$

Devido à escolha do vector d , existe um valor positivo z_0^* de z_0 tal que o sistema (1.19) tem solução $z = 0$, $w = q + dz_0^*$ e $z_0 = z_0^*$. Além disso esse valor z_0^* é facilmente determinado a partir de

$$z_0^* = -\frac{q_r}{d_r} = \max \left\{ -\frac{q_i}{d_i} : q_i < 0 \right\}$$

Se efectuarmos uma operação pivotal com o pivot d_r , então obtém-se um sistema na forma

$$\bar{w} = \bar{q} + \bar{A}\bar{z} \tag{1.20}$$

com $\bar{q} \geq 0$, $(\bar{w}, \bar{z})$ uma permutação de (w, z, z_0) e $\bar{q}$ e $\bar{A}$ são obtidos de q e de $[M \mid d]$ através da operação pivotar com pivot d_r .

O método de Lemke é um algoritmo pivotar não principal que utiliza apenas soluções complementares admissíveis de (1.19). Em cada iteração usa soluções básicas do sistema (1.19) até obter uma solução desse sistema com $z_0 = 0$ ou mostrar que tal não é possível. Os passos deste algoritmo são apresentados a seguir.

Inicialização:

Passo 0: Se $q \geq 0$ então $z = 0$ e $w = q$ é solução do LCP. Caso contrário determine o índice r tal que:

$$-\frac{q_r}{d_r} = \max \left\{ -\frac{q_i}{d_i} : q_i < 0 \right\}$$

e efectue uma operação pivotar com pivot d_r . Nesta iteração a variável z_0 passa a básica por troca com a variável básica w_r .

Em cada iteração:

Passo 1: Seja s o índice da coluna da matriz $\bar{A}$ definida em (1.20) e associada à variável complementar da variável que passou a não básica na iteração anterior.

Passo 2: Se $\bar{a}_{is} \geq 0, \forall i \in \{1, \dots, n\}$, o algoritmo termina numa aresta ilimitada. Caso contrário, determine r tal que

$$-\frac{\bar{q}_r}{\bar{a}_{rs}} = \min \left\{ -\frac{\bar{q}_i}{\bar{a}_{is}} : \bar{a}_{is} < 0 \right\}$$

Passo 3: Efectue uma operação pivotar com o pivot $\bar{a}_{rs}$.

- Se $z_0 = 0$ termine com a solução $(\bar{w} = \bar{q}, \bar{z} = 0)$.
- Se $z_0 > 0$ volte ao Passo 1.

Este algoritmo tem convergência finita desde que as soluções básicas utilizadas durante a resolução do LCP sejam todas não degeneradas, isto é, $\bar{q}_i > 0$ para todo $i = 1, \dots, n$ [92]. Convém salientar que este algoritmo pode ter duas terminações possíveis, a saber, uma aresta ilimitada ou uma solução do LCP.

De um modo geral, se o algoritmo terminar numa aresta ilimitada, não significa obrigatoriamente que o LCP não tenha solução. Por este motivo, a sua aplicabilidade é restrita, sendo normalmente utilizado para resolver LCPs cujas matrizes pertencem a determinadas classes [26, 92]. Assim, se $M \in P$ (e também PD) o método de Lemke termina sempre com uma solução do LCP. Além disso, se $M \in RSuf$ (e também PSD) as duas terminações são possíveis, mas se o algoritmo terminar numa aresta ilimitada, o LCP é inadmissível e consequentemente não tem solução [26, 92].

ALGORITMOS PIVOTAIS PRINCIPAIS

Os algoritmos pivotaís principais são processos que se baseiam em operações pivotaís principais simples ou por blocos. Em cada iteração utilizam soluções complementares (isto é, $w = q + Mz$ e $w_i z_i = 0$, $i = 1, \dots, n$) e terminam quando obtêm uma solução admissível, isto é, quando $z \geq 0$ e $w \geq 0$. Ao contrário do que se passa com o método de Lemke, estes métodos não incorporam qualquer variável artificial.

Suponhamos que as submatrizes principais de M são todas não singulares. Isso acontece por exemplo se $M \in P$ (e também se $M \in PD$). Então M_{FF} é não singular para todo o $F \subseteq \{1, \dots, m\}$ e uma solução complementar associada a F é dada por:

$$\bar{z} = \begin{bmatrix} \bar{z}_F \\ 0 \end{bmatrix}, \quad \bar{w} = \begin{bmatrix} 0 \\ \bar{w}_G \end{bmatrix}$$

onde $G = \{1, \dots, n\} \setminus F$ e $\bar{z}_F$ e $\bar{w}_G$ satisfazem

$$M_{FF} \bar{z}_F = -q_F$$

$$\bar{w}_G = q_G + M_{GF} \bar{z}_F$$

Seja $(\bar{w}, \bar{z})$ uma solução básica complementar e consideremos o conjunto das inadmissibilidades

$$H = \{i \in F : \bar{z}_i < 0\} \cup \{j \in G : \bar{w}_j < 0\}$$

O número de elementos de H é denominado de número de inadmissibilidades associado à solução $(\bar{w}, \bar{z})$. É óbvio que $(\bar{w}, \bar{z})$ é uma solução do LCP se e só se $|H| = 0$. Em cada

iteração do algoritmo pivotal principal, um conjunto $H_1 \subset H$ é considerado e o conjunto F é modificado da seguinte forma

$$F = F \setminus (F \cap H_1) \cup (G \cap H_1)$$

Se em cada iteração $|H_1| = 1$ então o algoritmo diz-se Simples. Caso contrário o algoritmo diz-se por Blocos. Quando no algoritmo pivotal principal consideramos

$$H_1 = \{\min \{i \in H\}\}$$

obtemos o chamado método de Murty [90], que tem convergência finita desde que a matriz M seja P [90, 92]. Em particular o algoritmo pode ser usado para processar LCPs estritamente monótonos, isto é, LCPs cuja matriz M é PD . Outros algoritmos pivotaes simples têm vindo a ser sugeridos na literatura, tais como os métodos de Graves [53], de Dantzig-Cottle [28, 25] e de Keller [73]. Como F só é alterado num elemento por iteração então a sua convergência pode tornar-se muita lenta para LCPs com um número elevado de variáveis. Por outro lado, se considerarmos $H_1 = H$ em todas as iterações, obtemos um algoritmo pivotal principal desenvolvido por Kostreva [81], que não possui convergência assegurada. Júdice e Pires [68] desenvolveram um método híbrido que utiliza operações pivotaes por blocos e simples e possui convergência finita. Experiência computacional [68] indica que esse processo é bastante eficiente para a resolução de LCPs com matrizes P de ordem elevada.

1.4 Métodos Iterativos para o Problema Linear Complementar

Um método iterativo começa com uma estimativa inicial x^0 da solução $x = (w, z)$ do LCP e gera uma sucessão de pontos $x^1, x^2, \dots, x^k, \dots$, obtida através de uma determinada fórmula. Esta sequência de pontos em certas condições converge para a solução do LCP, isto é,

$$\lim_{k \rightarrow \infty} x^k = x \quad (1.21)$$

Se os métodos não terminarem ao fim de um número finito de iterações, teoricamente têm que continuar indefinidamente. Esta é uma das características fundamentais que distingue estes processos dos métodos directos. A convergência destes métodos só se verifica no limite,

pelo que só se consegue obter uma solução aproximada. Esta aproximação pode, contudo, ser suficientemente boa para a finalidade pretendida, pelo que estes métodos não devem ser considerados menos eficazes do que os métodos directos. Sugerimos [26, 35] para discussões aprofundadas dos vários métodos iterativos que têm vindo a ser propostos na literatura para a resolução de problemas complementares. Neste trabalho vamos apenas mencionar uma classe de métodos iterativos que se tem mostrado muito interessante e eficiente para a resolução do LCP, os chamados algoritmos de pontos interiores. Este tipo de processos teve origem num trabalho desenvolvido por Karmakar em 1984 [72] para resolver programas lineares. Como o próprio nome indica, estes métodos operam no interior do octante não negativo $\mathbb{R}_+^{2n}$, isto é, $w^k > 0$ e $z^k > 0$ para cada iterando $x^k = (w^k, z^k)$, num esforço de reduzir $z^T w$ a zero e obter admissibilidade $w = q + Mz$. Para este fim, é usada uma direcção d^k em cada iteração que consiste numa modificação ao método de Newton [76]. Um passo α_k é calculado de modo a manter os iterandos no interior do octante não negativo de $\mathbb{R}^{2n}$. Para uma descrição desse processo, comecemos por reescrever o LCP como um sistema de equações não lineares

$$w - Mz - q = 0 \quad (1.22)$$

$$ZWe = 0 \quad (1.23)$$

$$w > 0, \quad z > 0 \quad (1.24)$$

onde Z e W são matrizes diagonais cujos elementos diagonais são z_i e w_i respectivamente e e é um vector de uns de dimensão n .

Uma forma de resolver este problema consiste em aplicar directamente o método de Newton à resolução do sistema não linear constituído pelas restrições (1.22) e (1.23). Contudo, esta aplicação pode introduzir problemas de fronteira, isto é, numa determinada iteração o algoritmo pode não ser capaz de se mover por ter encontrado um ponto fronteiro ($z_i w_i = 0$). Deste modo, os métodos de pontos interiores mais conhecidos e eficientes seguem o denominado caminho central [121] em $\mathbb{R}_{++}^n = \{(w, z) : w > 0, z > 0\}$ numa tentativa de conseguir obter uma solução complementar e admissível. A determinação de uma direcção de pesquisa é então associada ao seguinte sistema

$$\begin{aligned} w - Mz - q &= 0 \\ ZWe &= \mu_k e \end{aligned} \quad (1.25)$$

onde μ_k é o denominado parâmetro central e satisfaz

$$\mu_k = \delta \frac{(z^k)^T w^k}{n}$$

com $0 < \delta < 1$ e (w^k, z^k) é o ponto corrente.

O Jacobiano associado ao sistema (1.25) no ponto (w^k, z^k) é dado por

$$\begin{pmatrix} -I & M \\ Z^k & W^k \end{pmatrix}.$$

Desta forma a direcção de Newton é calculada pela resolução do sistema

$$\begin{pmatrix} -I & M \\ Z^k & W^k \end{pmatrix} \begin{pmatrix} \Delta w^k \\ \Delta z^k \end{pmatrix} = \begin{pmatrix} w^k - Mz^k - q \\ -Z^k W^k e + \mu_k e \end{pmatrix},$$

O próximo iterando (w^{k+1}, z^{k+1}) é então determinado por

$$(w^{k+1}, z^{k+1}) = (w^k, z^k) + \alpha_k(\Delta w^k, \Delta z^k),$$

onde o tamanho do passo α_k é escolhido de forma a garantir que os iterandos permaneçam no interior do octante não negativo e haja algum progresso em relação à satisfação da complementaridade e admissibilidade. O algoritmo termina quando essas condições (1.22) e (1.23) são aproximadamente satisfeitas, isto é, se (w^k, z^k) verifica

$$(z^k)^T w^k < \epsilon_1 \quad e \quad \| -w^k + Mz^k + q \| < \epsilon_2$$

para uma certa norma e para tolerâncias positivas ϵ_1 e ϵ_2 . Sugerimos [36, 76, 120, 121] para uma descrição detalhada deste tipo de métodos.

Capítulo 2

ALGORITMOS DE PROGRAMAÇÃO QUADRÁTICA E ENUMERATIVOS PARA PROBLEMAS LINEARES COMPLEMENTARES

Os métodos enumerativos são os únicos algoritmos que resolvem problemas lineares complementares sem imporem restrições aos dados. Neste tipo de processos é explorada a redução do LCP ao programa quadrático

$$\begin{array}{ll} \text{Minimizar} & g_1(z, w) = z^T w \\ \text{sujeito a} & w - Mz = q \\ & z \geq 0, \quad w \geq 0 \end{array} \quad (2.1)$$

referida no capítulo anterior. Neste capítulo é discutido um método enumerativo para o LCP, que incorpora um método de Gradiente Reduzido Modificado ou um método de Restrições Activas para tratar o programa quadrático (2.1). Esses processos serão discutidos em detalhe neste capítulo conjuntamente com a descrição do método enumerativo. Na última secção, são apresentados alguns resultados computacionais com o método enumerativo na resolução de LCPs NP-difíceis.

2.1 Algoritmo de Gradiente Reduzido Modificado

O Algoritmo de Gradiente Reduzido Modificado (MRG) [3] procura encontrar um *Mínimo Local Estrela (MLE)*, do inglês *Local Star Minimum*, da função g_1 no conjunto admissível K (1.12) do LCP. Um *MLE* de g_1 em K é um ponto extremo $\bar{x} = (\bar{z}, \bar{w})$ de K que satisfaz

$$g_1(\bar{z}, \bar{w}) \leq g_1(z, w)$$

para todos os pontos extremos adjacentes $x = (z, w) \in K$ a $(\bar{z}, \bar{w})$.

Para descrever o algoritmo MRG, seja $\bar{x} = (\bar{z}, \bar{w})$ um ponto extremo de K associado a uma solução básica admissível com base B . Se (z^*, w^*) é um ponto extremo adjacente, então

$$(z^*, w^*) = (\bar{z}, \bar{w}) + \mu(d_z, d_w)$$

onde μ é o tamanho do passo máximo usado no método simplex e $d = (d_z, d_w)$ é uma direcção admissível na qual d_z e d_w são vectores que contêm todas as componentes d_i de d associadas às variáveis z_i e w_i respectivamente. Esta direcção admissível pode ser definida em termos da base B e das colunas da matriz M ou da matriz identidade I . De facto, sejam P e T os conjuntos dos índices das variáveis x básicas e não básicas respectivamente e seja x_s a variável não básica escolhida para entrar na base. Incrementando esta variável de zero, gera-se uma solução básica admissível adjacente associada a $x^* = (z^*, w^*)$. Então a direcção d é definida por

$$\begin{aligned} d_s &= 1 \\ d_i &= 0 \text{ para todo } i \in T \setminus \{s\} \\ d_P &= \begin{cases} -B^{-1}M_{\cdot r} & \text{se } x_s = z_r \\ B^{-1}e^r & \text{se } x_s = w_r \end{cases} \end{aligned} \quad (2.2)$$

onde $M_{\cdot r}$ e e^r são as r -ésimas colunas das matrizes M e I respectivamente. O valor da função quadrática $g_1(z, w) = z^T w$ no novo ponto extremo (z^*, w^*) é então

$$g_1(z^*, w^*) = \bar{z}^T \bar{w} + \mu(\bar{z}^T d_w + \bar{w}^T d_z) + \mu^2 d_z^T d_w$$

Consequentemente existe uma direcção descendente para um novo ponto extremo adjacente ($\mu > 0$) se e só se

$$\bar{z}^T d_w + \bar{w}^T d_z + \mu d_z^T d_w < 0 \quad (2.3)$$

Em cada iteração do algoritmo MRG é procurada uma direcção descendente admissível d que, para um passo máximo μ (calculado como no método simplex), satisfaz (2.3). Se esse vector d existe, o algoritmo move-se para um ponto extremo adjacente com um decréscimo da função objectivo g_1 . De outro modo, o algoritmo termina. Casos degenerados onde $\mu = 0$ podem ser tratados através da chamada regra de Bland [63] ou por um outro esquema designado para ultrapassar essa dificuldade [91]. Seja

$$\varphi(z, w) = \bar{z}^T \bar{w} + \bar{z}^T w + \bar{w}^T z$$

uma aproximação linear de $g_1(z, w)$ no ponto extremo $(\bar{z}, \bar{w})$. Então é fácil mostrar [63] que $\bar{z}^T d_w + \bar{w}^T d_z$ é o coeficiente de custo reduzido $\bar{c}_s$ de $\varphi(z, w)$ associado à variável não básica de índice s a entrar na base e que é incrementada para gerar o novo ponto extremo adjacente. Deste modo, se $d_z^T d_w \leq 0$ e $\bar{c}_s < 0$ então

$$g_1(z^*, w^*) \leq g(\bar{z}, \bar{w}) + \bar{c}_s \mu < g(\bar{z}, \bar{w})$$

com $\mu > 0$. Assim, tal como no método simplex para programação linear, um coeficiente de custo reduzido negativo significa uma direcção descendente admissível. Por isso se existir a garantia de que $d_z^T d_w \leq 0$ se verifica sempre, então o método MRG reduz-se ao método simplex usual. Essa propriedade é verificada em alguns problemas lineares complementares associados à resolução de problemas de optimização global [63].

2.2 Algoritmo de Restrições Activas

O algoritmo apresentado nesta secção procura obter um mínimo local $(\bar{z}, \bar{w})$ para o programa quadrático (2.1) ou mostrar que este é inadmissível. De um modo geral, os algoritmos de restrições activas [95] começam por particionar para cada solução admissível as restrições de desigualdade em dois conjuntos de restrições:

Activas: Restrições de desigualdade que se verificam como igualdade (as restrições de igualdade são sempre activas).

Inactivas: Restrições de desigualdade que se verificam estritamente.

Em qualquer iteração k , o conjunto inactivo é ignorado, enquanto o conjunto activo passa a ser o *conjunto de trabalho*. Um novo iterando é então gerado por um movimento na superfície definida pelo *conjunto de trabalho*.

Para uma descrição geral do algoritmo de restrições activas (ASET), consideremos o programa não linear na sua forma simplificada

$$\begin{array}{ll} \text{Minimizar} & f(x) \\ \text{sujeito a} & Ax \geq b \end{array}$$

onde $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ e $f \in C^2(D)$, com D um conjunto aberto e convexo que contém o conjunto admissível

$$K = \{x \in \mathbb{R}^n : Ax \geq b\}$$

Um elemento $\bar{x}$ de K é um ponto estacionário de f em K se e só se as seguintes condições são verificadas

$$\begin{aligned} \nabla f(x) = \bar{A}^T \mu &\Leftrightarrow Z^T \nabla f(x) = 0 \\ \bar{A}x &= \bar{b} \\ Ax &\geq b \\ \mu &\geq 0 \end{aligned} \tag{2.4}$$

onde $\bar{A} \in \mathbb{R}^{t \times n}$ contém as linhas $i \in I(\bar{x})$ de A correspondentes às restrições activas em $\bar{x}$, t é o número de elementos de $I(\bar{x})$ ($t = |I(\bar{x})|$), $\mu \in \mathbb{R}^t$ é o vector dos multiplicadores de Lagrange associado a essas restrições activas e $Z \in \mathbb{R}^{n \times (n-t)}$ é a matriz constituída pelos vectores de uma base do subespaço nulo de $\bar{A}$. O vector $Z^T \nabla f(\bar{x}) \in \mathbb{R}^{(n-t)}$ é denominado gradiente reduzido de f em $\bar{x}$.

O método de restrições activas é um algoritmo descendente (ou de gradientes) com as seguintes características:

- (i) Apenas soluções admissíveis $x^k \in K$ são usadas.
- (ii) Em cada iteração k é escolhida uma direcção descendente d^k para um problema com igualdades lineares constituídas pelas restrições activas em x^k .
- (iii) Se x^k é um ponto estacionário de f no conjunto definido pelas suas restrições activas, então satisfaz as três primeiras condições de (2.4) e dois casos podem acontecer:

- a) Se $\mu \geq 0$, então x^k é ponto estacionário de f em K e o algoritmo termina.
- b) Se tal não acontece, então uma restrição activa deve ser retirada do conjunto de restrições activas em x^k .

Esta decisão é baseada no seguinte resultado:

Teorema 2.1 *Seja $x^k \in K$ que satisfaz $A_i x^k = b_i$, $i \in I(x^k)$. Sejam ainda $|I(x^k)| = t$, $\bar{A} = [A_i]_{i \in I(x^k)}$ e $\mu \in \mathbb{R}^t$ o vector solução de*

$$\bar{A}^T \mu = \nabla f(x^k)$$

Se $\mu_r < 0$ então qualquer direcção $d \in \mathbb{R}^n$ que satisfaça

$$A_r d > 0, \quad A_i d = 0, \quad i \in I(x^k) \setminus \{r\}$$

é descendente em x^k .

Demonstração:

Com efeito, tem-se

$$\begin{aligned} \nabla f(x^k) d &= (\bar{A}^T \mu)^T d = \mu^T (\bar{A} d) = \sum_{i \in I(x^k)} \mu_i (A_i d) \\ &= \sum_{i \in I(x^k) \setminus \{r\}} \mu_i (A_i d) + \mu_r (A_r d) < 0 \end{aligned}$$

pois $\mu_r < 0$, $A_r d > 0$ e $A_i d = 0$ para todo o $i \in I(x^k) \setminus \{r\}$. ◇

De acordo com este resultado, se supirmos a restrição activa $A_r x \geq b_r$ correspondente a um multiplicador $\mu_r < 0$, fazendo $I(x^k) = I(x^k) \setminus \{r\}$, então qualquer direcção d que satisfaça $A_r d > 0$ e seja admissível no espaço das restantes restrições é descendente em x^k .

Devido às observações apresentadas, dado um vector $x^k \in K$ que não seja ponto estacionário de f em K , é sempre possível encontrar uma direcção descendente d^k em x^k e, por isso, um movimento nessa direcção possibilita uma redução no valor da função. Tal como é comum nos métodos de gradientes, é necessário determinar um passo $\alpha_k > 0$ tal que $x^k + \alpha_k d^k$ é o novo iterando. Para o cálculo de α_k , seja d^k a direcção obtida na iteração k . Então para qualquer $i \in I(x^k)$, $A_i d^k = 0$ e

$$A_i. \left(x^k + \alpha d^k \right) = A_i.x^k = b_i, \quad i \in I(x^k)$$

Portanto todas as restrições activas em x^k mantêm-se activas ao longo da semi-recta que emana de x^k segundo a direcção d^k . Como se pretende que todos os iterandos permaneçam em K , então o passo $\alpha \geq 0$ deve satisfazer

$$A_i. \left(x^k + \alpha d^k \right) \geq b_i$$

para todas as restrições inactivas $i \notin I(x^k)$. Portanto

$$A_i.x^k + \alpha \left(A_i.d^k \right) \geq b_i$$

Se $A_i.d^k \geq 0$ para todo $i \notin I(x^k)$, então α pode tomar um valor qualquer não negativo. Por outro lado, se $A_i.d^k < 0$ para certo $i \notin I(x^k)$, então α tem de satisfazer

$$\alpha \left(-A_i.d^k \right) \leq A_i.x^k - b_i$$

Portanto α_k deve ser um número não negativo inferior ou igual a

$$\alpha_{max} = \begin{cases} \min \left\{ \frac{A_i.x^k - b_i}{-A_i.d^k} : i \notin I(x^k) \text{ e } A_i.d^k < 0 \right\} \\ +\infty \text{ se } A_i.d^k \geq 0 \text{ para todo } i \notin I(x^k) \end{cases} \quad (2.5)$$

É ainda de notar que se $\alpha_k = \alpha_{max}$, então há um número $l \geq 1$ de restrições inactivas que passam a activas. Essas restrições correspondem aos índices i onde α_{max} foi atingido. A convergência global do processo é assegurada desde que o passo α_k seja calculado por uma técnica de pesquisa unidireccional baseada no Critério de Armijo, Minimização ou outro semelhante [95] que procure minimizar $f(x^k + \alpha d^k)$ para $0 \leq \alpha \leq \alpha_{max}$ de uma forma exacta ou inexacta.

Para facilitar a descrição do algoritmo de restrições activas, analisemos o número possível de restrições activas associadas a cada ponto $x^k \in K$. Se K é um conjunto definido por desigualdades lineares, então K contém pontos interiores e pontos fronteiros, entre os quais se destacam os pontos extremos. É sabido que os pontos extremos não se podem escrever como combinação convexa de pontos de K . Além disso apenas os pontos extremos de K estão nessas condições. Por outro lado, não há restrições activas para pontos interiores, os

pontos fronteiros têm associada pelo menos uma restrição activa, enquanto que para pontos extremos o número dessas restrições é maior ou igual a n . Para que esse número seja superior a n é necessário que existam restrições redundantes. Se a Hipótese de Não Degenerescência for verificada:

$$\bar{A} = [A_i.]_{i \in I(\bar{x})} \Rightarrow \text{característica de } \bar{A} = |I(\bar{x})| \leq n \quad (2.6)$$

então dois casos podem ocorrer para cada $\bar{x} \in K$:

1. Se a característica de $\bar{A} = [A_i.]_{i \in I(\bar{x})}$ é igual a $t = |I(\bar{x})| < n$, então $\bar{x}$ não é ponto extremo de K e existe uma matriz $Z \in \mathbb{R}^{n \times (n-t)}$ constituída pelos vectores de uma base do subespaço nulo de $\bar{A}$.
2. Se a característica de $\bar{A}$ é igual a n , então $\bar{x}$ é ponto extremo de K e a matriz Z não existe.

Notar que de acordo com essa hipótese, se para uma dada solução admissível $\bar{x}$, o número $|I(\bar{x})|$ de restrições de desigualdades activas em $\bar{x}$ é superior a n , então apenas n dessas restrições são consideradas como activas. Além disso para que a hipótese seja verdadeira, as $A_i x \geq b_i$, $i \in I(\bar{x})$ devem ser escolhidas de forma a que a matriz $\bar{A}$ em (2.6) seja não singular.

Após estas considerações passemos à descrição do algoritmo:

Algoritmo de Restrições Activas

Passo 0 Sejam $x^0 \in K$ e $I(x^0)$ o conjunto dos índices das linhas de A correspondentes às restrições activas em x^0 . Se $|I(x^0)| < n$ determine a matriz Z_0 cujas colunas são constituídas pelos vectores de uma base do subespaço nulo de $\bar{A} = [A_i.]_{i \in I(x^0)}$. Faça $k = 0$.

Passo 1 (i) Se $|I(x^k)| = n$ ou $Z_k^T \nabla f(x^k) = 0$, calcule μ a partir de

$$\bar{A}^T \mu = \nabla f(x^k)$$

Se $\mu \geq 0$, x^k é ponto estacionário de f em K e termine. De outro modo seja

$$\mu_r = \min \left\{ \mu_i : i \in I(x^k), \mu_i < 0 \right\}$$

Faça $I(x^k) = I(x^k) \setminus \{r\}$ e actualize Z_k com mais uma coluna. Vá para

Passo 2.

(ii) Se $Z_k^T \nabla f(x^k) \neq 0$ vá para o Passo 2.

Passo 2 Calcule uma direcção d^k descendente e admissível para o problema

$$\begin{aligned} & \text{Minimizar} && f(x) \\ & \text{sujeito a} && A_i x = b_i, \quad i \in I(x^k) \end{aligned}$$

isto é, um vector d^k satisfazendo

$$A_i d^k = 0, \quad i \in I(x^k), \quad \nabla f(x^k)^T d^k < 0$$

e $A_r d^k > 0$ se no Passo 1 houve um decréscimo em $I(x^k)$.

Passo 3 Calcule α_{max} a partir de (2.5) e determine α_k usando um critério que procure *minimizar* $f(x^k + \alpha d^k)$ para $0 \leq \alpha \leq \alpha_{max}$.

Passo 4 Faça

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

Se $\alpha_k < \alpha_{max}$, faça $Z_{k+1} = Z_k$. De outro modo ($\alpha_k = \alpha_{max}$) acrescente a $I(x^k)$ as linhas onde α_{max} é atingido. Se $|I(x^k)| > n$, escolha $I(x^k)$ de forma a que $|I(x^k)| = n$ e $\bar{A} = [A_i]_{i \in I(x^k)}$ seja não singular. Se $|I(x^k)| < n$, actualize Z_{k+1} .

Vá para o Passo 1 com $k = k + 1$.

A convergência global do algoritmo está assegurada desde que se verifique a hipótese de não degenerescência (2.6) e $\alpha_{max} > 0$ em cada iteração. A hipótese de não degenerescência é normalmente verificada na prática. Por outro lado, em geral o algoritmo é capaz de determinar um ponto estacionário de f em K , mesmo quando α_{max} seja nulo em algumas iterações [95]. Notar que se o programa não linear contiver igualdades, então essas restrições são sempre activas. Além disso os correspondentes multiplicadores não têm restrição de sinal pelo que o seu sinal nunca será objecto de estudo no Passo 1.

A direcção d^k pode ser obtida por uma técnica de descida máxima, Newton modificado ou de secante [95]. O processo da descida máxima tem uma implementação simples, mas é em geral bastante lento. Em contrapartida, uma das grandes vantagens do processo de Newton é a rapidez da sua convergência. Este processo garante que em cada iteração a direcção obtida d^k seja sempre descendente, uma vez que neste processo é adicionada à Hessiana reduzida $Z_k^T \nabla^2 f(x^k) Z_k$ uma matriz diagonal de elementos não negativos de forma a que a matriz resultante seja simétrica PD . Contudo, este processo tem a desvantagem de necessitar do cálculo da Hessiana $\nabla^2 f(x^k)$ da função. Em compensação, no método de secante *BFGS* esse cálculo não é necessário. Apesar da sua taxa de convergência ser em geral inferior à do método de Newton, a vantagem anterior tem levado à sua utilização nas implementações dos algoritmos de restrições activas. Uma das implementações mais conhecidas para este tipo de processo está inserida no código comercial MINOS [50], que será brevemente descrito na secção seguinte.

2.3 Código MINOS

O MINOS [50] é um sistema desenvolvido em Fortran para resolver problemas de optimização de grande dimensão expressos na seguinte forma

$$\begin{array}{llll}
 \text{Minimizar} & F(t) & +c^T t & +d^T y \\
 \text{sujeito a} & f(t) & & +A_1 y = b^1 \\
 & & A_2 t & +A_3 y = b^2 \\
 & l & \leq & t, y \leq u
 \end{array}$$

onde os vectores c , d , b^1 , b^2 , l , u e as matrizes A_1 , A_2 , A_3 são dados, $F(t)$ é uma função escalar diferenciável, e $f(t)$ é um vector de funções diferenciáveis $\{f_i(t)\}$.

Este algoritmo implementa uma estratégia indirecta de restrições activas convertendo todas as restrições de desigualdade em igualdades através da adição de variáveis de folga que estão sujeitas a limites. Esta aproximação faz com que todas as restrições, quer lineares quer não lineares, façam sempre parte do conjunto de trabalho. Deste modo, o problema anterior é escrito da seguinte forma

$$\begin{aligned}
\text{Minimizar} \quad & F(t) + c^T t + d^T y = \bar{F}(t, y) \\
\text{sujeito a} \quad & f(t) + A_1 y + I s = 0 \\
& A_2 t + A_3 y + I v = 0 \\
& l \leq t, y, s, v \leq u
\end{aligned} \tag{2.7}$$

onde s e v são as variáveis de folga. Por razões computacionais, os termos do segundo membro, b^1 e b^2 , são incorporados nos limites das variáveis s e v .

Este algoritmo é usado para resolver problemas de programação linear e não linear. Para resolver programas lineares o MINOS usa uma implementação do método primal simplex [27]. Para programas não lineares existem duas técnicas aliadas ao tipo de não linearidade existente no problema em questão que são discutidas a seguir.

- Se a não linearidade estiver confinada à função objectivo, o MINOS resolve estes problemas usando um algoritmo de restrições activas que tal como o anteriormente descrito utiliza o gradiente reduzido [119] e um algoritmo de secante BFGS para a determinação da direcção [29].
- Se a não linearidade estiver também confinada às restrições, o MINOS usa o algoritmo da projecção lagrangeana aumentado, baseado num método devido a Robinson [103].

No caso de $f(t) = 0$, se $x = (t^T, y^T, s^T, v^T)^T$ então o problema (2.7) pode ser reescrito da seguinte forma:

$$\begin{aligned}
\text{Minimizar} \quad & \bar{F}(x) \\
\text{sujeito a} \quad & Ax = 0 \\
& l \leq x \leq u
\end{aligned} \tag{2.8}$$

Na implementação do método, é assumido que a matriz A tem característica completa. Isso é suficiente para garantir a hipótese de não degenerescência (2.6). As restrições activas $Ax = 0$ são escritas na forma

$$Bx_B + Sx_S + Nx_N = 0$$

onde B é uma matriz *Base* quadrada não singular e x_B , x_S e x_N são respectivamente os vectores das variáveis básicas, superbásicas e não básicas. Em qualquer solução, as variáveis básicas e superbásicas tomam valores entre os seus limites enquanto que as não básicas são

fixas em um dos seus limites.

As variáveis superbásicas são vistas pelo método de restrições activas como um conjunto de variáveis livres, uma vez que se podem mover em qualquer direcção, nomeadamente numa que melhore o valor da função objectivo.

Em cada iteração, sempre que um multiplicador associado a uma restrição activa $x_r = l_r$ ou $x_r = u_r$ não tem o sinal adequado, a correspondente variável não básica x_r é seleccionada a passar a superbásica ou básica. Em contrapartida, sempre que uma variável básica ou superbásica atinja um dos seus limites, passará a não básica desde que a hipótese de não degenerescência se mantenha verdadeira. Assim mudanças no conjunto das restrições activas correspondem a mudanças no estatuto das variáveis básicas, superbásicas e não básicas. Notar ainda que este algoritmo se reduz ao método simplex quando a função $\bar{F}$ é linear.

Um dos passos mais importantes do MINOS é a utilização do método de secante para a determinação da direcção de pesquisa $d = Zd_S$, onde d_S é um vector de dimensão igual ao número de variáveis superbásicas e que é obtido resolvendo um sistema da forma

$$R^T R d_S = -Z^T \nabla \bar{F}(\bar{x}).$$

onde $Z^T \nabla \bar{F}(\bar{x})$ é o gradiente reduzido de $\bar{F}$ no ponto corrente $\bar{x}$ e R é uma matriz triangular superior, que é actualizada de modo a aproximar-se da Hessiana reduzida, isto é, $R^T R \approx Z^T \nabla^2 \bar{F}(\bar{x}) Z$. A matriz Z é constituída pelos vectores de uma base do espaço nulo da matriz $\bar{A}$ associada às restrições activas em $\bar{x}$ e tem a forma

$$Z = \begin{pmatrix} B^{-1}S \\ I \\ 0 \end{pmatrix}$$

onde I é a matriz identidade de ordem igual ao número de variáveis superbásicas. A determinação do passo α_k é feita utilizando uma técnica de pesquisa unidimensional

$$\min_{\alpha} \bar{F}(\bar{x} + \alpha d) \quad \text{sujeito a } 0 \leq \alpha \leq \alpha_{max},$$

onde α_{max} é determinado por um critério da admissibilidade em relação aos limites inferiores e superiores não activos.

No caso de existirem restrições não lineares, isto é, quando $f(t) \neq 0$, o MINOS é uma

implementação de um algoritmo de projecção lagrangeano aumentado. Este processo é constituído por uma sequência de iterações maiores, cada uma das quais requer a solução de um subproblema com restrições lineares. Cada um destes subproblemas contém linearizações das restrições não lineares, assim como as restrições lineares originais e os limites das variáveis. Para mais detalhes sobre o código MINOS sugerimos [50, 88].

2.4 Método Enumerativo

No capítulo 1 mostrámos que existem vários algoritmos directos e iterativos para a resolução do LCP. Contudo, estes processos só são capazes de encontrar uma solução do LCP ou mostrar que não existe tal solução quando a matriz M pertence a determinadas classes de matrizes. No caso geral, Chung [22] mostrou que o $\text{LCP}(q, M)$ é NP-difícil. Consequentemente, os algoritmos enumerativos são os únicos que conseguem resolver o $\text{LCP}(q, M)$ sem impor restrições à classe da sua matriz.

Se o LCP tem solução, então cada par de variáveis (z_i, w_i) tem de satisfazer a condição de complementaridade $z_i w_i = 0$. Portanto em qualquer solução do LCP, $z_i = 0$ ou $w_i = 0$ para cada $i = 1, \dots, n$. Essa propriedade dá ao LCP um carácter combinatório, que serve de motivação à utilização de um método enumerativo. Assim um algoritmo desse tipo encontra uma solução para o LCP explorando uma árvore da forma:

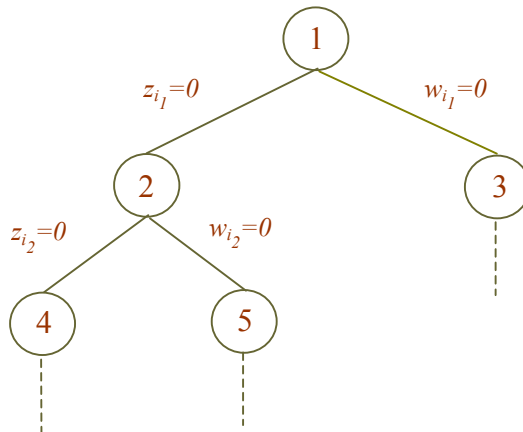


Figura 2.1: Esquema de ramificação do método enumerativo

Em cada nó da árvore é considerado um programa quadrático, obtido do QP apresentado nas secções anteriores por adição de algumas restrições do tipo $z_i = 0$ ou $w_i = 0$. Esse programa

tem então a forma

$$\begin{aligned}
 & \text{Minimizar} && g_1(z, w) = z^T w \\
 & \text{sujeito a} && w - Mz = q \\
 & && z_L = 0 \\
 & && w_J = 0 \\
 & && z \geq 0, \quad w \geq 0
 \end{aligned} \tag{2.9}$$

onde

- $J = \{i : w_i \text{ está fixa em zero}\}$
- $L = \{i : z_i \text{ está fixa em zero}\}$

satisfazem $J \cap L = \emptyset$ e $J \cup L \subseteq \{1, \dots, n\}$. Assim, por exemplo no nó 5 da árvore anterior, o programa quadrático a resolver consiste em

$$\begin{aligned}
 & \text{Minimizar} && g_1(z, w) = z^T w \\
 & \text{sujeito a} && w - Mz = q \\
 & && z_{i_1} = 0 \\
 & && w_{i_2} = 0 \\
 & && z \geq 0, \quad w \geq 0
 \end{aligned} \tag{2.10}$$

Um algoritmo de optimização quadrática global [44, 61] pode tornar-se bastante eficiente na resolução dos LCPs, uma vez que $g_1(z, w) = 0$ fornece um critério de paragem para o processo. Esta característica importante é essencial para um bom desempenho do método enumerativo na resolução de LCPs NP-difíceis. Além disso, os algoritmos MRG e ASET descritos nas secções anteriores parecem ser ferramentas fundamentais para determinar um ponto estacionário ou um *MLE* para o programa quadrático correspondente a cada nó. Os passos do algoritmo enumerativo são apresentados a seguir.

Algoritmo Enumerativo

Passo 0 Seja $\mathcal{L} = \{1\}$ a lista inicial dos nós em aberto e QP(1) o programa quadrático QP (2.9) com $J_1 = L_1 = \emptyset$. Faça $k = 1$.

Passo 1 Se $\mathcal{L} = \emptyset$, páre: o LCP não tem solução. De outro modo, escolha um nó $t \in \mathcal{L}$.

Passo 2 Retire t de $\mathcal{L}$. Aplique um dos algoritmos MRG ou ASET ao programa quadrático $QP(t)$ associado ao nó t . Se $QP(t)$ é inadmissível, volte para Passo 1. De outro modo, seja $(\bar{z}, \bar{w})$ a solução obtida para $QP(t)$.

Passo 3 Se $g_1(\bar{z}, \bar{w}) = \bar{z}^T \bar{w} = 0$, páre: $(\bar{z}, \bar{w})$ é solução do LCP. De outro modo, seja (z_r, w_r) um par de variáveis básicas positivas em $(\bar{z}, \bar{w})$.

Passo 4 Adicione dois novos nós $k+1$ e $k+2$ à lista $\mathcal{L}$, com os programas quadráticos $QP(k+1)$ e $QP(k+2)$ definidos por:

$$QP(k+1) \quad : \quad L_{k+1} = L_t \cup \{r\}, \quad J_{k+1} = J_t$$

$$QP(k+2) \quad : \quad J_{k+2} = J_t \cup \{r\}, \quad L_{k+2} = L_t$$

Volte ao Passo 1.

Um algoritmo enumerativo só é eficiente se pesquisar poucos nós antes de se obter uma solução para o LCP ou mostrar que esta não existe. Para isso, Al-Khayyal [3] sugere que se utilizem as seguintes heurísticas:

- **Escolha do nó (Passo 1):** aquele para o qual o produto $z^T w$ é mínimo.
- **Escolha do par de variáveis complementares (Passo 3):** aquele para o qual o produto $z_r w_r$ é máximo.

Em [62] é sugerido

- **Escolha do nó (Passo 1):** aquele para o qual é mínima a expressão

$$\text{NCP} + \frac{z^T w}{\lambda},$$

onde NCP é o número de variáveis complementares básicas e λ é um factor normalmente igual a 10.

A experiência computacional até agora efectuada demonstra que estas heurísticas aceleram a pesquisa da solução do LCP [62].

Suponhamos que o algoritmo do Gradiente Reduzido Modificado (MRG) é aplicado em cada nó do método enumerativo para encontrar um *MLE* da função g_1 num conjunto $\bar{K}$ que contém as restrições lineares associadas a esse nó. Da descrição do algoritmo MRG impõe-se, em cada

nó, a determinação de uma solução admissível para o programa quadrático correspondente. Para isso, é necessário resolver um programa linear da forma:

$$\begin{aligned}
 & \text{Minimizar} && z_0 \\
 & \text{sujeito a} && w = q + z_0 p + Mz \\
 & && z \geq 0, \quad w \geq 0, \quad z_0 \geq 0
 \end{aligned} \tag{2.11}$$

onde z_0 é uma variável artificial e p um vector não negativo tal que $p_i > 0$ para todo o i tal que $q_i < 0$. O método simplex pode ser usado para esse fim. Dois casos podem acontecer:

1. $\min z_0 > 0$ e o conjunto admissível do LCP é vazio.
2. $\min z_0 = 0$ e a solução obtida é uma solução admissível para o LCP.

Como o nosso objectivo é determinar uma solução admissível e complementar, então na implementação utilizada do método simplex foi efectuada uma modificação de modo a escolher, sempre que possível, para variável a entrar na base, uma variável que, além de melhorar o valor da função objectivo, não piore o número de pares de variáveis complementares básicas [62]. Esta modificação na implementação do método simplex também é utilizada em cada um dos restantes nós da árvore binária, para determinar uma solução admissível para o QP(t) correspondente, resolvendo desta forma o seguinte programa linear:

$$\begin{aligned}
 & \text{Minimizar} && x_i \\
 & \text{sujeito a} && w = q + Mz \\
 & && z \geq 0, \quad w \geq 0 \\
 & && z_{L_t} = 0 \\
 & && w_{J_t} = 0
 \end{aligned} \tag{2.12}$$

com x_i a representar uma variável complementar z_i ou w_i .

Da descrição do algoritmo pode concluir-se que apenas soluções básicas admissíveis do LCP são utilizadas ao longo do algoritmo. Por este motivo o método enumerativo baseado no Algoritmo do Gradiente Reduzido Modificado, que denotaremos por EMRG, pode ser implementado para LCPs de grandes dimensões usando as técnicas de reinversão e actualização da factorização LU que exploram a esparsidade das matrizes bases [91]. Sugere-se [62] para uma descrição mais detalhada desta implementação.

A segunda versão do método enumerativo utiliza um método de restrições activas para

determinar em cada nó um ponto estacionário do programa quadrático correspondente. O método enumerativo resultante será denotado por EASET e a sua implementação é baseada no código MINOS descrito anteriormente. É importante notar que LCPs monótonos ou com matrizes $RSuf$ são resolvidos no nó 1, pois o ponto estacionário do QP é solução do LCP. Como é discutido em [62, 66], o método enumerativo é normalmente eficaz para determinar uma solução de um LCP quando ela existe. Essa característica é consequência da incorporação de um método local que determina um MLE ou um ponto estacionário do programa quadrático associado a cada nó e das heurísticas discutidas anteriormente. Se o LCP é admissível e não tem solução, então o algoritmo necessita de uma pesquisa exaustiva para terminar, pois esses processos locais não permitem terminar o processo. Essas conclusões irão ser realçadas ao longo deste trabalho.

2.5 Experiência Computacional

Nesta secção é apresentada alguma experiência computacional com as duas versões EMRG e EASET do método enumerativo na resolução de LCPs NP-difíceis. Esta experiência computacional foi realizada num Pentium IV 2.4GHz com 256MB de RAM. Os problemas teste utilizados nesta experiência estão associados a problemas de mochila. As reduções desses problemas a LCPs são discutidas a seguir.

Redução de um Problema de Mochila a um LCP

Dado um número real positivo b e um vector positivo $a \in \mathbb{R}^n$, o Problema da Mochila consiste em encontrar um vector $x \in \mathbb{R}^n$ tal que

$$\begin{aligned} a^T x &= b \\ x_i &\in \{0, 1\}, \quad i = 1, 2, \dots, n \end{aligned} \tag{2.13}$$

Nesta secção são apresentadas três formulações diferentes do Problema da Mochila como um LCP. A primeira formulação baseia-se no facto de que para cada $i = 1, 2, \dots, n$, $x_i \in \{0, 1\}$ é equivalente a

$$w_i = 1 - x_i, \quad x_i \geq 0, \quad w_i \geq 0, \quad x_i w_i = 0.$$

É então fácil de mostrar [22] que o Problema da Mochila é equivalente ao seguinte LCP de dimensão $(n + 2)$:

$$\begin{aligned} w &= e - x \\ w_{n+1} &= -b + a^T x - \alpha x_{n+1} \\ w_{n+2} &= b - a^T x - \beta x_{n+2} \\ x_i &\geq 0, \quad w_i \geq 0, \quad x_i w_i = 0, \quad i = 1, 2, \dots, n+2 \end{aligned}$$

onde $e \in \mathbb{R}^n$ é um vector de uns e α, β são dois números reais positivos. A matriz M deste LCP tem então a seguinte forma:

$$M = \begin{bmatrix} -I & 0 & 0 \\ a^T & -\alpha & 0 \\ -a^T & 0 & -\beta \end{bmatrix} \in \mathbb{R}^{(n+2) \times (n+2)}$$

onde I é a matriz identidade de ordem n . Esta matriz é negativa semi-definida (*NSD*) se α e β são escolhidos de forma a satisfazerem as seguintes inequações

$$\alpha > \theta \frac{a^T a}{4}, \quad \beta > \theta \alpha \frac{a^T a}{4\alpha - a^T a}$$

onde $\theta > 1$ é um número escolhido arbitrariamente. Por outro lado, M é uma matriz indefinida (*IND*) se

$$\alpha > \theta \frac{a^T a}{4}, \quad \beta < \frac{\alpha}{\theta} \frac{a^T a}{4\alpha - a^T a}$$

onde $\theta > 1$, como anteriormente.

Como é estabelecido em [93], é também possível explorar a redução anterior para reduzir o Problema da Mochila a um $\text{LCP}(q, M)$, onde $q = [a, -b, b]^T \in \mathbb{R}^{n+2}$ e

$$M = \begin{bmatrix} -I & e & -e \\ e^T & -2n & 0 \\ -e^T & 0 & -2n \end{bmatrix} \in \mathbb{R}^{(n+2) \times (n+2)}$$

com I a matriz identidade de ordem n e $e \in \mathbb{R}^n$ um vector de uns. Notar que a matriz M é neste caso simétrica *NSD*.

Finalmente em [76], é apresentada a equivalência do Problema da Mochila a um $\text{LCP}(q, M)$, onde

(i) $q = [p \ p \ \dots \ p \ -b \ b]^T \in \mathbb{R}^{4n+2}$, com $p = [0 \ 0 \ -1 \ 1]^T \in \mathbb{R}^4$.

(ii)

$$M = \begin{bmatrix} B & 0 & \cdots & 0 & 0 & 0 \\ 0 & B & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & B & 0 & 0 \\ & & & \bar{a}^T & & \\ & & & -\bar{a}^T & & \end{bmatrix} \in \mathbb{R}^{(4n+2) \times (4n+2)}$$

onde

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{4 \times 4}$$

e as componentes $\bar{a}_i$ de $\bar{a} \in \mathbb{R}^{4n+2}$ satisfazem

$$\bar{a}_i = \begin{cases} a_j, & \text{se } i = 4j - 3, \quad j = 1, 2, \dots, n \\ 0, & \text{caso contrário} \end{cases}$$

Como é discutido em [76], a matriz M do LCP pertence à classe P_0 .

Designemos por 2N e 3P as duas últimas formulações LCP do Problema da Mochila nas quais a matriz associada é simétrica NSD e P_0 respectivamente, e por 1N e 1I as duas versões possíveis da primeira formulação do Problema da Mochila, consoante a matriz M associada seja NSD ou IND respectivamente.

Estes problemas foram gerados para estudar o comportamento do método enumerativo face à complexidade computacional destes LCPs. Uma vez que o Problema da Mochila é NP-difícil, então o mesmo acontece para o LCP quando a sua matriz M é simétrica ou não simétrica NSD , IND ou P_0 .

Nos problemas teste usados todas as componentes a_i do vector $a \in \mathbb{R}^n$ são geradas aleatoriamente no intervalo $[1, 50]$. O número real b é definido da seguinte forma

$$b = \sum_{i \in I} a_i$$

onde I é o subconjunto $\{1, \dots, p\}$ de $\{1, \dots, n\}$ correspondente às variáveis x_i que são iguais a um numa solução do problema da Mochila. Na nossa experiência computacional foram considerados três valores diferentes para p , a saber $\frac{3n}{4}$, $\frac{n}{2}$ e $\frac{n}{4}$. Portanto o vector b é gerado de modo a que uma solução do problema de mochila possua uma percentagem de variáveis iguais a um de 75%, 50% e 25%, respectivamente. Na notação de cada problema teste, após as duas siglas que identificam a formulação LCP utilizada, é incorporada uma outra sigla que é uma letra A, B ou C, e que corresponde à percentagem de variáveis iguais a um ser 75%, 50% ou 25%, respectivamente. Além disso, são também usados os dígitos 1, 2, 3 ou 4 dependendo da dimensão da matriz do problema da mochila ser 22, 52, 102 ou 152. No entanto, na terceira formulação, os números 2 e 4 significam que a dimensão do problema é 50 e 150, respectivamente.

Consequentemente, a identificação de cada um dos problemas teste associados aos problemas da mochila é efectuada através de um conjunto de quatro números ou letras. Assim, por exemplo, 1NA4 representa o LCP correspondente à primeira reformulação do problema de mochila, cuja matriz é *NSD* de dimensão 152 e no qual o vector b foi gerado com 75% das variáveis iguais a 1.

Comportamento do Método Enumerativo

Nesta subsecção é apresentada a experiência computacional com o método enumerativo e as suas versões EMRG e EASET discutidas anteriormente. Além da utilização das duas primeiras heurísticas para as escolhas de nós e do par de variáveis a ramificar, respectivamente, a implementação MINOS de EASET poderá ainda incluir alguns melhoramentos relacionados com a base inicial utilizada em cada nó e com o valor dos parâmetros associados ao uso da tolerância "Linesearch" e ao "Partial Price" [89]. Breves descrições sobre estes dois parâmetros são apresentadas a seguir.

- Linesearch tolerance (l), $l \in [0, 1]$

Para problemas não lineares, este parâmetro controla a precisão com que o tamanho do passo α é encontrado no problema unidimensional

$$\min_{\alpha} \bar{F}(\bar{x} + \alpha d) \quad \text{sujeito a} \quad 0 \leq \alpha \leq \alpha_{max},$$

O valor por defeito é 0.1, o que exige uma pesquisa com uma precisão moderada. Contudo, se todos os gradientes são conhecidos deve-se utilizar uma melhor precisão.

- Partial Price (p).

Este parâmetro é recomendado para problemas grandes que têm significativamente mais variáveis do que restrições. Reduz o trabalho exigido para cada operação "pricing" quando uma variável não básica é seleccionada para se tornar superbásica ou básica. Por defeito esse valor é 1, o que faz com que todas as colunas da matriz sejam pesquisadas. Caso contrário a matriz é particionada em p segmentos todos iguais. Se a pesquisa anterior foi bem sucedida a próxima pesquisa começa com o segmento posterior ao utilizado na pesquisa anterior. Se nessa pesquisa não foi seleccionada nenhuma variável não básica, a pesquisa continua nos segmentos seguintes.

Na versão EASET1 estes parâmetros assumem os valores por defeito e a base inicial utilizada em cada nó é a base final do problema anterior. Na versão EASET2 os parâmetros utilizados são $l=0.95$ e $p=5$ e em cada nó é utilizada como base inicial a base final do nó pai com uma pequena alteração que consiste em retirar da base através de uma operação pivotal a variável que é fixa em zero nesse nó.

O comportamento destas duas versões do método enumerativo EASET assim como o da versão EMRG na resolução de LCPs NP-difíceis pode ser observado nas Tabelas 2.1 e 2.2.

	EMRG			EASET1			EASET2		
	N _I	T	N _D	N _I	T	N _D	N _I	T	N _D
PIOR	1996	0.22	156	1003	0.13	91	912	0.14	130
MEDIA	218	0.02	25	211	0.03	26	131	0.02	20
MELHOR	11	0.00	1	12	0.00	1	9	0.00	1

Tabela 2.1: Comparação entre as versões EMRG e EASET do método enumerativo

Nestas tabelas PROB designa o problema teste a resolver, enquanto que N_I e N_D são o número total de iterações (operações pivotais) e nós, respectivamente e T é o tempo total CPU em segundos requerido pelos algoritmos enumerativos para resolverem o LCP.

Os resultados apresentados nas Tabelas 2.1 e 2.2 mostram que as versões EMRG e EASET são capazes de resolver todos os LCPs NP-difíceis num tempo razoável, sendo a eficiência dos dois métodos comparáveis. Além disso EASET2 é mais robusto e mais eficiente do que as restantes

PROB	EMRG			EASET					
	NI	T	ND	EASET1			EASET2		
				NI	T	ND	NI	T	ND
1NA1	140	0.00	31	171	0.03	31	168	0.03	75
1NA2	98	0.00	14	620	0.06	90	71	0.02	10
1NA3	272	0.05	66	147	0.00	20	57	0.00	2
1NA4	570	0.11	156	228	0.08	31	171	0.06	25
1NB1	21	0.00	5	352	0.05	91	167	0.00	78
1NB2	150	0.02	35	60	0.02	9	113	0.02	31
1NB3	104	0.02	16	207	0.05	33	34	0.00	4
1NB4	174	0.03	32	200	0.06	32	48	0.00	2
1NC1	17	0.00	5	226	0.02	55	12	0.00	5
1NC2	237	0.02	55	47	0.00	8	57	0.03	17
1NC3	38	0.00	4	195	0.05	30	108	0.02	24
1NC4	113	0.02	23	131	0.03	22	62	0.02	12
1IA1	140	0.00	31	171	0.03	31	168	0.02	75
1IA2	98	0.00	14	620	0.09	90	71	0.00	10
1IA3	272	0.02	66	147	0.05	20	57	0.02	2
1IA4	570	0.20	156	228	0.06	31	171	0.06	25
1IB1	21	0.00	5	352	0.03	91	167	0.06	78
1IB2	150	0.03	35	60	0.02	9	113	0.02	31
1IB3	104	0.02	16	207	0.03	33	34	0.00	4
1IB4	174	0.02	32	200	0.06	32	48	0.02	2
1IC1	17	0.00	5	226	0.03	55	12	0.02	5
1IC2	237	0.02	55	47	0.00	8	57	0.02	17
1IC3	38	0.00	4	195	0.05	30	108	0.05	24
1IC4	113	0.00	23	131	0.02	22	62	0.02	12
2NA1	23	0.00	2	24	0.00	2	75	0.00	12
2NA2	58	0.00	3	174	0.03	15	60	0.02	6
2NA3	81	0.02	1	82	0.02	1	142	0.05	13
2NA4	117	0.00	2	115	0.02	2	113	0.03	1
2NB1	343	0.00	46	46	0.00	4	11	0.00	2
2NB2	74	0.02	7	69	0.02	6	44	0.00	6
2NB3	51	0.00	1	48	0.02	1	55	0.00	1
2NB4	83	0.02	1	80	0.00	1	115	0.03	10
2NC1	34	0.00	5	29	0.02	3	70	0.00	16
2NC2	39	0.00	5	202	0.06	20	20	0.00	2
2NC3	37	0.00	2	46	0.00	3	47	0.00	4
2NC4	56	0.02	2	121	0.02	10	71	0.02	7
3PA1	40	0.00	5	31	0.00	5	27	0.00	5
3PA2	212	0.00	17	600	0.03	59	115	0.00	14
3PA3	687	0.09	19	385	0.05	15	833	0.11	60
3PA4	1996	0.22	33	812	0.13	38	203	0.03	7
3PB1	25	0.00	4	13	0.00	2	9	0.00	1
3PB2	122	0.00	8	53	0.00	4	111	0.00	17
3PB3	897	0.03	89	1003	0.11	81	912	0.14	130
3PB4	763	0.09	37	413	0.08	20	724	0.11	33
3PC1	11	0.00	1	12	0.00	1	9	0.00	1
3PC2	20	0.00	1	24	0.02	1	82	0.02	16
3PC3	153	0.02	7	103	0.05	6	126	0.02	8
3PC4	666	0.03	24	480	0.09	30	165	0.03	7

Tabela 2.2: Comportamento das versões EMRG e EASET do método enumerativo na resolução de LCPs NP-difíceis

versões. Nas Figuras 2.2, 2.3 e 2.4 são destacados os comportamento das três versões por dimensão dos problemas de Mochila em relação ao número de iterações, nós e tempo de CPU gasto na resolução desses problemas. Numa apreciação global destes gráficos podemos

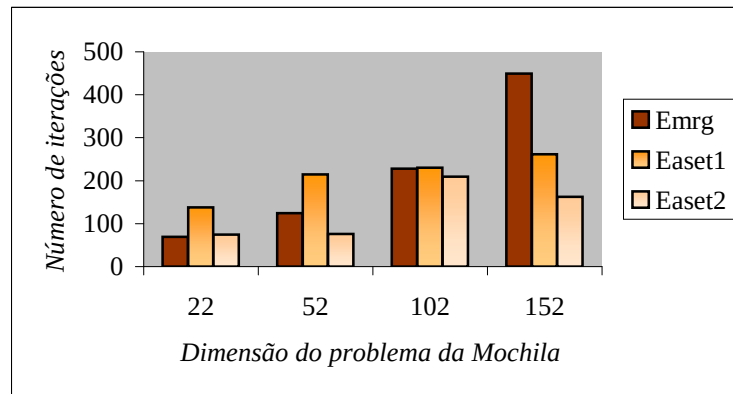


Figura 2.2: Comparação do número médio de iterações gasto por dimensão do problema

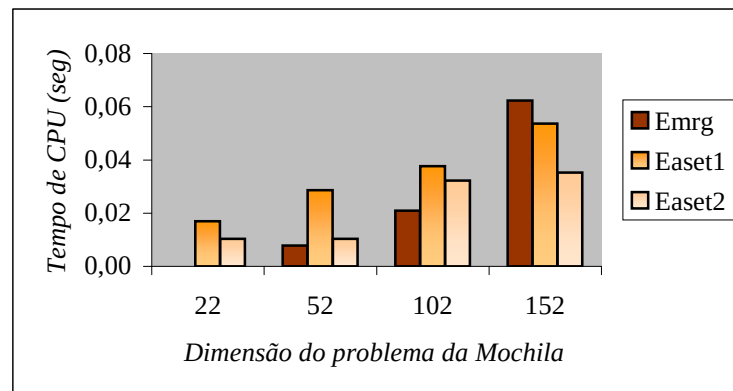


Figura 2.3: Comparação do tempo médio de CPU gasto por dimensão do problema

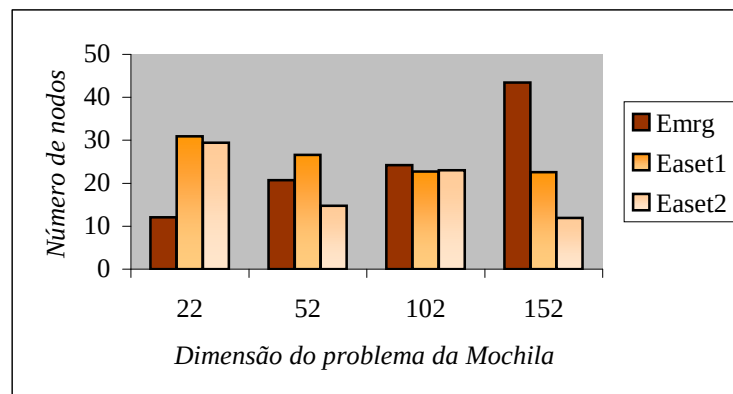


Figura 2.4: Comparação do número médio de nós gasto por dimensão do problema

concluir que EASET2 é a versão que requer menor esforço computacional para encontrar a solução. Essa conclusão torna-se mais notória à medida que a dimensão do LCP aumenta. Assim propomos a utilização de um método de restrições activas no método enumerativo e o uso dos parâmetros alternativos e das heurísticas discutidas nesta secção.

É importante acrescentar que esta versão EASET tem ainda a vantagem acrescida de resolver LCPs monótonos e com matrizes $RSuf$ no nó 1. Com efeito, um ponto estacionário do programa quadrático associado a esse nó é solução do LCP nesses casos.

Capítulo 3

PROBLEMA DE PROGRAMAÇÃO MATEMÁTICA COM RESTRIÇÕES DE COMPLEMENTARIDADE

Os Problemas de Programação Matemática com Restrições de Complementaridade ou de Equilíbrio (MPEC) são problemas de optimização nos quais as restrições incluem problemas complementares ou de desigualdades variacionais paramétricas. As desigualdades variacionais estão tipicamente relacionadas com fenómenos de equilíbrio que aparecem nas aplicações de engenharia e economia. Os MPECs são vistos como extensões dos programas de dois níveis, também conhecidos como programas matemáticos com restrições de optimização. Estes problemas tiveram origem nos anos sessenta num trabalho desenvolvido por K. Kirchgässner [74] e desde então têm sido objecto de intensa investigação. Um estudo mais detalhado sobre este tipo de problemas é apresentado em [85].

Neste capítulo iremos inicialmente apresentar a definição do MPEC e discutir algumas reduções de problemas de optimização a MPECs. Seguidamente faremos uma revisão de algumas técnicas locais para a determinação de um ponto estacionário de um MPEC. Em particular será discutido em detalhe um algoritmo de restrições activas modificado para esse fim. Finalmente alguns resultados computacionais obtidos na resolução de MPECs são incluídos de modo a testar a eficiência destes processos.

3.1 Definição do Programa com Restrições de Complementaridade

Na introdução deste capítulo foi referido que o Problema de Programação Matemática com Restrições de Equilíbrio contém restrições de desigualdade variacionais paramétricas, que constituem um VI paramétrico. Neste tipo de VI existe um parâmetro que pode variar num certo subconjunto do Espaço Euclidiano $\mathbb{R}^m$. Formalmente estes problemas podem ser definidos, a partir da definição 1.4, da seguinte forma:

Definição 3.1 *Seja $F : P \times D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ uma função de dois argumentos $(\mathbf{p}, z)$; seja $K : \mathbb{R}^m \rightarrow D$ uma "multifunção" com valores em D , isto é, para cada $\mathbf{p} \in \mathbb{R}^m$, $K(\mathbf{p})$ é um subconjunto (possivelmente vazio) de D .*

Uma Família Paramétrica de VIs é definida por

$$\{VI(K(\mathbf{p}), F(\mathbf{p}, z)) : \mathbf{p} \in P\}$$

onde $\mathbf{p}$ é o parâmetro que varia num dado conjunto P .

Assim, o Problema de Programação Matemática com Restrições de Equilíbrio, denotado por MPEC, é um problema de optimização que contém uma família paramétrica de VIs:

Definição 3.2 *Dada uma função com valores reais $f : \mathbb{R}^{m+n} \rightarrow \mathbb{R}$ e dois subconjuntos Y e Z de $\mathbb{R}^m$ e $\mathbb{R}^n$, respectivamente, o **Problema de Programação Matemática com Restrições de Equilíbrio (MPEC)** consiste em*

$$\begin{aligned} &\text{Minimizar} && f(y, z) \\ &\text{sujeito a} && y \in Y, z \in Z, \\ &&& z \in SOL(K(y), F(y, z)), \end{aligned} \tag{3.1}$$

onde $SOL(K(y), F(y, z))$ é o conjunto de soluções da família paramétrica de VIs com parâmetro y .

Um caso particular bastante importante do MPEC (3.1) surge quando $K(y)$ é um cone convexo em $\mathbb{R}^n$ para todo o $y \in Y$. Neste caso, o $VI(K(y), F(y, z))$ é equivalente ao problema complementar sobre o cone $K(y)$ [58]:

$$z \in K(y), \quad F(y, z) \in K(y)^*, \quad z^T F(y, z) = 0 \tag{3.2}$$

onde $K(y)^*$ é o cone dual de $K(y)$ definido em (1.7). Um dos casos mais interessantes do MPEC obtém-se quando se pode escrever $K(y)$ na forma $K(y) = \mathbb{R}^{n_1} \times \mathbb{R}_+^{n_2}$, com n_1 e n_2 inteiros não negativos tais que $n_1 + n_2 = n$. Particionando os vectores z e $F(y, z)$ do mesmo modo

$$z = \begin{pmatrix} z^1 \\ z^2 \end{pmatrix}, \quad F(y, z) = \begin{pmatrix} F_1(y, z) \\ F_2(y, z) \end{pmatrix}$$

com $z^1, F_1(y, z) \in \mathbb{R}^{n_1}$ e $z^2, F_2(y, z) \in \mathbb{R}^{n_2}$, o problema (3.2) reduz-se a:

$$\begin{aligned} F_1(y, z) &= 0 \\ z^2 &\geq 0, \quad F_2(y, z) \geq 0, \quad (z^2)^T F_2(y, z) = 0, \end{aligned} \tag{3.3}$$

que é um problema complementar não linear misto. Quando $n_1 = 0$ obtemos o problema complementar não linear NCP(F_2).

O objectivo deste trabalho é estudar e desenvolver algoritmos associados a programas matemáticos com restrições de equilíbrio afins em que $F(y, z)$ é uma função linear. Assim, o nosso estudo vai incidir sobre o seguinte problema

$$\begin{aligned} \text{MPEC} \quad & \text{Minimizar} \quad f(y, z) \\ & \text{sujeito a} \quad Ew = q + Mz + Ny \\ & \quad z \geq 0, \quad w \geq 0, \quad y \in K_y \\ & \quad z^T w = 0 \end{aligned} \tag{3.4}$$

onde $q \in \mathbb{R}^l$, $z, w \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, $E \in \mathbb{R}^{l \times n}$, $M \in \mathbb{R}^{l \times n}$, $N \in \mathbb{R}^{l \times m}$, $f : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ é duas vezes continuamente diferenciável num conjunto aberto que contém o conjunto admissível associado ao MPEC. Além disso $K_y \subseteq \mathbb{R}^m$ é um poliedro convexo em y , que ao longo deste trabalho irá ser definido da seguinte forma

$$K_y = \{y \in \mathbb{R}^m : Ay = b, y \geq 0\}$$

com $A \in \mathbb{R}^{p \times m}$ e $b \in \mathbb{R}^p$.

Notar que restrições de desigualdade na definição de K_y podem ser sempre reduzidas a igualdades por introdução de variáveis de folga sem acarretar qualquer dificuldade acrescida

ao problema. Como veremos neste e em capítulos posteriores, em muitas aplicações do MPEC as matrizes E e M são quadradas, sendo E a identidade e $M \in PSD$. A existência de restrições de complementaridade na formulação anterior do MPEC faz com que este problema também se designe por Problema de Programação Matemática com Restrições de Complementaridade. Consoante a função f seja linear ou não linear este problema é designado por MPEC linear ou não linear, respectivamente.

Na formulação original do MPEC, y é a variável (de desenho) do primeiro nível e z é a variável (de estado) do segundo nível. Na formulação linear do MPEC (3.4) as restrições de equilíbrio constituem um problema linear complementar paramétrico com parâmetro y e com as variáveis primárias z e w a satisfazer

$$0 \leq z \perp w \geq 0$$

Estas restrições são equivalentes a

$$z_i = 0 \quad \vee \quad w_i = 0 \quad \forall i$$

Podemos assim concluir que o MPEC (3.4) é um problema de optimização NP-difícil, uma vez que a determinação de uma sua solução admissível é um GLCP, que tem essa complexidade [69].

3.2 Redução de um Problema de Dois Níveis a um Programa com Restrições de Complementaridade

Nos problemas de programação de dois níveis existe uma hierarquia de problemas de optimização onde as restrições de um dos problemas (o chamado problema do nível superior ou do primeiro nível) são definidas em parte por um segundo problema de optimização paramétrico (o problema do nível inferior ou do segundo nível). Este tipo de problemas ocorre em diversas aplicações nas mais distintas áreas da ciência, entre as quais se destacam a engenharia civil, a engenharia química e a economia [77, 78, 45].

Um programa de programação de dois níveis (BP), na sua forma mais geral, apresenta a seguinte definição:

Definição 3.3 *Dados os subconjuntos X e Y de $\mathbb{R}^n$ e $\mathbb{R}^m$, respectivamente, as funções reais f e θ tais que $f, \theta : \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ e as funções vectoriais reais g e h tais que $g : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^r$ e $h : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^s$ com n, m, r e $s \in \mathbb{N}$, o **Problema de Programação de Dois Níveis (BP)** consiste em:*

$$\begin{aligned} & \text{Minimizar}_{x,y} && f(x, y) \\ & \text{sujeito a} && x \in X \\ & && g(x, y) \geq 0 \\ & && y \in \operatorname{argmin}\{\theta(x, y) : y \in Y, h(x, y) \geq 0\} \end{aligned}$$

Os processos de resolução do BP mais frequentes são os que exploram a redução destes problemas a outros problemas equivalentes de otimização. Quando as funções envolvidas são continuamente diferenciáveis, a função $\theta(x, y)$ é convexa para cada $x \in X$ e $h(x, y)$ é linear para todos x e y , o problema do segundo nível pode ser substituído pelas suas condições necessárias e suficientes de optimalidade. Deste modo, o BP é equivalente ao seguinte MPEC

$$\begin{aligned} & \text{Minimizar} && f(x, y) \\ & \text{sujeito a} && g(x, y) \geq 0 \\ & && \nabla_y \theta(x, y) - \mu^T \nabla_y h(x, y) = 0 \\ & && h(x, y) \geq 0, \mu \geq 0 \\ & && \mu^T h(x, y) = 0 \\ & && y \in Y, x \in X \end{aligned} \tag{3.5}$$

onde $\mu \in \mathbb{R}^s$. Se a função θ do segundo nível é quadrática, a função g é linear e X e Y são poliedros convexos, então as restrições do problema (3.5) são lineares e obtém-se um MPEC da forma (3.4), com $M \in PSD$. Se, além disso $Y = \mathbb{R}_+^m$, então E é a matriz identidade.

3.3 Redução de um Problema Bilinear a um Programa com Restrições de Complementaridade

O Problema de Programação Bilinear aparece como uma extensão natural do problema de programação linear [79]. Esse problema é normalmente definido na seguinte forma

Definição 3.4 *Sejam $c \in \mathbb{R}^n$, $d \in \mathbb{R}^m$, $A \in \mathbb{R}^{t \times n}$, $B \in \mathbb{R}^{t \times m}$, $C \in \mathbb{R}^{n \times m}$ e $a \in \mathbb{R}^t$.*

O Problema de Programação Bilinear *consiste em*

$$\begin{aligned} & \text{Minimizar} && c^T x + d^T y + x^T C y \\ & \text{sujeito a} && Ax + By = a \\ & && x \geq 0 \quad y \geq 0 \end{aligned}$$

Num caso particular importante as restrições nas variáveis x e y da função bilinear estão separadas em dois conjuntos S_1 e S_2 , respectivamente, que definimos, por conveniência, da seguinte forma

$$S_1 = \{x \in \mathbb{R}^n : Ax \geq a, x \geq 0\} \quad \text{e} \quad S_2 = \{y \in \mathbb{R}^m : By \geq b, y \geq 0\}$$

Um programa bilinear que possui esta característica designa-se por Programa Bilinear Separado e é definido a seguir.

Definição 3.5 *Sejam $c \in \mathbb{R}^n$, $d \in \mathbb{R}^m$, $A \in \mathbb{R}^{r \times n}$, $B \in \mathbb{R}^{s \times m}$, $C \in \mathbb{R}^{n \times m}$, $a \in \mathbb{R}^r$ e $b \in \mathbb{R}^s$. O Problema de Programação Bilinear Separado, designado por BLP, consiste em*

$$\begin{aligned} & \text{Minimizar} && c^T x + d^T y + x^T C y \\ & \text{sujeito a} && Ax \geq a, x \geq 0 \\ & && By \geq b, y \geq 0 \end{aligned} \tag{3.6}$$

Historicamente, foi H. Mills quem pela primeira vez discutiu a utilidade dos problemas bilineares, quando procurava determinar um ponto de equilíbrio do jogo de duas matrizes [87]. Posteriormente, vários matemáticos debruçaram-se sobre este tipo de problemas, dos quais se destacam O. Mangasarian, H. Stone, M. Altman, A. Cabot e R. Francis, tendo proposto algoritmos para obtenção de ótimos locais e globais. Um destes algoritmos deve-se a Konno [80] e procura determinar um ponto estacionário do BLP a partir da resolução de um número finito de programas lineares alternados em x e em y . Esses problemas lineares podem ser resolvidos através do método simplex ou de um método de pontos-interiores, sendo o primeiro mais utilizado, uma vez que na sucessão desses problemas os conjuntos de restrições são sempre os mesmos. Este algoritmo teve como base o seguinte teorema [80]

Teorema 3.1 *Se o BLP tem uma solução óptima, então existe uma solução óptima (x^*, y^*) tal que x^* e y^* são pontos extremos de S_1 e S_2 respectivamente.*

Embora seja simples a aplicação deste algoritmo, o ponto estacionário obtido não é normalmente um óptimo global do programa bilinear. No entanto, esse processo pode ser útil em métodos sequenciais que tentem determinar a solução óptima do programa bilinear pesquisando um subconjunto dos seus pontos estacionários.

Um outro processo para a resolução deste tipo de problema e que será objecto de estudo neste capítulo, é o que se baseia na sua redução a um MPEC. Para isso, consideramos o BLP reescrito na seguinte forma

$$\begin{aligned} & \min_y \{d^T y + g(y), \quad y \in S_2\} \\ \text{com} \quad & g(y) = \min_x \{(c + Cy)^T x, \quad x \in S_1\}, \end{aligned} \quad (3.7)$$

O dual do programa linear (3.7) consiste em

$$\begin{aligned} & \max_u \quad a^T u \\ \text{sujeito a} \quad & A^T u \leq c + Cy \\ & u \geq 0 \end{aligned} \quad (3.8)$$

De acordo com a teoria da dualidade da programação linear, se (3.7) tem uma solução óptima para qualquer $y \in S_2$, então x é uma solução óptima de (3.7) se e só se $x \in S_1$, u é uma solução dual admissível e

$$\alpha^T x = \beta^T u = 0$$

onde β e α são as variáveis desvio do problema (3.7) e do seu dual (3.8), respectivamente. Além disso $(c + C)^T x = a^T u$ e podemos reformular estas condições da seguinte forma:

$$\begin{aligned} \alpha &= c + Cy - A^T u \\ \beta &= -a + Ax \\ \alpha, \beta, u, y, x &\geq 0 \\ \alpha^T x &= \beta^T u = 0 \\ g(y) &= a^T u \end{aligned}$$

Portanto, se S_1 for limitado, o BLP é equivalente ao seguinte MPEC:

$$\begin{aligned}
& \text{Minimizar} && d^T y + a^T u \\
& \text{sujeito a} && \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} c \\ -a \end{bmatrix} + \begin{bmatrix} 0 & -A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + \begin{bmatrix} C \\ 0 \end{bmatrix} y \\
& && \alpha, \beta, u, x \geq 0 \\
& && \alpha^T x = \beta^T u = 0 \\
& && By \geq b, y \geq 0
\end{aligned}$$

onde $x, \alpha, c \in \mathbb{R}^n$, $d, y \in \mathbb{R}^m$, $u, \beta, a \in \mathbb{R}^r$, $b \in \mathbb{R}^s$, $A \in \mathbb{R}^{r \times n}$, $B \in \mathbb{R}^{s \times m}$ e $C \in \mathbb{R}^{n \times m}$.

Notar que este MPEC tem a forma (3.4), com $M \in PSD$ e E a matriz identidade.

3.4 Redução de um Problema Linear Complementar a um Programa com Restrições de Complementaridade

Consideremos novamente o LCP

$$\begin{aligned}
w &= q + Mz, \quad z \geq 0, \quad w \geq 0 \\
z^T w &= 0
\end{aligned} \tag{3.9}$$

Tal como é referido em [86], o LCP é equivalente ao seguinte Problema Linear Complementar Aumentado (ALCP)

$$\begin{aligned}
\begin{bmatrix} \beta \\ \gamma \end{bmatrix} &= \begin{bmatrix} e \\ q \end{bmatrix} + \begin{bmatrix} 0 & -I \\ M & 0 \end{bmatrix} \begin{bmatrix} z \\ y \end{bmatrix} \\
\beta, \gamma, z, y &\geq 0 \\
\gamma^T y &= z^T \beta = 0
\end{aligned} \tag{3.10}$$

Por outro lado, o ALCP é equivalente ao Programa Bilinear Separado

$$\begin{aligned}
& \text{Minimizar} && g(z, y) = e^T z + q^T y + y^T (M - I) z \\
& \text{sujeito a} && q + Mz \geq 0, \quad z \geq 0 \\
& && 0 \leq y \leq e
\end{aligned}$$

De facto:

Teorema 3.2 $(\bar{\beta}, \bar{\gamma}, \bar{z}, \bar{y})$ é solução do ALCP se e só se $(\bar{z}, \bar{y})$ é solução óptima do BLP com valor óptimo igual a zero.

Demonstração:

Para provar esta equivalência basta observar que as restrições lineares do ALCP e BLP coincidem e $g(z, y) = \gamma^T y + z^T \beta$, para qualquer solução admissível dos dois problemas. $\diamond$

De acordo com o que foi mencionado na secção anterior, este BLP é equivalente ao seguinte MPEC

$$\begin{aligned} & \text{Minimizar} && q^T y - q^T u \\ & \text{sujeito a} && \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} e \\ q \end{bmatrix} + \begin{bmatrix} 0 & -M^T \\ M & 0 \end{bmatrix} \begin{bmatrix} z \\ u \end{bmatrix} + \begin{bmatrix} M^T - I \\ 0 \end{bmatrix} y \\ & && \alpha, \beta, u, y, z \geq 0 \\ & && \alpha^T z = \beta^T u = 0 \\ & && 0 \leq y \leq e \end{aligned}$$

Notar que, tal como anteriormente, as matrizes associadas às variáveis complementares são PSD e a identidade. Do teorema anterior podemos concluir o seguinte resultado:

Teorema 3.3 $(\bar{z}, \bar{w})$ é solução do LCP se e só se $(\bar{\alpha}, \bar{\beta}, \bar{v}, \bar{u}, \bar{y}, \bar{z})$ é solução do MPEC com valor óptimo igual a zero.

3.5 Algoritmos de Programação Não Linear

O MPEC ocorre com frequência em vários domínios científicos e técnicos, designadamente em engenharia e economia. Por este motivo, desde há vários anos têm sido propostos diversos algoritmos que garantem óptimos locais na resolução desses problemas. Nesta secção faz-se um inventário dos mais importantes destes algoritmos, assumindo-se para o MPEC a seguinte formulação:

$$\begin{aligned} & \text{Minimizar} && f(x) \\ & \text{sujeito a} && h_1(x) = 0 \\ & && h_2(x) \geq 0 \\ & && x^1 \geq 0, x^2 \geq 0 \\ & && x_i^1 x_i^2 = 0, \quad i = 1, \dots, n \end{aligned} \tag{3.11}$$

onde $x = (x^0, x^1, x^2)$ é uma decomposição das variáveis do problema nas variáveis de desenho $x^0 \in \mathbb{R}^m$ e nas de estado $(x^1, x^2) \in \mathbb{R}^{2n}$.

As monografias [85, 97] dão-nos um estudo amplo deste tipo de problemas e apresentam vários algoritmos locais. Uma vez que (3.11) é um programa não linear, as primeiras experiências numéricas efectuadas para a resolução do MPEC utilizaram algoritmos de programação não linear (PNL). No entanto, como foi demonstrado em [21], uma das condições suficientes para a estabilidade de um PNL é violada em todos os pontos admissíveis deste problema, uma vez que esses pontos não satisfazem a chamada Restrição de Qualificação de Mangasarian-Fromovitz (MFCQ) [107]. Por isso podem surgir várias dificuldades quando se tenta resolver o MPEC através de algoritmos de programação não linear. Uma outra proposta para a resolução do MPEC (3.11) consiste na substituição das condições de complementaridade por uma inequação não linear ou por conjunto de inequações não lineares, tais como $(x^1)^T x^2 \leq 0$ ou $X_1 x^2 \leq 0$, onde $X_1 = \text{diag}(x^1)$. No entanto, a Restrição de Qualificação de Mangasarian-Fromovitz para os programas não lineares resultantes continua a não ser verificada.

As primeiras experiências numéricas para a resolução de problemas de dois níveis usando um método de gradiente projectado para o PNL (3.11) redundaram num fracasso [8]. Recentemente novos desenvolvimentos têm demonstrado que, contrariamente aos resultados iniciais, é possível resolver localmente com grande confiança e eficácia uma grande classe de MPECs através das suas reformulações PNL [42, 4]. O reaparecimento do interesse na abordagem de programação não linear na resolução do MPEC foi suscitado pelo sucesso dos métodos de Programação Quadrática Sequenciais (SQP). Assim Jiang e Ralph [70] propuseram dois métodos SQP diferenciáveis para resolver MPECs (3.11), cuja restrição de complementaridade é substituída pela função diferenciável de Fischer-Burmeister

$$\Psi_\tau(x_{1i}, x_{2i}) = \sqrt{x_{1i}^2 + x_{2i}^2} + \tau - x_{1i} - x_{2i} = 0.$$

O primeiro método aplica a técnica SQP ao problema, onde τ é um parâmetro que é reduzido em cada iteração, enquanto que o segundo processo considera uma restrição adicional $e^\tau - 1 = 0$, com τ uma variável. Esta nova equação garante que $\tau = 0$ no óptimo. Fukushima et al. [46] propuseram um outro algoritmo sequencial de programação quadrática, para reformular as condições de complementaridade como um sistema de equações semidiferenciáveis, desenvolvido com o intuito de obter um ponto estacionário para o MPEC, usando a funcional

de Fischer-Burmeister [38, 39, 40].

O algoritmo FILTERSQP é uma outra técnica SQP que tem sido utilizada com algum sucesso na resolução de MPECs. Esse processo utiliza regiões de confiança e explora o conceito de filtro, que só aceita um determinado ponto se houver redução no valor da função objectivo ou da função de violação das restrições. Convém salientar que normalmente estas duas funções são combinadas através de uma função penalty que penaliza a violação das restrições. No entanto, neste algoritmo as duas funções são consideradas em separado, dando origem a um problema de optimização multi-objectivo.

Uma outra abordagem que tem sido explorada consiste em substituir a restrição $(x^1)^T x^2 \leq 0$ por $(x^1)^T x^2 \leq \tau_k$ e resolver uma sequência de PNLs, com τ_k a tender para zero [37, 107]. Uma alternativa é penalizar a restrição de complementaridade, resolvendo uma sequência de PNLs, onde a função objectivo é a seguinte função penalty

$$\text{Minimizar } f(x) + \nu_k (x^1)^T x^2$$

para uma sequência crescente de parâmetros penalty $\nu_k > 0$. Esta aproximação é também discutida em [4], onde se mostra que sob suposições razoáveis, a penalização é exacta.

Uma outra ideia interessante deve-se a Facchinei et al. [34] que consiste em substituir a restrição de complementaridade por uma função diferenciável

$$\xi_\tau(x_{1i}, x_{2i}) = \sqrt{(x_{1i} - x_{2i})^2 + 4\tau} - x_{1i} - x_{2i} = 0$$

onde $\tau > 0$ é um parâmetro. Para $\tau = 0$, pode ver-se que $\xi_0(a, b) = -2\min(a, b)$. O PNL resultante satisfaz a restrição de MFCQ e é diferenciável com os gradientes limitados, desde que $(x_{1i}, x_{2i}, \tau_k) \neq 0$. Em [34], são sugeridos vários algoritmos baseados na resolução de sequências de PNLs, para valores decrescentes $\tau_k \rightarrow 0$ e estudadas as suas propriedades de convergência.

A aproximação de programação implícita, devida a Outrata et al. [97], é outra abordagem para o MPEC que consiste em resolver o problema de complementaridade do nível inferior fixando as variáveis do primeiro nível e calculando as sensibilidades dessas variáveis. No problema do nível superior, um problema de optimização não diferenciável é então resolvido usando técnicas de regiões de confiança.

Os Métodos de Pontos Interiores (IPM) não são tão robustos como os algoritmos SQPs

para resolver o MPEC. Contudo, os resultados obtidos com estes métodos são promissores, apesar da restrição de qualificação de Mangasarian-Fromovitz não ser satisfeita e o caminho central não existir. A razão para a não existência deste caminho central é a condição de complementaridade, uma vez que as restrições de desigualdade pela qual é substituída juntamente com as condições de não negatividade das variáveis complementares faz com que não exista interior no conjunto admissível.

Um dos processos desta classe mais referenciados na literatura é o Algoritmo Penalty de Pontos Interiores (PIPA) [85]. Este método esforça-se por manter $x^1 > 0$ e $x^2 > 0$ durante a utilização de uma aproximação SQP para resolver o PNL. Em cada passo resolve-se um programa quadrático para obter uma direcção de pesquisa. A pesquisa em linha nesta direcção é realizada com o objectivo de manter a admissibilidade estrita e de reduzir suficientemente a função penalty $f(x) + \gamma(x^1)^T x^2$ onde o parâmetro penalty $\gamma > 0$ é actualizado em cada iteração. No entanto, como é discutido em [84], o PIPA pode convergir para pontos não estacionários em alguns problemas. Uma melhoria desse processo foi proposta em [46] para lidar com MPECs de forma semelhante à apresentada em (3.4).

Em [42] são apresentados resultados com dois algoritmos de pontos interiores e dois algoritmos SQPs, implementados nos códigos KNITRO [17], LOQO [117], FILTERSQP [41, 43] e SNOPT [49] na resolução de 137 MPECs. Os métodos SQP não conseguiram obter a solução somente em 7 e 5 casos, respectivamente, enquanto que os algoritmos de pontos interiores falharam mais vezes, respectivamente em 25 e 20 casos. Nesse artigo também se pode constatar que os métodos SQP parecem ser mais robustos do que alguns dos algoritmos especialmente desenvolvidos para MPECs que têm sido publicados na literatura. Além disso, Dirkse et al. [31] investigaram técnicas automáticas de reformulação de MPECs em linguagem GAMS. Nesse artigo são consideradas algumas reformulações diferenciáveis e é incluída experiência computacional com um grande número de MPECs.

3.6 Algoritmo de Restrições Activas Modificado

Nesta secção é discutido em detalhe um Algoritmo de Restrições Activas Modificado ASETM para a determinação de um ponto estacionário de um MPEC, isto é, de uma solução satisfazendo as condições necessárias de primeira ordem do PNL que corresponde ao MPEC (3.4)

considerando as restrições de complementaridade $z_i w_i = 0, i = 1, \dots, n$ como restrições não lineares. Assim esse PNL tem a forma

$$\begin{aligned}
 \text{PNL} \quad & \text{Minimizar} \quad f(y, z) \\
 & \text{sujeito a} \quad Ew = q + Mz + Ny \\
 & \quad \quad \quad Ay = b \\
 & \quad \quad \quad z \geq 0, \quad w \geq 0, \quad y \geq 0 \\
 & \quad \quad \quad z_i w_i = 0, \quad i = 1, \dots, n
 \end{aligned} \tag{3.12}$$

onde $q \in \mathbb{R}^l$, $z, w \in \mathbb{R}^n$, $b \in \mathbb{R}^p$, $y \in \mathbb{R}^m$, $M \in \mathbb{R}^{l \times n}$, $E \in \mathbb{R}^{l \times n}$, $N \in \mathbb{R}^{l \times m}$ e $A \in \mathbb{R}^{p \times m}$ com $p < m$. Este algoritmo é semelhante a um processo introduzido por Scholtes [108] e consiste essencialmente em utilizar uma técnica de restrições activas no conjunto de soluções do GLCP do MPEC. Assim em cada iteração k , os iterandos (w, z, y) satisfazem as restrições do PNL anterior e o conjunto de restrições activas é constituído por

$$\left. \begin{aligned}
 Ew - Mz - Ny &= q \\
 Ay &= b \\
 w_i &= 0, \quad i \in L_w \subseteq \{1, \dots, n\} \\
 z_i &= 0, \quad i \in L_z \subseteq \{1, \dots, n\} \\
 y_i &= 0, \quad i \in L_y \subseteq \{1, \dots, m\}
 \end{aligned} \right\} \tag{3.13}$$

onde L_w, L_z e L_y são os conjuntos das restrições activas correspondentes às restrições de não negatividade nas variáveis w, z e y , respectivamente.

Estas restrições activas (3.13) constituem um sistema linear da seguinte forma:

$$D_k x = g^k$$

onde $x = (w^T, z^T, y^T)^T$ e $D_k \in \mathbb{R}^{t \times (2n+m)}$, com $t = l + p + |L_w| + |L_z| + |L_y| + |H|$ é o cardinal do conjunto H .

As condições de optimalidade de primeira ordem para o problema

$$\text{Minimizar} \left\{ f(x) : D_k x = g^k \right\}$$

podem ser escritas na forma

$$\begin{aligned}\nabla f(x) &= D_k^T \mu \\ D_k x &= g^k.\end{aligned}\tag{3.14}$$

Tal como referimos no capítulo anterior, para que seja possível calcular univocamente o vector dos multiplicadores de Lagrange μ iremos impor a seguinte condição

Hipótese de Não Degenerescência:

$$t \leq 2n + m \quad \text{e} \quad \text{característica}(D_k) = t$$

Deste modo, o conjunto das restrições activas é sempre linearmente independente. Além disso, particionemos o vector dos Multiplicadores de Lagrange μ em três subvectores denotados por

$$\begin{aligned}\beta &\rightarrow \text{subvector associado ao primeiro conjunto das restrições de igualdade} \\ \vartheta &\rightarrow \text{subvector associado ao segundo conjunto das restrições de igualdade} \\ \lambda_i^x &\rightarrow \text{subvector associado a } x_i = 0 \text{ para } x = z, y, w\end{aligned}$$

Os passos do algoritmo complementar de restrições activas são apresentados a seguir.

ALGORITMO DE RESTRIÇÕES ACTIVAS MODIFICADO - ASETM

Passo 0

Faça $k = 1$ e determine uma solução x^k do GLCP associado ao MPEC. Seja $D_k x = g^k$ o conjunto das restrições activas em x^k que satisfazem a hipótese de não degenerescência e L_y , L_z e L_w os conjuntos de índices correspondentes às restrições activas de não negatividade $y_i = 0$, $z_i = 0$ e $w_i = 0$, respectivamente.

Passo 1 *Condições de optimalidade*

Se x^k não é um ponto estacionário do Problema de Igualdades

$$\begin{aligned}PI \quad &\text{Minimizar} \quad f(x) \\ &\text{sujeito a} \quad D_k x = g^k.\end{aligned}\tag{3.15}$$

vá para o Passo 2. De outro modo existe um único μ tal que

$$D_k^T \mu = \nabla f(x^k).$$

Dois casos podem acontecer:

1. Se

$$\begin{aligned}\lambda_i^y &\geq 0 && \text{para todo } i \in L_y \\ \lambda_i^z &\geq 0 && \text{para todo } i \in L_z \cap L_w \\ \lambda_i^w &\geq 0 && \text{para todo } i \in L_z \cap L_w\end{aligned}$$

páre: x^k é um ponto estacionário do MPEC.

2. Se existe pelo menos um i tal que

$$\begin{aligned}\lambda_i^y &< 0 && \text{para } i \in L_y \\ \text{ou } \lambda_i^z &< 0 && \text{para } i \in L_z \cap L_w \\ \text{ou } \lambda_i^w &< 0 && \text{para } i \in L_z \cap L_w\end{aligned}$$

retire uma restrição activa $y_i = 0$ ou $z_i = 0$ ou $w_i = 0$, associada ao multiplicador de Lagrange mais negativo. Seja $D_{k_i}x = g_i^k$ a linha que é retirada de $D_kx = g^k$, e reordene as linhas da seguinte forma

$$D_k = \begin{bmatrix} \bar{D}_k \\ D_{k_i} \end{bmatrix}, \quad g^k = \begin{bmatrix} \bar{g}^k \\ g_i^k \end{bmatrix}.$$

Encontre a direcção d tal que $\nabla f(x^k)^T d < 0$, $\bar{D}_k d = 0$, e $D_{k_i} d > 0$ (o teorema 3.5 mostra que d existe). Substitua D_k por $\bar{D}_k$ e vá para o Passo 3.

Passo 2 *Cálculo da Direcção de Pesquisa*

Encontre uma direcção descendente para f no conjunto das restrições activas, isto é, determine d tal que:

$$\begin{aligned}\nabla f(x^k)^T d &< 0 \\ D_k d &= 0\end{aligned}$$

Passo 3 *Cálculo do Passo*

1. Determine o maior valor α_{max} de α tal que:

$$x^k + \alpha d \geq 0$$

a partir de

$$\alpha_{max} = \min \left\{ \frac{x_i^k}{-d_i} : d_i < 0 \right\} \quad (3.16)$$

2. Calcule $0 < \alpha_k \leq \alpha_{max}$ tal que

$$x^k + \alpha_k d$$

proporcione um decréscimo significativo para f , usando uma Pesquisa Unidireccional.

Se $\alpha_k = +\infty$, páre. O MPEC é ilimitado.

Passo 4 *Actualização do iterando*

Faça

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

Se $\alpha_k = \alpha_{max}$, adicione ao conjunto de restrições activas, as restrições $x_i \geq 0$ onde α_{max} foi atingido, de forma a que a hipótese de não degenerescência se mantenha verdadeira. Volte ao Passo 1.

3.7 Convergência do Método de Restrições Activas Modificado

Nesta secção será demonstrada a convergência do Método de Restrições Activas Modificado para um ponto estacionário do MPEC. Nesse sentido, será inicialmente estabelecida uma condição suficiente para um ponto estacionário do problema (3.15) ser um ponto estacionário do MPEC.

Teorema 3.4 *Em qualquer iteração k , seja x^k um ponto estacionário para o problema*

$$\text{Minimizar } \left\{ f(x) : D_k x = g^k \right\} \quad (3.17)$$

onde $D_k x = g^k$ representa todas as restrições activas do MPEC em x^k , e as componentes de x^k para as restrições inactivas são positivas. Se $\lambda_i^y \geq 0 \quad \forall i \in L_y$ e $(\lambda_i^z, \lambda_i^w) \geq 0 \quad \forall i \in L_z \cap L_w$, então x^k é um ponto estacionário para o MPEC.

Demonstração:

Por definição, x^k é admissível para as restrições lineares do MPEC. Além disso x^k satisfaz as restrições de complementaridade uma vez que $L_z \cup L_w \equiv \{1, \dots, n\}$ em cada iteração. Então, pelas restrições de um ponto estacionário para (3.17) tem-se que

$$D_k^T \mu = \nabla f(x^k). \quad (3.18)$$

Para provar que x^k é um ponto estacionário do MPEC, é preciso mostrar que existem $\bar{\mu}$ e $\bar{\psi}$ tais que

$$D_k^T \bar{\mu} + \sum_{i=1}^n [\xi_i] \bar{\psi}_i = \nabla f(x^k) \quad (3.19)$$

$$\bar{\lambda}_i^y \geq 0 \quad \forall i \in L_y, \bar{\lambda}_i^z \geq 0 \quad \forall i \in L_z, \bar{\lambda}_i^w \geq 0 \quad \forall i \in L_w \quad (3.20)$$

onde para cada $i = 1, \dots, n$

$$[\xi_i] = \begin{bmatrix} 0 & \dots & 0 & w_i^k & z_i^k & 0 & \dots & 0 \end{bmatrix}^T \quad (3.21)$$

$\uparrow \quad \uparrow$
 posição para z_i posição para w_i

Seja

$$\bar{\lambda}_i^y = \lambda_i^y \quad \forall i \in L_y \quad (3.22)$$

$$\bar{\lambda}_i^z = \begin{cases} \lambda_i^z & \text{se } \lambda_i^z \geq 0 \\ 0 & \text{caso contrário} \end{cases} \quad \forall i \in L_z \quad (3.23)$$

$$\bar{\lambda}_i^w = \begin{cases} \lambda_i^w & \text{se } \lambda_i^w \geq 0 \\ 0 & \text{caso contrário} \end{cases} \quad \forall i \in L_w \quad (3.24)$$

$$\bar{\psi}_i = \begin{cases} \frac{\lambda_i^z}{w_i^k} & \text{se } \lambda_i^z < 0 \\ \frac{\lambda_i^w}{z_i^k} & \text{se } \lambda_i^w < 0 \\ 0 & \text{caso contrário} \end{cases} \quad \forall i = 1, \dots, n \quad (3.25)$$

Então (3.20) é satisfeita e $\bar{\psi}$ está bem definido, uma vez que $\{i \in L_z : \lambda_i^z < 0\} \cap \{i \in L_w : \lambda_i^w < 0\} = \emptyset$. Com efeito $(\lambda_i^z, \lambda_i^w) \geq 0 \quad \forall i \in L_z \cap L_w$. Além disso, $\lambda_i^z < 0 \Rightarrow i \in L_z \setminus (L_z \cap L_w) \Rightarrow i \notin L_w \Rightarrow w_i^k > 0$, e analogamente, $\lambda_i^w < 0 \Rightarrow z_i^k > 0$. Por outro lado, se $\lambda_i^z < 0$, seja e_i^z o vector unitário de $\mathbb{R}^n$ que representa a coluna de D_k^T correspondente a λ_i^z em (3.18). De (3.22) e (3.25) vem

$$[e_i^z] \lambda_i^z = [\xi_i] \bar{\psi}_i. \quad (3.26)$$

Analogamente

$$\lambda_i^w < 0 \Rightarrow [e_i^w] \lambda_i^w = [\xi_i] \bar{\psi}_i. \quad (3.27)$$

Agora, de (3.18) e de (3.22) a (3.27) tem-se que $\bar{\mu}$ e $\bar{\psi}$ dados por (3.22)-(3.25) satisfazem (3.19) e (3.20), o que completa a demonstração. $\diamond$

A seguir iremos mostrar que sempre que as condições suficientes identificadas no teorema 3.4 não são verificadas o valor da função objectivo do MPEC pode ser melhorado no Passo 1.

Teorema 3.5 *Em qualquer iteração k , considere-se a situação descrita no teorema 3.4 e que existe um multiplicador de Lagrange negativo, μ_i , tal que μ_i é λ_i^y para $i \in L_y$ ou λ_i^z para $i \in L_z \cap L_w$ ou λ_i^w para $i \in L_z \cap L_w$. Designemos por $D_{k_i}x = g_i^k$ a linha correspondente de $D_kx = g^k$, e reordenemos as linhas da seguinte forma:*

$$D_k = \begin{bmatrix} \bar{D}_k \\ D_{k_i} \end{bmatrix}, g^k = \begin{bmatrix} \bar{g}^k \\ g_i^k \end{bmatrix}, \mu_k = \begin{bmatrix} \bar{\mu} \\ \mu_i \end{bmatrix} \quad (3.28)$$

Então, o problema

$$\text{Minimizar } \left\{ f(x) : \bar{D}_k x = \bar{g}^k \right\} \quad (3.29)$$

tem uma direcção admissível descendente d que satisfaz

$$\nabla f(x^k)^T d < 0, \quad \bar{D}_k d = 0 \quad e \quad D_{k_i} d > 0 \quad (3.30)$$

Demonstração:

Notar que uma das condições de *KKT* para (3.17) é

$$\bar{D}_k^T \bar{\mu} + D_{k_i}^T \mu_i = \nabla f(x^k) \quad (3.31)$$

Seja d uma solução qualquer do sistema

$$\begin{bmatrix} \bar{D}_k \\ D_{k_i} \end{bmatrix} d = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad (3.32)$$

que existe pela hipótese de não degenerência. Além disso, de (3.31) e (3.32), tem-se

$$\nabla f(x^k)^T d = \bar{\mu}^T \bar{D}_k d + \mu_i D_{k_i} d = \mu_i < 0 \quad (3.33)$$

Mas, de (3.32), $\bar{D}_k d = 0$ e $D_{k_i} d > 0$. Donde (3.30) é satisfeita, o que completa a demonstração. $\diamond$

O teorema a seguir investiga a convergência do algoritmo ASETM num número finito de iterações para um ponto estacionário do MPEC, caso exista.

Teorema 3.6 *Consideremos o Algoritmo ASETM e suponhamos que o ponto estacionário do problema (3.17) satisfaz a condição do teorema 3.4 e é finitamente determinado (se existir). Além disso, suponhamos que o conjunto dos pontos estacionários para qualquer um destes problemas (3.17) aceita apenas um número finito de valores da função objectivo. Então o ASETM termina num número finito de iterações com um ponto estacionário do MPEC ou com a indicação de que o MPEC é inadmissível ou ilimitado.*

Demonstração:

Se o GLCP é inadmissível então também o é o MPEC. De outro modo o MPEC é admissível. Se ao resolver o problema (3.17), for detectado que o problema é ilimitado, então o MPEC também o é pois, a condição de complementaridade é satisfeita em cada iteração. Caso contrário, o algoritmo volta ao Passo 1 num número finito de iterações.

Neste caso, ou é detectado um ponto estacionário para o MPEC no Passo 1, ou é determinada uma direcção descendente admissível d que satisfaz (3.30) do teorema 3.5. No primeiro caso, o algoritmo termina, enquanto que no último caso, pela restrição (3.30) e pelo facto de $\alpha_{max} > 0$, o próximo iterando x^{k+1} irá satisfazer $f(x^{k+1}) < f(x^k)$. Esta propriedade juntamente com a hipótese de que é finito o número de valores possíveis da função objectivo no conjunto de pontos estacionários para qualquer combinação de restrições activas, garante que dado qualquer conjunto de restrições activas é finito o número de passagens pelo Passo 1. Como o número de possíveis conjuntos de restrições activas é finito, a demonstração está completa. $\diamond$

Um caso particular deste teorema é apresentado a seguir.

Teorema 3.7 *Se f é convexa ou se o ponto estacionário x^k do problema (3.17) é um mínimo global, em cada iteração, então o algoritmo ASETM é finitamente convergente.*

Demonstração:

Se f é convexa, então os pontos estacionários para o problema (3.17) têm um único valor para a função objectivo. Por outro lado, se em cada iteração o ponto estacionário de (3.17) é um seu mínimo global, então a demonstração segue outra vez o argumento apresentado na demonstração do teorema 3.6. $\diamond$

Notar que se f não é convexa, nem sempre é garantida a determinação de um mínimo global do problema (3.17). Na próxima secção, iremos mostrar que essa propriedade é verdadeira se a função objectivo do MPEC é quadrática.

3.8 Caso Especial da Função Objectivo Quadrática

Suponhamos que f é uma função quadrática da seguinte forma

$$f(x) = c^T x + \frac{1}{2} x^T H x$$

onde a Hessiana H pode ser indefinida. Então o Problema (3.17) é da forma

$$\begin{array}{ll} QP & \text{Minimizar} \quad c^T x + \frac{1}{2} x^T H x \\ & \text{sujeito a} \quad Dx = g \end{array} \quad (3.34)$$

e as condições de KKT são:

$$\begin{array}{l} c + Hx - D^T \mu = 0 \\ Dx = g \end{array} \quad (3.35)$$

Para qualquer ponto estacionário x de (3.34), tem-se

$$f(x) = \frac{1}{2} c^T x + \frac{1}{2} (c^T x + x^T H x) = \frac{1}{2} c^T x + \frac{1}{2} \mu^T D x = \frac{1}{2} (c^T x + \mu^T g) \quad (3.36)$$

Deste modo, o problema de encontrar a melhor solução para QP reduz-se ao programa linear

$$\begin{array}{ll} LP & \text{Minimizar} \quad c^T x + g^T \mu \\ & \text{sujeito a} \quad -Hx + D^T \mu = c \\ & \quad \quad \quad Dx = g \end{array} \quad (3.37)$$

e verifica-se o seguinte resultado.

Teorema 3.8 *Considere os problemas QP e LP dados por (3.34) e (3.37), respectivamente. Se LP é inadmissível, então ou QP é inadmissível ou é ilimitado. De outro modo, LP fornece um ponto estacionário para QP que é um seu mínimo global.*

Demonstração:

O programa dual DL do LP (3.37) pode ser escrito da seguinte forma:

$$\begin{aligned}
 DL \quad & \text{Maximizar} \quad c^T y \quad + \quad g^T \gamma \\
 & \text{sujeito a} \quad -Hy \quad + \quad D^T \gamma \quad = \quad c \\
 & \quad \quad \quad Dy \quad \quad \quad = \quad g.
 \end{aligned} \tag{3.38}$$

Como (3.37) e (3.38) têm as mesmas restrições, então

$$\text{LP é inadmissível} \Leftrightarrow \text{DL é inadmissível}$$

Mas se DL é inadmissível, então [10] existe (d_x, d_μ) tal que

$$-Hd_x + D^T d_\mu = 0 \tag{3.39}$$

$$Dd_x = 0 \tag{3.40}$$

$$c^T d_x + g^T d_\mu < 0 \tag{3.41}$$

Se QP é inadmissível o teorema está demonstrado. De outro modo seja $\bar{x}$ uma solução admissível de QP e consideremos a trajectória $\bar{x} + \alpha d_x$, para $\alpha \geq 0$. Esta trajectória é admissível, pois, por (3.40),

$$D(\bar{x} + \alpha d_x) = D\bar{x} + \alpha Dd_x = g$$

Além disso

$$\begin{aligned}
 f(\bar{x} + \alpha d_x) &= c^T(\bar{x} + \alpha d_x) + \frac{1}{2}(\bar{x} + \alpha d_x)^T H(\bar{x} + \alpha d_x) \\
 &= f(\bar{x}) + \alpha [c^T d_x + \bar{x}^T H d_x] + \frac{1}{2}\alpha^2 d_x^T H d_x.
 \end{aligned} \tag{3.42}$$

Mas de (3.39) e (3.40), tem-se

$$d_x^T H d_x = d_\mu^T D d_x = 0 \tag{3.43}$$

Como $D\bar{x} = g$, então

$$\bar{x}^T H d_x = d_\mu^T D \bar{x} = g^T d_\mu. \tag{3.44}$$

Substituindo (3.43) e (3.44) em (3.42), obtém-se

$$f(\bar{x} + \alpha d_x) = f(\bar{x}) + \alpha [c^T d_x + g^T d_\mu]$$

Donde por (3.41), $f(\bar{x} + \alpha d_x) \rightarrow -\infty$ quando $\alpha \rightarrow \infty$.

Finalmente, como os problemas primal LP e dual DL têm as mesmas restrições, se LP é admissível, então não pode ser ilimitado, e por isso tem uma solução óptima. A definição de LP implica que essa solução fornece um ponto estacionário para QP, que é um seu mínimo global. $\diamond$

Para este tipo de problemas, o Algoritmo ASETM deve resolver LP em vez de tratar o problema (3.15). Se LP é inadmissível, então, de acordo com a demonstração do teorema anterior, obtém-se uma direcção d_x ilimitada para o Problema (3.15) e, por isso, uma direcção descendente para o Passo 3. Caso contrário, o iterando corrente é um mínimo global de (3.15). A convergência finita do algoritmo de restrições activas está assim garantida neste caso. É ainda de notar que se a função objectivo é linear, o algoritmo ASETM reduz-se ao método simplex com Base Restrita (Basis Restricted Simplex Method) discutido por vários autores [64, 65, 20, 96].

3.9 Implementação e Casos Degenerados

Da descrição do Algoritmo de Restrições Activas Modificado apresentado na secção 3.6 facilmente se conclui que o código MINOS pode servir para implementar o algoritmo, desde que a escolha da restrição activa a passar a inactiva seja feita de modo a não prejudicar a complementaridade. Por outro lado, o algoritmo de restrições activas necessita de uma solução do GLCP para se iniciar. Essa solução pode ser em geral obtida usando o método enumerativo EASET discutido no capítulo 2.

Se no GLCP (3.4), E é a identidade e $M \in PSD$, então pelo teorema 1.2 essa solução pode ser determinada a partir de um ponto estacionário do programa quadrático QP_{GLCP} associado. Portanto, nesse caso o método enumerativo EASET termina no nó 1 com uma solução do GLCP ou com a indicação que o GLCP e o MPEC são inadmissíveis.

Como foi mencionado na secção 3.6, toda a teoria desenvolvida neste capítulo tem como base a hipótese de não degenerescência. Se essa condição não é verificada pela solução do GLCP obtida pelo algoritmo EASET, então é necessário determinar uma outra solução do GLCP antes de iniciar o Algoritmo das Restrições Activas Modificado de forma a que esse processo possa ser usado. Para explicar melhor a necessidade dessa nova solução do GLCP, apresentamos um pequeno exemplo no qual a hipótese de não degenerescência não é

verificada. Para isso, consideremos o seguinte problema

$$\begin{aligned}
 & \text{Minimizar} \quad f(w_1, w_2, z_1, z_2, y) = -y \\
 & \text{sujeito a} \\
 & \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} y + \begin{bmatrix} 0 \\ -1 \end{bmatrix} \\
 & z_i \geq 0, \quad w_i \geq 0, \quad i \in \{1, 2\} \\
 & 0 \leq y \leq 2 \\
 & z_i w_i = 0, \quad i \in \{1, 2\}
 \end{aligned} \tag{3.45}$$

O algoritmo EASET determina a solução do GLCP

$$\bar{w} = (0, 0)^T, \quad \bar{z} = (1, 0)^T, \quad \bar{y} = 0$$

Portanto no início da algoritmo ASETM, $x^1 = (\bar{w}, \bar{z}, \bar{y})$ e o vector g^1 e a matriz D_1 correspondentes às restrições activas em x^1 são dados por:

$$x^1 = (0, 0, 1, 0, 0), \quad D_1 = \begin{bmatrix} 1 & 0 & 0 & -1 & -1 \\ 0 & 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{e} \quad g^1 = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Portanto $D^1 \in \mathbb{R}^{6 \times 5}$ e a hipótese de não degenerescência não é verificada. Por consequência, o algoritmo ASETM não poderá ser utilizado com essa solução inicial x^1 . Tal como referimos no capítulo 2, se retirarmos uma restrição activa de não negatividade do conjunto de restrições activas a hipótese de não degenerescência passa a ser verificada. Contudo, a escolha arbitrária de uma dessas restrições não é em geral suficiente quando se pretende resolver o MPEC. Com efeito é necessário que para cada par de variáveis complementares pelo menos uma dessas duas variáveis corresponda a uma restrição activa. Se não for possível obter uma solução do GLCP nessas condições, então o algoritmo ASETM não pode ser usado. No exemplo anterior, isso é

conseguido considerando $y = 0$ como inactiva. Tendo em conta que as restrições activas de não negatividade correspondem às variáveis não básicas, então sugerimos o seguinte processo para ser aplicado antes do início do método ASETM.

Modificação 1

Verifique se a solução do GLCP contém algum par de variáveis complementares em que nenhuma dessas variáveis seja não básica. Em caso afirmativo, para cada um desses pares retire uma variável superbásica nula da lista das superbásicas tornando-a não básica. Quanto aos pares de variáveis básicas, efectue operações pivotais de modo a retirar da base uma variável nula de cada um desses pares. Estas operações não podem permitir que se tornem básicas variáveis não básicas cujas complementares são básicas ou superbásicas. Se não for possível efectuar uma dessas operações termine com insucesso. De outro modo, a solução do GLCP assim obtida satisfaz a hipótese de não degenerescência.

Se no exemplo anterior esta modificação for efectuada antes da utilização do algoritmo ASETM, obtém-se uma solução do GLCP com variáveis básicas $y = 0$ e $z_1 = 1$ e as restantes variáveis não básicas. Essa solução satisfaz a hipótese de não degenerescência e está em condições de ser a solução inicial para o método ASETM.

Uma outra dificuldade associada a problemas degenerados é não haver a garantia de que a solução obtida pelo algoritmo de restrições activas modificado ASETM seja um ponto estacionário do MPEC. Para lidar com este tipo de problemas, uma outra modificação terá que ser incorporada no algoritmo. Para ilustrar essa necessidade, consideremos o seguinte problema:

$$\begin{aligned}
 &\text{Minimizar} && f(w_1, w_2, z_1, z_2, y) = w_2 - y \\
 &\text{sujeito a} && \\
 &&& \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} y + \begin{bmatrix} -1 \\ -1 \end{bmatrix} \\
 &&& z_i \geq 0, \quad w_i \geq 0, \quad i \in \{1, 2\} \\
 &&& 0 \leq y \leq 2 \\
 &&& z_i w_i = 0, \quad i \in \{1, 2\}
 \end{aligned} \tag{3.46}$$

Ao ser aplicado o algoritmo ASETM para a obtenção de um ponto estacionário para o problema (3.46) usando como base inicial

$$B = \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix},$$

associada às variáveis w_1 e y , respectivamente, obtém-se a seguinte solução:

EXIT -- OPTIMAL SOLUTION FOUND									
NO. OF ITERATIONS		0		OBJECTIVE VALUE		-1.000000000000E+00			
NUMBER	.COLUMN.	STATE	...ACTIVITY...	.OBJ	GRADIENT.	..LOWER LIMIT.	..UPPER LIMIT.	REDUCED GRADNT	M+J
1	w	1 D BS	.00000	.00000	.00000	.00000	NONE	.00000	4
2	w	2 A LL	.00000	1.00000	.00000	.00000	NONE	.00000	5
3	z	1 N LL	.00000	.00000	.00000	.00000	NONE	-1.00000	6
4	z	2 A LL	.00000	.00000	.00000	.00000	NONE	.00000	7
5	y	1 BS	1.00000	-1.00000	.00000	2.00000	.00000	.00000	8

Figura 3.1: Solução do MPEC sem Modificação 2

A análise da solução apresentada na Figura 3.1, mostra que existe um par de variáveis complementares (z_1, w_1) ambas nulas, sendo z_1 não básica e w_1 básica. Portanto $w_1 \geq 0$ é considerada como inactiva e o processo pára sem ter a garantia de se ter obtido um ponto estacionário do MPEC. Como o multiplicador λ_1^z (gradiente reduzido) associado a $z_1 = 0$ é negativo e $w_1 = 0$ há toda a vantagem de aumentar o valor de z_1 . Para efectuar essa operação, há a necessidade de tentar passar previamente a variável w_1 a não básica. Se tal for possível, então z_1 poderá posteriormente aumentar o seu valor. Essas considerações sugerem a seguinte modificação ao algoritmo ASETM:

Modificação 2

No Passo 1 do algoritmo ASETM, verifique se existem variáveis básicas ou super-básicas nulas complementares de variáveis associadas a

$$\lambda_i^z < 0 \quad \vee \quad \lambda_i^w < 0 \quad \text{para} \quad i \in (L_z \cup L_w) \setminus (L_z \cap L_w)$$

Se existirem, retire as variáveis superbásicas nulas da lista das superbásicas tornando-as não básicas. Quanto às variáveis básicas nulas, efectue operações pivotais para as retirar da base. Estas operações não podem permitir que se tornem básicas variáveis não básicas cujas complementares são básicas ou superbásicas.

Ao resolver o problema (3.46) com o algoritmo ASETM com estas modificações obtém-se a solução da Figura 3.2 cuja função objectivo tem menor valor do que o correspondente à solução apresentada na Figura 3.1.

EXIT -- OPTIMAL SOLUTION FOUND										
NO. OF ITERATIONS				1	OBJECTIVE VALUE		-2.000000000000E+00			
NUMBER	.COLUMN.	STATE	...ACTIVITY...	.OBJ	GRADIENT.	..LOWER LIMIT.	..UPPER LIMIT.	REDUCED GRADNT	M+J	
1	w	1 A	LL	.00000	.00000	.00000	NONE	.00000	4	
2	w	2	LL	.00000	1.00000	.00000	NONE	1.00000	5	
3	z	1	BS	1.00000	.00000	.00000	NONE	.00000	6	
4	z	2	BS	1.00000	.00000	.00000	NONE	.00000	7	
5	y	1	UL	2.00000	-1.00000	.00000	2.00000	-1.00000	8	

Figura 3.2: Solução do MPEC com Modificação 2

Notar que a solução obtida é o mínimo global do MPEC do exemplo anterior, pois os multiplicadores associados às restrições activas são não negativos.

Convém salientar que esta modificação nem sempre pode ser usada, por a retirada de uma variável básica não poder ser feita nas condições definidas pelo processo.

3.10 Experiência Computacional

Nesta secção é apresentada alguma experiência computacional com o Algoritmo de Restrições Activas Modificado discutido na Secção 3.6 na obtenção de um ponto estacionário para o MPEC, assim como a comparação deste algoritmo com alguns métodos existentes na literatura. Estas experiências foram realizadas num Pentium IV 2.4GHz com 256 MB de RAM.

Problemas Teste

O gerador de problemas lineares e quadráticos de dois níveis utilizado nas nossas experiências foi desenvolvido por P. Calamai e L. Vicente [18, 19] e permite construir problemas com soluções óptimas conhecidas. Nesse gerador o problema de dois níveis (BP) é apresentado na seguinte forma

$$\begin{aligned}
& \text{Minimizar}_{x,y} \quad Q(x,y) \\
& \text{sujeito a} \quad x \geq 0 \quad \text{e} \quad y \text{ é solução óptima de:} \\
& \quad \text{Minimizar}_y \quad q(x,y) \\
& \quad \text{sujeito a} \quad A_1 y + A_2 x \leq b
\end{aligned}$$

onde Q e q são funções duas vezes continuamente diferenciáveis num conjunto aberto que contém o conjunto admissível do BP, $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, $A_1 \in \mathbb{R}^{r \times m}$, $A_2 \in \mathbb{R}^{r \times n}$ e $b \in \mathbb{R}^r$.

Na Tabela 3.1 são apresentadas as características dos problemas lineares de dois níveis (LBP) obtidos por este gerador [18] e que foram resolvidos nas nossas experiências.

PROB	n	m	OP.OBJ.	NMG	NML
Lv1	4	2	6.000	1	3
Lv2	10	5	16.000	2	14
Lv3	15	10	30.200	1	511
Lv4	20	7	6.000	1	3
Lv5	20	7	14.000	4	12
Lv6	30	20	7.500	1	7
Lv7	30	20	23.500	8	56
Lv8	30	50	50.500	4	1020

Tabela 3.1: Características dos LBPs

Nesta e nas restantes tabelas desta secção, OP.OBJ. é o valor óptimo da função objectivo, enquanto que NMG e MNL são o número total de mínimos globais e locais do problema, respectivamente.

Com este gerador é possível efectuar transformações aos problemas apresentados na Tabela 3.1, originando novos LBPs que possuem exactamente as mesmas características e soluções que os LBPs iniciais, mas que são menos esparsos. Neste trabalho, só foram efectuadas transformações nos três primeiros problemas, que deram origem a três novos LBPs que irão ser designados com o mesmo nome do problema inicial mas com a adição da letra 'M' (Lv1M, Lv2M, e Lv3M, respectivamente). Para efectuar essas modificações foram utilizados os valores descritos na Tabela 3.2 para os parâmetros discutidos em [18].

Finalmente, neste grupo de problemas teste, foram também gerados cinco problemas quadráticos de dois níveis QBP [19] cujas características estão apresentadas na Tabela 3.3.

Assumimos adicionalmente nesses problemas que a região admissível está restringida ao octante não negativo isto é, que $x \geq 0$ e $y \geq 0$. Essa modificação não traz qualquer alteração

PROB	vx	vy	D
LV1M	(-0.7,0.7,0.1,-0.1)	(0.8,-0.6)	(20,10,20,10,10,20)
LV2M	(-0.8,0,0,0.5,0, -0.3,0,0,-0.1,0.1)	(0.6,0,0.7, 0.4,-0.3)	(50,30,80,60,24,64,78,82, 38,60,90,28,85,70,38)
LV3M	(-0.8,0,0,-0.5,0,0.3, 0,0,0.3,0.1,0,0,0)	(-0.6,0.7,0, 0,-0.4,0,0, 0,0,-0.3)	(50,30,80,60,24,64,78,82, 38,60,90,28,85,70,38,25, 45,36,87,37,75,86,65,24,57)

Tabela 3.2: Características dos LBPs Modificados

PROB	n	m	OP.OBJ.	NMG	NML
NLV1	4	2	0.313	1	3
NLV2	6	4	0.593	1	15
NLV3	10	5	0.790	2	14
NLV4	25	15	1.040	4	28
NLV5	30	20	2.436	32	8160

Tabela 3.3: Características dos QBPs

aos mínimos locais e globais dos problemas teste.

Como segunda classe de problemas-teste considerámos os MPECs introduzidos por Jiang e Ralph em [71]. Esses problemas são definidos da seguinte forma:

$$\begin{aligned}
LCP - QP \quad & \text{Minimizar} \quad \frac{1}{2} \begin{bmatrix} x \\ z \end{bmatrix}^T P \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} c \\ d \end{bmatrix}^T \begin{bmatrix} x \\ z \end{bmatrix} \\
& \text{sujeito a} \quad Gx + Hz + a \leq 0 \\
& \quad \quad \quad w = Nx + Mz + q \\
& \quad \quad \quad z \geq 0, w \geq 0 \\
& \quad \quad \quad w^T z = 0
\end{aligned}$$

onde $x \in \mathbb{R}^n$, $z \in \mathbb{R}^m$, $w \in \mathbb{R}^m$, $P \in \mathbb{R}^{(n+m) \times (n+m)}$, $c \in \mathbb{R}^n$, $d \in \mathbb{R}^m$, $G \in \mathbb{R}^{p \times n}$, $H \in \mathbb{R}^{p \times m}$, $a \in \mathbb{R}^p$, $N \in \mathbb{R}^{m \times n}$, $M \in \mathbb{R}^{m \times m}$ e $q \in \mathbb{R}^m$.

Com o objectivo de gerar este tipo de problemas foi utilizado o gerador QPECgen, implementado em MATLAB e desenvolvido por H. Jiang e D. Ralph [71]. Os problemas teste gerados dependem de parâmetros definidos em [71] e da versão de MATLAB usada. No artigo foram sugeridas as versões 4.0 e 5.0. Como só temos disponível a versão 5.0 para estudantes, foi apenas possível gerar os 4 problemas teste de menor dimensão, introduzidos em [71]. Estes problemas serão designados por M5_. Usou-se também a versão MATLAB 6.1 para gerar

novos problemas teste com os mesmos parâmetros definidos em [71], que se representam por M6_.

Os parâmetros usados são os apresentados na Tabela 8 de [71]. A Tabela 3.4 inclui as características desses problemas teste. Nesta tabela (m, n, p) representa a dimensão dos problemas gerados e FGEN_RALPH é o valor da função objectivo na solução óptima gerada pelo QPECgen.

PROBLEMA	(m,n,p)	M5_STUD	M6.1
		FGEN_RALPH	FGEN_RALPH
1	(50,8,4)	-142.829	-176.645
2	(100,8,4)	-664.389	-663.294
3	(150,8,4)	— — —	-581.974
4	(200,8,4)	— — —	14.217
5	(50,8,4)	-41.876	-108.553
6	(100,8,4)	-599.936	-614.683
7	(150,8,4)	— — —	-562.634
8	(200,8,4)	— — —	126.845

Tabela 3.4: Características dos LCP-QPs

Como terceira classe de problema-teste, considerámos os LCPs NP-difíceis associados a problemas de mochila e discutidos no capítulo 2. Como vimos nesse capítulo, estes LCPs podem ser transformados em MPECs usando as transformações descritas na secção 3.4.

Determinação de um ponto estacionário do MPEC

Uma vez que os MPECs associados aos problemas teste descritos anteriormente podem ser degenerados, torna-se indispensável a incorporação da *Modificação 1* no Algoritmo das Restrições Activas Modificado ASETM. A incorporação da *Modificação 2*, versão ASETM-MOD2, não é indispensável mas pode tornar o processo mais eficiente. Esta conclusão é evidenciada na análise dos resultados das Tabelas 3.5 e 3.6. Nas tabelas deste capítulo, N representa a dimensão do MPEC, NC e Ni são, respectivamente, o número total de pares de variáveis complementares e iterações executadas pelo Algoritmo das Restrições Activas Modificado, T é o tempo total de CPU em segundos necessário para resolver o MPEC, OBJ. é o valor obtido para a função objectivo e tal como anteriormente OP.OBJ. é o valor óptimo conhecido.

As Tabelas 3.5 e 3.6 incluem os resultados das experiências computacionais com o algoritmo

ASETM na resolução de MPECs associados aos problemas de 2 níveis do gerador de Calamai e Vicente [18, 19] e dos LCPs correspondentes a problemas de mochila.

A análise desses resultados mostra que de um modo geral o Algoritmo da Restrições Activas Modificado consegue encontrar um ponto estacionário para todos os problemas num tempo de execução razoável e com poucas iterações. Esse ponto estacionário é nalguns casos o mínimo global do problema de dois níveis. Para os LCPs associados aos problemas de Mochila a versão ASETM-MOD2 consegue por vezes melhorar o valor da função objectivo relativamente à versão ASETM com poucas mais iterações e num tempo de execução da mesma ordem de grandeza.

Na Tabela 3.7 são apresentados os resultados médios, pior e melhor dos resultados incluídos as Tabelas 3.5 e 3.6, onde o parâmetro DIFOBJ é definido por

$$\text{DIFOBJ} = \begin{cases} \frac{\text{OBJ} - \text{OP.OBJ.}}{\text{OP.OBJ.}} & \text{se } \text{OP.OBJ.} \neq 0 \\ \text{OBJ} - \text{OP.OBJ.} & \text{se } \text{OP.OBJ.} = 0 \end{cases}$$

Também nestes resultados é evidenciado uma melhoria clara do valor da função objectivo na versão ASETM-MOD2 em relação a ASETM, sem que haja uma diferença significativa entre os esforços computacionais. Portanto há vantagem em usar ASETM-MOD2 em geral.

Para comparar a eficiência do método de restrições activas modificado em relação a outras técnicas de optimização não linear, considerámos duas formulações PNL1 e PNL2 de programação não linear do MPEC, onde:

- PNL1 - Programação Não Linear (3.12) contendo as restrições lineares do MPEC e as restrições $w_i z_i = 0, i = 1, \dots, n$
- PNL2 - Programação Não Linear com as restrições lineares do MPEC e a restrição não linear $z^T w \leq 0$

As Tabelas 3.8 e 3.9 ilustram a experiência computacional realizada com o código MINOS na resolução dos problemas teste. Nestas tabelas o parâmetro INFEAS significa que o MINOS encontrou uma solução inadmissível, enquanto que INTINF significa que o programa parou porque atingiu o limite máximo de 10000 iterações sem conseguir obter uma solução admissível.

Os resultados apresentados nas Tabelas 3.8 e 3.9 mostram que MINOS não conseguiu resolver os PNLs associados a alguns dos problemas teste, enquanto que o ASETM-MOD2 determinou

				ASETm			ASETm-Mod2			
	PROB	Nc	N	Ni	T	OBJ.	Ni	T	OBJ.	Op.OBJ.
L B P	Lv1	12	12×28	8	0.02	8.000	8	0.03	8.000	6.000
	Lv1M	12	12×28	10	0.00	6.000	10	0.03	6.000	6.000
	Lv2	30	30×70	20	0.00	20.000	20	0.02	20.000	16.000
	Lv2M	30	30×70	21	0.02	16.000	21	0.01	16.000	16.000
	Lv3	55	55×125	35	0.00	40.000	35	0.03	40.000	30.200
	Lv3M	55	55×125	36	0.00	33.200	36	0.03	33.200	30.200
	Lv4	48	48×116	29	0.00	8.000	29	0.02	8.000	6.000
	Lv5	48	48×116	31	0.00	16.000	31	0.00	16.000	14.000
	Lv6	110	110×250	53	0.02	12.000	53	0.02	12.000	7.500
	Lv7	110	110×250	57	0.02	28.000	57	0.05	28.000	23.500
	Lv8	190	190×410	94	0.03	56.000	94	0.05	56.000	50.500
Q B P	Nlv1	8	8×20	6	0.06	0.500	6	0.00	0.500	0.313
	Nlv2	16	16×38	10	0.03	1.000	10	0.00	1.000	0.593
	Nlv3	20	20×50	16	0.02	1.000	16	0.02	1.000	0.790
	Nlv4	60	60×145	50	0.01	1.250	50	0.00	1.250	1.040
	Nlv5	80	80×190	57	0.03	3.250	57	0.02	3.250	2.436
M O C H I L A	1Na1	44	44×110	40	0.02	0.499	41	0.02	0.280	0.000
	1Na2	104	104×260	120	0.05	0.500	121	0.03	0.440	0.000
	1Na3	204	204×510	203	0.05	0.346	203	0.05	0.346	0.000
	1Na4	304	304×760	295	0.08	0.321	295	0.09	0.321	0.000
	1Nb1	44	44×110	12	0.03	0.499	13	0.02	0.310	0.000
	1Nb2	104	104×260	59	0.05	0.353	59	0.02	0.353	0.000
	1Nb3	204	204×510	111	0.06	0.263	111	0.03	0.263	0.000
	1Nb4	304	304×760	168	0.10	0.500	169	0.05	0.200	0.000
	1Nc1	44	44×110	9	0.08	0.497	10	0.00	0.030	0.000
	1Nc2	104	104×260	35	0.01	0.425	35	0.02	0.425	0.000
	1Nc3	204	204×510	63	0.05	0.318	63	0.01	0.318	0.000
	1Nc4	304	304×760	96	0.05	0.500	97	0.03	0.080	0.000
	1IA1	44	44×110	40	0.03	0.499	41	0.00	0.280	0.000
	1IA2	104	104×260	120	0.03	0.500	121	0.02	0.440	0.000
	1IA3	204	204×510	203	0.06	0.346	203	0.05	0.346	0.000
	1IA4	304	304×760	295	0.08	0.321	295	0.08	0.321	0.000

Tabela 3.5: Ponto estacionário usando ASETm e ASETm-Mod2

	PROB	NC	N	ASETM			ASETM-Mod2			Op.OBJ.
				Ni	T	OBJ.	Ni	T	OBJ.	
M O C H I L A	1IB1	44	44×110	12	0.00	0.499	13	0.03	0.310	0.000
	1IB2	104	104×260	59	0.03	0.353	59	0.02	0.353	0.000
	1IB3	204	204×510	111	0.06	0.263	111	0.05	0.263	0.000
	1IB4	304	304×760	168	0.05	0.500	169	0.05	0.200	0.000
	1IC1	44	44×110	9	0.03	0.497	10	0.03	0.030	0.000
	1IC2	104	104×260	35	0.02	0.425	35	0.02	0.425	0.000
	1IC3	204	204×510	63	0.05	0.318	63	0.02	0.318	0.000
	1IC4	304	304×760	96	0.03	0.500	97	0.03	0.080	0.000
	2NA1	44	44×110	64	0.03	12.000	64	0.01	12.000	0.000
	2NA2	104	104×260	146	0.03	6.000	146	0.03	6.000	0.000
	2NA3	204	204×510	307	0.08	8.000	307	0.05	8.000	0.000
	2NA4	304	304×760	436	0.16	19.433	439	0.11	5.000	0.000
	2NB1	44	44×110	22	0.03	16.397	23	0.02	14.000	0.000
	2NB2	104	104×260	88	0.03	8.000	88	0.02	8.000	0.000
	2NB3	204	204×510	178	0.06	12.438	179	0.06	6.000	0.000
	2NB4	304	304×760	302	0.12	22.417	303	0.09	10.000	0.000
	2NC1	44	44×110	17	0.08	11.000	17	0.02	11.000	0.000
	2NC2	104	104×260	56	0.02	4.000	56	0.01	4.000	0.000
	2NC3	204	204×510	103	0.06	5.000	103	0.03	5.000	0.000
	2NC4	304	304×760	146	0.08	8.000	146	0.06	8.000	0.000
	3PA1	44	44×110	25	0.02	0.121	25	0.00	0.121	0.000
	3PA2	100	100×250	61	0.00	0.457	61	0.02	0.457	0.000
	3PA3	204	204×510	143	0.06	0.180	143	0.01	0.180	0.000
	3PA4	304	300×750	163	0.06	0.146	163	0.05	0.146	0.000
	3PB1	44	44×110	27	0.02	0.242	27	0.00	0.242	0.000
	3PB2	104	100×250	69	0.03	0.121	69	0.03	0.121	0.000
	3PB3	204	204×510	177	0.08	0.103	177	0.03	0.103	0.000
	3PB4	304	300×750	251	0.08	0.483	252	0.06	0.483	0.000
	3PC1	44	44×110	31	0.05	0.000	31	0.02	0.000	0.000
	3PC2	104	100×250	72	0.03	0.000	72	0.03	0.000	0.000
	3PC3	204	204×510	161	0.05	0.320	161	0.01	0.320	0.000
	3PC4	304	300×750	237	0.08	0.167	237	0.08	0.167	0.000

Tabela 3.6: Ponto estacionário usando ASETM e ASETM-Mod2 (Cont.)

	ASETm			ASETm-MOD2		
	Ni	T	DIFOBJ	Ni	T	DIFOBJ
PIOR	436	0.16	22.417	439	0.11	14.000
MÉDIA	97	0.04	2.056	98	0.03	1.515
MELHOR	6	0.00	0.000	6	0.00	0.000

Tabela 3.7: Comparação entre os algoritmos ASETm e ASETm-MOD2

sempre pontos estacionários do MPEC. Além disso, o algoritmo ASETm-MOD2 determina em geral soluções com melhor valor óptimo em menos iterações e com menor tempo de execução. Essa conclusão é ainda mais evidente no resumo da experiência apresentado na Tabela 3.10. Na segunda experiência aplicámos o algoritmo ASETm-MOD2 na resolução dos problemas teste LCP-QP. Os resultados são apresentados na Tabela 3.11, onde OP.OBJ. é o valor da função objectivo, FGEN_RALPH, da solução óptima fornecida pelo gerador de QPECgen. Notar que de acordo com [71] não há a garantia das soluções apresentadas serem óptimos globais dos problemas. Os resultados apresentados na Tabela 3.11 mostram que o ASETm-MOD2 conseguiu obter sempre a solução óptima dada pelo gerador em pouco tempo e num número razoável de iterações. Não é possível, no entanto, comparar estes resultados com os obtidos pelo algoritmo PIPA, uma vez que o número de iterações mencionadas em [71] não corresponde ao número de operações pivotais do ASETm-MOD2, mas sim ao número de vezes que uma determinada sequência de passos do algoritmo é efectuada. Por outro lado não é possível contabilizar no algoritmo ASETm-MOD2 o número de operações em vírgula flutuante.

Em [84] Sven Leyffer considera o seguinte MPEC

$$\begin{aligned}
&\text{Minimizar} && x + w \\
&\text{Sujeito a} && -1 \leq x \leq 1 \\
&&& -1 + x + y = 0 \\
&&& 0 \leq y \perp w \geq 0
\end{aligned}$$

A solução óptima deste problema é $(x, y, w)^* = (-1, 2, 0)$. No entanto, em [84] é mostrado que o PIPA converge para um ponto limite $(x^\infty, y^\infty, w^\infty)$ com $-0.4828 \leq x^\infty \leq 0$, $1 \leq y^\infty \leq 1.4828$ e $w^\infty = 0.0$ que satisfaz as hipóteses exigidas em [85] para a convergência do algoritmo PIPA mas que não é um ponto estacionário deste problema. Ao analisar o comportamento do ASETm-MOD2 na resolução deste problema verificamos que o algoritmo obtém um ponto

		ASET-M-Mod2			MINOS/PnL1			MINOS/PnL2			
	PROB	Ni	T	OBJ.	Ni	T	OBJ.	Ni	T	OBJ.	Op.OBJ.
L B P	Lv1	8	0.03	8.000	47	0.10	8.000	12	0.10	8.000	6.000
	Lv1M	10	0.03	6.000	50	0.13	6.000	16	0.13	6.000	6.000
	Lv2	20	0.02	20.000	185	0.23	26.333	22	0.13	17.000	16.000
	Lv2M	21	0.01	16.000	166	0.23	26.333	26	0.11	16.000	16.000
	Lv3	35	0.03	40.000	259	7.28	52.281	38	0.13	37.000	30.200
	Lv3M	36	0.03	33.200	264	0.32	51.766	40	0.07	33.200	30.200
	Lv4	29	0.02	8.000	151	0.13	8.000	24	0.12	8.000	6.000
	Lv5	31	0.00	16.000	185	0.23	17.666	26	0.12	16.000	14.000
	Lv6	53	0.02	12.000	138	0.11	78.833	65	0.09	12.000	7.500
	Lv7	57	0.05	28.000	236	0.11	76.000	81	0.11	28.000	23.500
Lv8	94	0.05	56.000	262	0.24	56.567	246	0.09	52.500	50.500	
Q B P	NLv1	6	0.00	0.500	27	0.11	0.313	29	0.11	0.313	0.313
	NLv2	10	0.00	1.000	95	0.23	0.593	47	0.23	0.753	0.593
	NLv3	16	0.02	1.000	99	0.24	0.790	52	0.24	0.790	0.790
	NLv4	50	0.00	1.250	78	0.13	1.250	89	0.12	1.250	1.040
	NLv5	57	0.02	3.250	94	0.12	3.250	86	0.11	3.250	2.436
M O C H I L A	1NA1	41	0.02	0.280	267	0.11	INFEAS	49	0.12	0.060	0.000
	1NA2	121	0.03	0.440	271	0.22	INFEAS	300	0.31	0.102	0.000
	1NA3	203	0.05	0.346	345	0.19	2.430	300	0.39	0.500	0.000
	1NA4	295	0.09	0.321	381	0.19	6.050	491	0.78	0.500	0.000
	1NB1	13	0.02	0.310	261	0.16	INFEAS	53	0.09	0.260	0.000
	1NB2	59	0.02	0.353	343	0.22	1.550	148	0.10	0.353	0.000
	1NB3	111	0.03	0.263	254	0.09	0.263	161	0.08	0.263	0.000
	1NB4	169	0.05	0.200	388	0.17	0.320	249	0.14	0.320	0.000
	1Nc1	10	0.00	0.030	288	0.21	0.980	70	0.11	0.020	0.000
	1Nc2	35	0.02	0.425	265	0.11	INFEAS	167	0.11	0.425	0.000
	1Nc3	63	0.01	0.318	264	0.11	INFEAS	245	0.08	INFEAS	0.000
	1Nc4	97	0.03	0.080	547	0.31	0.720	220	0.15	0.080	0.000
	1IA1	41	0.00	0.280	283	0.18	1.918	48	0.11	0.060	0.000
	1IA2	121	0.02	0.440	271	0.19	INFEAS	278	0.32	0.102	0.000
	1IA3	203	0.05	0.346	345	0.20	2.430	300	0.43	0.500	0.000
	1IA4	295	0.08	0.321	376	0.21	16.493	1360	3.09	INTINF	0.000

Tabela 3.8: Ponto estacionário usando MINOS e ASETM-Mod2

		ASETm-Mod2			MINOS/PnL1			MINOS/PnL2			Op.OBJ.
PROB		Ni	T	OBJ.	Ni	T	OBJ.	Ni	T	OBJ.	
M O C H I L A	1IB1	13	0.03	0.310	261	0.13	INFEAS	48	0.09	0.260	0.000
	1IB2	59	0.02	0.353	343	0.22	1.550	149	0.08	0.353	0.000
	1IB3	111	0.03	0.263	254	0.09	0.263	161	0.06	0.263	0.000
	1IB4	169	0.05	0.200	387	0.25	0.320	246	0.13	0.320	0.000
	1IC1	10	0.03	0.030	286	0.22	0.980	68	0.11	0.020	0.000
	1IC2	35	0.02	0.425	268	0.09	INFEAS	168	0.11	0.425	0.000
	1IC3	63	0.02	0.318	264	0.15	INFEAS	245	0.10	INFEAS	0.000
	1IC4	97	0.03	0.080	553	0.29	3.910	259	0.15	0.080	0.000
	2NA1	64	0.01	12.000	267	0.12	INFEAS	161	0.09	12.000	0.000
	2NA2	146	0.03	6.000	144	0.09	6.000	161	0.09	6.000	0.000
	2NA3	307	0.05	8.000	606	0.19	21.000	500	0.15	24.000	0.000
	2NA4	439	0.11	5.000	834	0.42	34.000	702	0.25	5.000	0.000
	2NB1	23	0.02	14.000	244	0.13	INFEAS	99	0.13	14.000	0.000
	2NB2	88	0.02	8.000	397	0.44	41.00	275	0.11	INFEAS	0.000
	2NB3	179	0.06	6.000	618	0.31	9.000	394	0.17	6.000	0.000
	2NB4	303	0.09	10.000	631	0.28	8.000	486	0.23	10.000	0.000
	2NC1	17	0.02	11.000	261	0.11	INFEAS	188	0.11	11.000	0.000
	2NC2	56	0.01	4.000	312	0.12	INFEAS	158	0.19	2.000	0.000
	2NC3	103	0.03	5.000	558	0.31	64.000	260	0.06	INFEAS	0.000
	2NC4	146	0.06	8.000	703	0.37	10.000	413	0.23	14.000	0.000
	3PA1	25	0.00	0.121	210	0.23	0.620	33	0.13	0.080	0.000
	3PA2	61	0.02	0.457	226	0.23	2.250	199	0.31	0.320	0.000
	3PA3	143	0.01	0.180	359	0.17	2.205	266	0.17	0.360	0.000
	3PA4	163	0.05	0.146	383	0.20	1.9545	212	0.15	0.146	0.000
	3PB1	27	0.00	0.242	273	0.24	1.080	38	0.13	0.242	0.000
	3PB2	69	0.03	0.121	298	0.22	3.280	50	0.10	0.121	0.000
	3PB3	177	0.03	0.103	404	0.23	0.796	421	0.73	0.103	0.000
	3PB4	252	0.06	0.483	785	0.42	1.338	269	0.14	0.500	0.000
	3PC1	31	0.02	0.000	192	0.23	0.000	52	0.11	0.000	0.000
	3PC2	72	0.03	0.000	326	0.22	0.750	70	0.12	0.000	0.000
	3PC3	161	0.01	0.320	340	0.21	1.313	204	0.09	0.242	0.000
	3PC4	237	0.08	0.167	766	0.30	1.592	249	0.14	0.179	0.000

Tabela 3.9: Ponto estacionário usando MINOS e ASETm-Mod2 (Cont.)

	ASETM-MOD2			MINOS/PNL1			MINOS/PNL2		
	Ni	T	DIFOBJ	Ni	T	DIFOBJ	Ni	T	DIFOBJ
Pior	439	0.11	14.000	834	7.29	64.000	702	0.78	24.000
MÉDIA	98	0.03	1.515	329	0.36	5.237	175	0.17	1.8736
MELHOR	6	0.00	0.000	27	0.09	0.000	12	0.06	0.000
FRACASSOS	0			13			5		

Tabela 3.10: Comparação entre MINOS e ASETM-MOD2

				ASETM-MOD2			Op.OBJ.
				Ni	T	OBJ	
LCP-QP	PROB	Nc	N				
	M5_1	50	54×108	203	0.08	-142.829	-142.829
	M5_2	100	104×208	368	0.33	-664.389	-664.389
	M5_5	50	54×108	168	0.03	-41.876	-41.876
	M5_6	100	104×208	384	0.33	-599.936	-599.936
	M6_1	50	54×108	177	0.05	-176.645	-176.645
	M6_2	100	104×208	385	0.36	-663.294	-663.294
	M6_3	150	154×308	610	1.19	-581.974	-581.974
	M6_4	200	204×408	3219	7.59	14.217	14.217
	M6_5	50	54×108	186	0.08	-108.553	-108.553
	M6_6	100	104×208	345	0.31	-641.683	-641.683
	M6_7	150	154×308	687	1.44	-562.634	-562.634
	M6_8	200	204×408	5143	11.09	126.845	126.845

Tabela 3.11: Ponto estacionário para os problemas LCP-QPs usando ASETM-MOD2

estacionário, que é um mínimo global do problema, num reduzido esforço computacional.

Como referimos anteriormente, Fukushima, Luo e Pang introduziram em [46] um Algoritmo Sequencial de Programação Quadrática (SQP) para o cálculo de pontos estacionários de um programa matemático com restrições lineares de complementaridade

$$\begin{aligned}
 & \text{Minimizar} && f(x, y) \\
 & \text{sujeito a} && w = Nx + My + q \\
 & && Ax \leq b \\
 & && 0 \leq w \perp y \geq 0
 \end{aligned}$$

onde $f : \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ é uma função continuamente diferenciável, $N \in \mathbb{R}^{m \times n}$, $M \in \mathbb{R}^{m \times m}$ é uma matriz P_0 , $A \in \mathbb{R}^{p \times n}$, $q \in \mathbb{R}^m$, $b \in \mathbb{R}^p$ e $(x, y, w) \in \mathbb{R}^{n+2m}$. A eficiência do processo foi testada através da realização de uma pequena experiência computacional com alguns problemas testes definidos em [34] e usados na experiência computacional de métodos alternativos. Na Tabela 3.12 são apresentados os resultados obtidos usando o algoritmo ASETM-MOD2,

o algoritmo SQP descrito em [46] e o método As (Ponto Estacionário Aproximado - método diferenciável) introduzido em [34] na resolução desses problemas.

P.FUK.	P.FAC.	(m,n,p)	As		SQP		ASETM-MOD2	
			NI	OBJ	NI	OBJ	NI	OBJ
1	6	(1,1,2)	5	-3266.67	22(6)	-3266.67	2	-3266.67
2	9	(2,2,4)	7	0.36E-26	48(6)	0.36E-26	4	0.25E-29
3	10	(4,4,9)	51	-6600.00	31	-6600.00	15	-6600.00

Tabela 3.12: Comparação dos métodos ASETM-MOD2, SQP e As

Nesta tabela P.FUK. e P.FAC. representam a numeração que o mesmo problema teste possui em [46] e [34], respectivamente, enquanto que (m, n, p) representa a dimensão do problema. Na Tabela 3.12 são apresentados dois valores para o número de iterações no algoritmo SQP. O primeiro é obtido usando os mesmos valores para os parâmetros do programa SQP nos três problemas teste, enquanto que o valor dentro de parêntesis é obtido com valores especificamente escolhidos para o problema em questão.

Os resultados apresentados na Tabela 3.12 mostram que os três métodos conseguiram obter uma solução com o mesmo valor da função objectivo, que aliás é o mínimo global do problema [34]. No entanto, o ASETM-MOD2 necessitou de um menor número de iterações para as obter. Como a dimensão dos problemas é pequena, o tempo de execução é pouco significativo e por isso não é apresentado.

Como conclusão final das experiências computacionais efectuadas, podemos afirmar que o método de restrições activas modificado é um processo robusto e eficiente para a determinação de um ponto estacionário de um MPEC. Apesar de ser difícil estabelecer conclusões em relação a outras técnicas alternativas por falta de informação disponível, alguns resultados indicam que o algoritmo é competitivo e até em muitos casos mais eficiente do que outros métodos locais que têm sido propostos para a determinação de um ponto estacionário do MPEC.

Capítulo 4

ALGORITMOS PARA OPTIMIZAÇÃO GLOBAL DE PROBLEMAS DE PROGRAMAÇÃO MATEMÁTICA COM RESTRIÇÕES DE COMPLEMENTARIDADE

Neste capítulo são discutidos algoritmos para a determinação de um mínimo global para o MPEC. O primeiro processo é um método de ramificação e limitação, do inglês Branch-and-Bound method. O algoritmo incorpora cortes disjuntivos para a determinação de limites inferiores e o Método de Restrições Activas Modificado ASETM para a obtenção de limites superiores. O Método Sequencial Complementar introduzido em [64] é também analisado neste capítulo. Experiência computacional com estes algoritmos na resolução de MPECs associados com Problemas de Dois Níveis e LCPs NP-difíceis é também apresentada neste capítulo e atesta a eficiência dos processos para os fins em vista.

4.1 Método de Ramificação e Limitação

Nesta secção iremos descrever um algoritmo de Ramificação e Limitação, do inglês Branch-and-Bound Algorithm, para a determinação de um mínimo global para o MPEC da forma

$$\begin{aligned}
 & \text{Minimizar} && f(z, y) \\
 & \text{sujeito a} && Ew = Mz + Ny + q \\
 & && z \geq 0, \quad w \geq 0 \\
 & && y \in K_y = \{y \in \mathbb{R}^m : Ay = b, y \geq 0\} \\
 & && z^T w = 0
 \end{aligned}$$

onde f é uma função convexa no conjunto definido pelas restrições lineares do problema, $M, E \in \mathbb{R}^{l \times n}$, $N \in \mathbb{R}^{l \times m}$, $w, z \in \mathbb{R}^n$, $y \in \mathbb{R}^m$ e $A \in \mathbb{R}^{p \times m}$, $b \in \mathbb{R}^p$ e $q \in \mathbb{R}^l$. Como as desigualdades em K_y podem ser transformadas em igualdades por introdução de variáveis de folga, a forma apresentada para definir K_y é a mais geral possível. Como vimos anteriormente, em muitas aplicações do MPEC, $l = n$, E é a matriz identidade e M é uma matriz *PSD*.

Tal como em programação inteira, o algoritmo de ramificação e limitação determina um mínimo global para o MPEC explorando uma árvore binária que vai sendo gerada ao longo do processo segundo um esquema baseado na dicotomia $z_i = 0$ ou $w_i = 0$ presente numa solução complementar. Assim se k é um nó dado, então podemos gerar dois nós filhos da seguinte forma

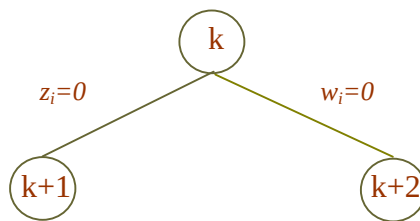


Figura 4.1: Árvore binária

Por isso em cada nó k podemos considerar os seguintes conjuntos

$$J_k = \{i : w_i \text{ fixa em zero no nó } k\}$$

$$L_k = \{i : z_i \text{ fixa em zero no nó } k\}$$

e o programa

$$\begin{aligned}
MPEC_{JL} \quad & \text{Minimizar} \quad f(z, y) \\
& \text{sujeito a} \quad Ew = Mz + Ny + q \\
& z \geq 0, \quad w \geq 0 \\
& y \in K_y = \{y \in \mathbb{R}^m : Ay = b, y \geq 0\} \\
& z^T w = 0 \\
& w_i = 0, \quad i \in J \\
& z_i = 0, \quad i \in L
\end{aligned} \tag{4.1}$$

com $J_k = J$ e $L_k = L$. À semelhança dos algoritmos de ramificação e limitação para programação inteira, é então possível determinar um limite inferior $LB(k)$ para o valor do mínimo global do MPEC. A relaxação convexa (linear se f é uma função linear) é usada para esse fim e consiste em omitir a restrição de complementaridade $z^T w = 0$ no $MPEC_{JL}$. Como veremos mais adiante, cortes disjuntivos [7, 110] são também úteis para esse efeito.

O algoritmo de Restrições Activas Modificado ASETM pode ser utilizado para a determinação de limites superiores para o valor óptimo do MPEC. Com efeito, se $(\bar{z}, \bar{y})$ é um ponto estacionário do $MPEC_{J_k L_k}$ associado ao nó k , então $UB = f(\bar{z}, \bar{y})$ é um limite superior pretendido.

O algoritmo ASETM é também usado no processo de ramificação empregue no algoritmo para a geração da árvore. Suponhamos que os nós da árvore gerada pelo algoritmo constituem os elementos da lista $\mathcal{L}$ e seja $|\mathcal{L}|$ o seu número de elementos. Consideremos que num nó k o algoritmo ASETM determinou um ponto estacionário $(\bar{z}, \bar{w}, \bar{y})$ do $MPEC_{J_k L_k}$. Se λ_i^w, λ_i^z são os multiplicadores de Lagrange associados às restrições $w_i \geq 0$ e $z_i \geq 0$, então há dois casos possíveis:

1. Se

$$\forall_{i \notin J_k} \quad \lambda_i^w \geq 0 \quad \text{e} \quad \forall_{i \notin L_k} \quad \lambda_i^z \geq 0$$

então $(\bar{z}, \bar{w}, \bar{y})$ é um mínimo global de $MPEC_{J_k L_k}$.

2. Se

$$\exists_{i \notin J_k} \quad \lambda_i^w < 0 \quad \text{ou} \quad \exists_{i \notin L_k} \quad \lambda_i^z < 0$$

então sejam $\lambda_{i_1}^w < 0, \dots, \lambda_{i_r}^w < 0$ e $\lambda_{i_1}^z < 0, \dots, \lambda_{i_s}^z < 0$, com $r + s \geq 1$. É então possível gerar um ramo da árvore binária baseado nesses multiplicadores. Para simplificar, consideremos o caso em que $r = s = 2$. Então podemos construir os seguintes ramos a partir do nó k .

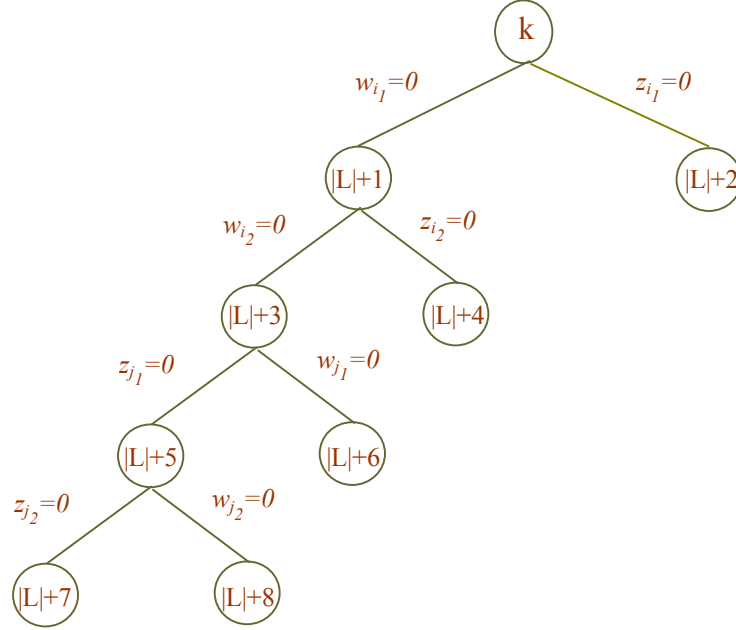


Figura 4.2: Esquema de Ramificação

Facilmente se conclui que esse ramo da árvore tem exactamente $(r + s + 1)$ nós abertos e que $(\bar{z}, \bar{w}, \bar{y})$ é um mínimo global para o MPEC_{JL} associado ao nó $|\mathcal{L}| + 2(r + s) - 1$. Portanto esse nó não necessita de ser considerado em futuros desenvolvimentos, pelo que apenas $(r + s)$ nós devem ser acrescentados à lista. É também importante notar que o ponto estacionário $(\bar{z}, \bar{w}, \bar{y})$ pode servir como solução inicial para cada um dos MPEC_{JL} associados aos $(r + s)$ novos nós que entraram na lista $\mathcal{L}$. Desta forma podemos associar a cada um desses nós t o número real $\text{VAL}(t) = f(\bar{z}, \bar{y})$.

Uma vez explicado o processo de ramificação podemos apresentar os passos do Algoritmo.

ALGORITMO DE RAMIFICAÇÃO E LIMITAÇÃO

Passo 0 Seja $\mathcal{L} = \{1\}$ a lista inicial de nós em abertos, $\text{UB} = +\infty$ e $\text{VAL}(1) = +\infty$.

Passo 1 *Optimalidade* - Se $\mathcal{L} = \emptyset$, termine o algoritmo. Se $\text{UB} = +\infty$, então o MPEC

não tem solução ótima (o GLCP não tem solução ou o MPEC é ilimitado). De outro modo, a solução $(\bar{z}, \bar{w}, \bar{y})$ associada a $UB < +\infty$ é o mínimo global do MPEC e o valor ótimo é UB .

Passo 2 *Escolha do nó* - Se $\mathcal{L} \neq \emptyset$, escolha um nó $k \in \mathcal{L}$ com conjuntos J_k e L_k de variáveis w_i e z_i fixas em zero e solução inicial $x^k = (z^k, w^k, y^k)$ associada a esse nó com valor $VAL(k)$. Faça $\mathcal{L} = \mathcal{L} \setminus \{k\}$.

Passo 3 *Cálculo do Limite Inferior* - Determine um limite inferior $LB(k)$ no nó k , usando como solução inicial x^k . Se $LB(k) \geq UB$ vá para o Passo 1.

Passo 4 *Cálculo de um Limite Superior* - Considere o $MPEC_{J_k L_k}$ da forma (4.1), com $J_k = J$ e $L_k = L$ e a solução inicial x^k associada a esse nó k . Com início em x^k , determine uma solução admissível do $MPEC_{J_k L_k}$ (isto é, do GLCP associado). Se essa solução não existe, vá para Passo 1. De outro modo, determine um ponto estacionário $x^k = (z^k, w^k, y^k)$ do $MPEC_{J_k L_k}$ usando o algoritmo ASETM e faça

$$UB = \min \left\{ UB, f(z^k, y^k) \right\}$$

Actualize a solução incumbente (z^k, w^k, y^k) , se UB reduziu de valor.

Passo 5 Sejam λ_i^w, λ_j^z os multiplicadores de Lagrange associados às restrições $w_i \geq 0$ e $z_j \geq 0$ no ponto estacionário $x^k = (z^k, w^k, y^k)$ calculado no Passo 4. Se $\lambda_i^w \geq 0$ para todo $i \notin J_k$ e $\lambda_j^z \geq 0$ para todo $j \notin L_k$, vá para Passo 1. De outro modo sejam $i = i_1, \dots, i_r \notin J_k$ e $j = j_1, \dots, j_s \notin L_k$ tais que $\lambda_i^w < 0$ e $\lambda_j^z < 0$, com $r + s \geq 1$. Acrescente os $(r + s)$ nós à lista $\mathcal{L}$ de acordo com o processo de ramificação explicado anteriormente. Para cada um desses nós, considere $x^k = (z^k, w^k, y^k)$ como solução inicial e faça $VAL(k) = f(z^k, y^k)$. Vá para Passo 2.

Da descrição deste algoritmo, chega-se à conclusão que em cada nó há necessidade de resolver um GLCP, que constitui as restrições do $MPEC_{JL}$ correspondente. Esse assunto será discutido na secção 4.3. Por outro lado, o processo de cálculo de limites inferiores é baseado na resolução do programa convexo relaxado que se obtém do $MPEC_{JL}$ por supressão da restrição de complementaridade $z^T w = 0$ e em cortes disjuntivos. Esse assunto será discutido na secção 4.2.

O algoritmo de ramificação e limitação só será eficiente se a lista $\mathcal{L}$ de nós em aberto não for excessivamente grande. A determinação de bons limites inferiores e superiores é fundamental neste aspecto. Algumas regras heurísticas simples podem também ser incorporadas no algoritmo e têm normalmente como consequência o controle do tamanho da lista $\mathcal{L}$.

- *Ordem dos Índices das Variáveis no Processo de Ramificação* (Passo 5)

Na construção do ramo da árvore associado ao processo de ramificação, os pares de variáveis complementares (w_i, z_i) são colocados por ordem crescente dos valores dos multiplicadores de Lagrange λ_i^w e λ_j^z negativos.

- *Escolha do nó* (Passo 2)

Escolher o nó k para o qual $\text{VAL}(k)$ é mínimo, onde $\text{VAL}(k)$ é o valor da função objectivo da solução básica correspondente ao nó k . Em caso de empate, escolher o nó do nível mais elevado.

4.2 Determinação de Limites Inferiores

Como afirmámos anteriormente, o processo de cálculo de limites inferiores baseia-se no tratamento do programa convexo relaxado RMPEC_{JL} obtido do MPEC_{JL} correspondente a cada nó por supressão da condição de complementaridade $z^T w = 0$. É importante que esses limites inferiores $\text{LB}(k)$ sejam boas aproximações da solução óptima do MPEC. Com efeito se tal acontecer, e se se conhecer um limite superior UB que também seja muito próximo desse valor óptimo, então será possível que num dado nó k se verifique $\text{LB}(k) \geq \text{UB}$, o que de acordo com o Passo 3 do algoritmo significa não ser necessário investigar a árvore a partir desse nó. Isso pode tornar a pesquisa da árvore muito mais reduzida, permitindo assim terminar o algoritmo num esforço computacional razoável.

A experiência computacional com MPECs [64, 65] tem mostrado que a simples determinação da solução óptima do programa relaxado em cada nó não conduz em geral a bons limites inferiores. Cortes disjuntivos têm vindo a ser sugeridos em optimização global com esse objectivo [110] e são discutidos nesta secção.

Consideremos a solução óptima $(\bar{z}, \bar{w}, \bar{y})$ do RMPEC_{JL} associado a um determinado nó da árvore. Se essa solução é complementar, então fornece um novo limite superior $\text{UB} = f(\bar{z}, \bar{y})$

e não há necessidade de continuar a pesquisa a partir desse nó. Caso contrário, existe pelo menos um par de variáveis complementares positivas. Para a explicação da geração dos cortes disjuntivos associados a essa solução $(\bar{z}, \bar{w}, \bar{y})$, iremos sem perda de generalidade considerar as restrições do programa relaxado RMPEC_{JL} na seguinte forma

$$\begin{aligned} Ew &= q + Mz + Ny \\ Ay &= b \\ w &\geq 0, \quad z \geq 0, \quad \alpha \leq y \leq \beta \\ z_i &= 0, \quad i \in L \\ w_i &= 0, \quad i \in J \end{aligned}$$

Notar que nesta formulação um dos limites α_i ou β_i pode ser infinito, mas não ambos. Podemos então escrever estas restrições na forma

$$\begin{aligned} Cx &= g \\ \alpha_i &\leq x_i \leq \beta_i \end{aligned}$$

com

$$g = \begin{bmatrix} q \\ b \end{bmatrix}, \quad C = \begin{bmatrix} -M & I & -N \\ 0 & 0 & A \end{bmatrix}, \quad x = \begin{bmatrix} z \\ w \\ y \end{bmatrix}$$

Iremos primeiramente considerar o caso da solução ótima do programa relaxado ser básica, para depois tratarmos o caso em que a solução ótima não satisfaz essa propriedade.

(I) Cortes Disjuntivos para Soluções Básicas

Suponhamos que $\bar{x} = (\bar{z}, \bar{w}, \bar{y})$ é uma solução básica não complementar. Então as variáveis complementares (z_k, w_k) positivas têm de ser ambas básicas. Sejam r e s as linhas de C correspondentes a essas variáveis e (L, U, J) a partição associada a essa solução básica, onde L e U são os índices das variáveis não básicas nos limites inferiores e superiores respectivamente e J o conjunto das variáveis básicas. Então podemos escrever

$$\begin{aligned}
w_k &= \bar{g}_s - \sum_{j \in L} \bar{c}_{sj} x_j - \sum_{j \in U} \bar{c}_{sj} x_j \\
z_k &= \bar{g}_r - \sum_{j \in L} \bar{c}_{rj} x_j - \sum_{j \in U} \bar{c}_{rj} x_j
\end{aligned}$$

onde $\bar{g}_i$ e $\bar{c}_{ij}$ são os elementos do quadro simplex associado à solução básica. Pela definição de corte disjuntivo [110], terá de acontecer uma das duas alternativas $w_k \leq 0$ ou $z_k \leq 0$. Então

$$\begin{aligned}
&\sum_{j \in L} \bar{c}_{sj} x_j + \sum_{j \in U} \bar{c}_{sj} x_j \geq \bar{g}_s \\
&\vee \\
&\sum_{j \in L} \bar{c}_{rj} x_j + \sum_{j \in U} \bar{c}_{rj} x_j \geq \bar{g}_r
\end{aligned} \tag{4.2}$$

Mas (4.2) pode ser escrito na forma

$$\begin{aligned}
&\sum_{j \in L} \bar{c}_{sj} (x_j - l_j) - \sum_{j \in U} \bar{c}_{sj} (-x_j + u_j) \geq \tilde{g}_s \\
&\vee \\
&\sum_{j \in L} \bar{c}_{rj} (x_j - l_j) - \sum_{j \in U} \bar{c}_{rj} (-x_j + u_j) \geq \tilde{g}_r
\end{aligned}$$

onde

$$\tilde{g}_s = \bar{g}_s - \sum_{j \in L} \bar{c}_{sj} l_j - \sum_{j \in U} \bar{c}_{sj} u_j \quad \text{e} \quad \tilde{g}_r = \bar{g}_r - \sum_{j \in L} \bar{c}_{rj} l_j - \sum_{j \in U} \bar{c}_{rj} u_j$$

que, por hipótese, são ambos positivos. Dividindo cada expressão por $\tilde{g}_s$ e $\tilde{g}_r$ respectivamente, vem

$$\begin{aligned}
&\sum_{j \in L} \frac{\bar{c}_{sj}}{\tilde{g}_s} (x_j - l_j) - \sum_{j \in U} \frac{\bar{c}_{sj}}{\tilde{g}_s} (-x_j + u_j) \geq 1 \\
&\vee \\
&\sum_{j \in L} \frac{\bar{c}_{rj}}{\tilde{g}_r} (x_j - l_j) - \sum_{j \in U} \frac{\bar{c}_{rj}}{\tilde{g}_r} (-x_j + u_j) \geq 1
\end{aligned}$$

Mas

$$x_j - l_j \geq 0 \quad \text{e} \quad -x_j + u_j \geq 0$$

para todo $j \in L \cup U$. Se agora escolhermos

$$d_j = \begin{cases} \max \left\{ \frac{\bar{c}_{sj}}{\tilde{g}_s}, \frac{\bar{c}_{rj}}{\tilde{g}_r} \right\}, & j \in L \\ \min \left\{ \frac{\bar{c}_{sj}}{\tilde{g}_s}, \frac{\bar{c}_{rj}}{\tilde{g}_r} \right\}, & j \in U \end{cases}$$

então obtemos o seguinte Corte Disjuntivo associado à solução básica definida pela partição (L, U, J) :

$$\sum_{j \in L} d_j(x_j - l_j) + \sum_{j \in U} (-d_j)(-x_j + u_j) \geq 1.$$

ou seja

$$\sum_{j \in L \cup U} d_j x_j \geq \lambda$$

onde $\lambda = 1 + \sum_{j \in L} d_j l_j + \sum_{j \in U} d_j u_j$.

(II) Cortes Disjuntivos para Soluções Não Básicas

De modo análogo ao caso anterior, suponhamos que existe pelo menos um par de variáveis complementares positivas. Como a solução não é básica, não temos a garantia de que ambas as variáveis desse par sejam básicas. Assim, para esse par pode acontecer um de três cenários:

- As variáveis são ambas básicas.
- Uma variável é básica e a outra é superbásica.
- As variáveis são ambas superbásicas.

Seja (z_k, w_k) o par das variáveis complementares positivas. Suponhamos ainda que essas duas variáveis são básicas nas linhas r e s associadas à solução admissível. Deste modo, tem-se

$$w_k = \bar{g}_s - \sum_{j \in L \cup U} \bar{c}_{sj} x_j - \sum_{j \in S} \bar{c}_{sj} x_j \quad \wedge \quad z_k = \bar{g}_r - \sum_{j \in L \cup U} \bar{c}_{rj} x_j - \sum_{j \in S} \bar{c}_{rj} x_j$$

onde S , L e U são os conjuntos dos índices das variáveis superbásicas e não básicas nos limites inferiores e superiores respectivamente. Tal como anteriormente, o corte disjuntivo é definido por

$$w_k \leq 0 \quad \vee \quad z_k \leq 0$$

ou seja, tem-se

$$\sum_{j \in N} \bar{c}_{sj} x_j + \sum_{j \in S} \bar{c}_{sj} x_j \geq \bar{g}_r \quad \vee \quad \sum_{j \in N} \bar{c}_{rj} x_j + \sum_{j \in S} \bar{c}_{rj} x_j \geq \bar{g}_s$$

Se $\bar{x}_j$, $j \in S$ são os valores das variáveis superbásicas, então

$$\sum_{j \in L} \bar{c}_{sj}(x_j - l_j) - \sum_{j \in U} \bar{c}_{sj}(-x_j + u_j) + \sum_{j \in S} \bar{c}_{sj}(x_j - \bar{x}_j) \geq \tilde{g}_s \quad (4.3)$$

$\vee$

$$\sum_{j \in L} \bar{c}_{rj}(x_j - l_j) - \sum_{j \in U} \bar{c}_{rj}(-x_j + u_j) + \sum_{j \in S} \bar{c}_{rj}(x_j - \bar{x}_j) \geq \tilde{g}_r \quad (4.4)$$

onde

$$\tilde{g}_s = \bar{g}_s - \sum_{j \in L} \bar{c}_{sj} l_j - \sum_{j \in U} \bar{c}_{sj} u_j - \sum_{j \in S} \bar{c}_{sj} \bar{x}_j \quad \text{e} \quad \tilde{g}_r = \bar{g}_r - \sum_{j \in L} \bar{c}_{rj} l_j - \sum_{j \in U} \bar{c}_{rj} u_j - \sum_{j \in S} \bar{c}_{rj} \bar{x}_j$$

que, por hipótese, são ambos positivos. Se $S = \emptyset$, então o corte disjuntivo pode ser gerado como no caso das soluções básicas. De outro modo procura-se retirar índices de S de forma a torná-lo vazio. Para isso, é suficiente substituir em (4.3) e (4.4) cada parcela associada a $j \in S$ por uma das expressões:

$$\bar{c}_{ij}(x_j - \bar{x}_j) = \bar{c}_{ij}(x_j - l_j) - \bar{c}_{ij}(\bar{x}_j - l_j) \quad (4.5)$$

$$\bar{c}_{ij}(x_j - \bar{x}_j) = \bar{c}_{ij}(x_j - u_j) - \bar{c}_{ij}(\bar{x}_j - u_j) \quad (4.6)$$

com $i = r, s$. Assim, se usar a expressão (4.5) tem-se,

$$\begin{aligned} S &= S \setminus \{j\}, \quad L = L \cup \{j\} \\ \tilde{g}_s &= \bar{g}_s + \bar{c}_{sj}(\bar{x}_j - l_j) \\ \tilde{g}_r &= \bar{g}_r + \bar{c}_{rj}(\bar{x}_j - l_j) \end{aligned} \quad (4.7)$$

Por outro lado (4.6) conduz às seguintes transformações

$$\begin{aligned} S &= S \setminus \{j\}, \quad U = U \cup \{j\} \\ \tilde{g}_s &= \bar{g}_s + \bar{c}_{sj}(\bar{x}_j - u_j) \\ \tilde{g}_r &= \bar{g}_r + \bar{c}_{rj}(\bar{x}_j - u_j) \end{aligned} \quad (4.8)$$

Como vimos anteriormente, a geração dos cortes disjuntivos necessita que após essas transformações $\tilde{g}_s$ e $\tilde{g}_r$ sejam ambas positivas. Para isso as alterações (4.7) e (4.8) devem ser escolhidas de forma a que essa condição seja verdadeira. Três casos são possíveis e são discutidos a seguir.

Caso 1 - $\bar{c}_{sj} \geq 0$ e $\bar{c}_{rj} \geq 0$

Se $l_j > -\infty$, então actualize S , L , $\tilde{g}_s$ e $\tilde{g}_r$ através de (4.7). De outro modo ($l_j = -\infty$), use (4.8) para actualizar S , U , $\tilde{g}_s$ e $\tilde{g}_r$.

Caso 2 - $\bar{c}_{sj} \leq 0$ e $\bar{c}_{rj} \leq 0$

Se $u_j < +\infty$, actualize S , U , $\tilde{g}_s$ e $\tilde{g}_r$ de acordo com (4.8). De outro modo ($u_j = +\infty$) actualize S , L , $\tilde{g}_s$ e $\tilde{g}_r$ de acordo com (4.7).

Caso 3 - $\bar{c}_{sj} > 0$ e $\bar{c}_{rj} < 0$ (ou vice-versa)

Nesse caso considere-se

$$\alpha = \min \{ \bar{c}_{sj}(u_j - \bar{x}_j), \bar{c}_{rj}(l_j - \bar{x}_j) \}$$

e há dois casos possíveis:

- i Se $\alpha = \bar{c}_{sj}(u_j - \bar{x}_j)$, actualize S , U , $\tilde{g}_s$ e $\tilde{g}_r$ de acordo com (4.8).
- ii Se $\alpha = \bar{c}_{rj}(l_j - \bar{x}_j)$, actualize S , L , $\tilde{g}_s$ e $\tilde{g}_r$ de acordo com (4.7).

Convém salientar que, nos casos 1 e 2 com limites finitos tem-se a certeza de que na actualização dos valores de $\tilde{g}_s$ e $\tilde{g}_r$ não existe decréscimo nos seus valores. No entanto, no final do processo, não se tem a garantia de que os valores de $\tilde{g}_s$ e $\tilde{g}_r$ sejam positivos, uma vez que, nos restantes casos, existe sempre um decréscimo de pelo menos um dos valores.

Deste modo, no fim do processo tem-se:

$$\begin{aligned} & \sum_{j \in L} \bar{c}_{sj}(x_j - l_j) + \sum_{j \in U} \bar{c}_{sj}(x_j - u_j) \geq \tilde{g}_s \\ \vee \\ & \sum_{j \in L} \bar{c}_{rj}(x_j - l_j) + \sum_{j \in U} \bar{c}_{rj}(x_j - u_j) \geq \tilde{g}_r \end{aligned}$$

Se $\tilde{g}_s > 0$ e $\tilde{g}_r > 0$, então o corte disjuntivo pode ser determinado como anteriormente. Caso contrário, o corte disjuntivo não será gerado.

Se na solução não básica uma das variáveis complementares z_i e w_i for superbásica, então

deve ser passada a básica por troca com uma variável nessas condições. Se essas modificações forem possíveis de modo a obter uma solução em que todas as variáveis complementares sejam básicas, então o processo para a construção do corte disjuntivo pode ser aplicado. De outro modo o corte disjuntivo não será construído.

(III) Exemplo de Construção dum Corte Disjuntivo

Considere-se um MPEC com o seguinte conjunto de restrições (GLCP):

$$\begin{aligned} 2x_1 + 3x_2 + x_3 &= 6 \\ -x_1 + x_2 + x_4 &= 1 \\ x_i &\geq 0, \quad i = 1, 2, 3, 4 \\ x_3x_4 &= 0 \end{aligned}$$

e seja $\bar{x} = (2, 0, 2, 3)$ uma solução não complementar não básica, com as variáveis básicas x_3 e x_4 , variável não básica x_2 no limite inferior nulo e variável superbásica x_1 . Como só há uma variável superbásica tem-se:

$$\begin{cases} x_3 = 2 - 2(x_1 - 2) - 3x_2 \\ x_4 = 3 - (-1)(x_1 - 2) - x_2 \end{cases}$$

Então

$$\begin{cases} x_3 \leq 0 & \Leftrightarrow 2(x_1 - 2) + 3x_2 \geq 2 \\ x_4 \leq 0 & \Leftrightarrow (-1)(x_1 - 2) + x_2 \geq 3 \end{cases}$$

e estamos no Caso 3. Como $u_1 = +\infty$, faz-se

$$S = S \setminus \{1\} \quad , \quad L = L \cup \{1\}$$

e actualiza-se $\tilde{g}_s$ e $\tilde{g}_r$ a partir de (4.5). Donde

$$\begin{aligned} \tilde{g}_r &= 2 + 2 \times 2 = 6 \\ \tilde{g}_s &= 3 + (-1) \times 2 = 1. \end{aligned}$$

Como $\tilde{g}_r > 0$, $\tilde{g}_s > 0$, então o corte disjuntivo pode ser gerado. Donde

$$\begin{array}{l} 2x_1 + 3x_2 \geq 6 \\ \text{ou} \\ -x_1 + x_2 \geq 1 \end{array} \left\{ \Rightarrow \begin{array}{l} \frac{1}{3}x_1 + \frac{1}{2}x_2 \geq 1 \\ \text{ou} \\ -x_1 + x_2 \geq 1 \end{array} \right\} \Rightarrow \frac{1}{3}x_1 + x_2 \geq 1$$

Como se pode observar na Figura a seguir, este corte restringe bastante a região admissível sem eliminar nenhuma das soluções complementares para o problema.

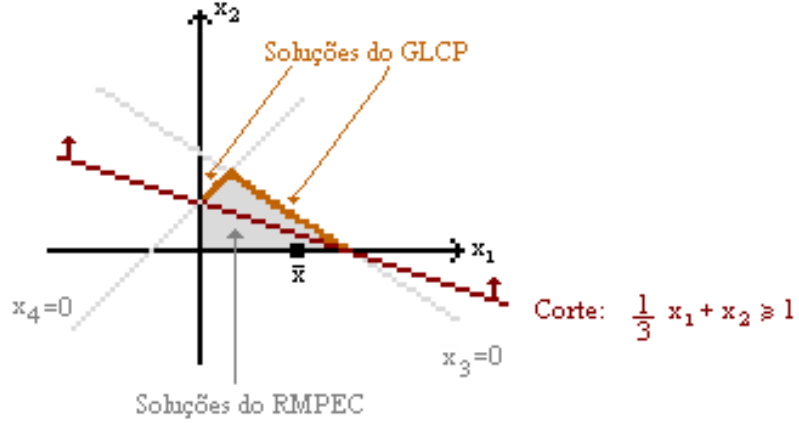


Figura 4.3: Exemplo de uma aplicação dos cortes disjuntivos

(IV) Aplicação dos Cortes Disjuntivos à Determinação de Limites Inferiores

A introdução de Cortes Disjuntivos é feita no Passo 3 do algoritmo de Ramificação e Limitação. Assim nesse passo é calculado primeiramente no nó t um limite inferior $LB(t)$ para o MPEC, resolvendo o problema relaxado $RMPEC_{J_t L_t}$, associado ao $MPEC_{J_t L_t}$, no qual a condição de complementaridade é relaxada.

Se $LB(t) \geq UB$, então o algoritmo volta ao Passo 1. Caso contrário, se a solução ótima é complementar, então essa solução fornece um novo limite superior UB . Como $UB=LB(t)$ então o algoritmo também volta ao Passo 1.

Suponhamos que a solução ótima do $RMPEC_{J_t L_t}$ é básica ou não básica não complementar. Para cada par (z_k, w_k) de variáveis complementares positivas e básicas, gera-se se possível o corte disjuntivo correspondente, de acordo com o processo discutido anteriormente. O programa relaxado com a adição desses cortes é resolvido e um novo limite inferior $LB(t)$ é determinado. O processo é repetido. Na prática, para cada solução não complementar deve-se procurar adicionar um número de cortes disjuntivos igual ao número de pares de

variáveis complementares positivas. No entanto essa opção poderá levar à geração de muitos cortes, tornando a resolução dos programas relaxados bastante pesada. Por isso, é normalmente recomendado que o número de cortes disjuntivos não seja muito elevado.

4.3 Determinação de Limites Superiores

Como referimos anteriormente, os limites inferiores são obtidos usando o algoritmo de restrições activas modificado ASETM para a determinação de um ponto estacionário do MPEC_{JL} associado a cada nó

$$\begin{aligned}
 & \text{Minimizar} && f(z, y) \\
 & \text{sujeito a} && Ew = q + Mz + Ny \\
 & && 0 = b - Ay \\
 & && w_i = 0, && i \in J \\
 & && z_i = 0, && i \in L \\
 & && z^T w = 0 \\
 & && z, w \geq 0
 \end{aligned} \tag{4.9}$$

com $z \in \mathbb{R}^n$, $w \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, $q \in \mathbb{R}^l$, $b \in \mathbb{R}^p$, $E \in \mathbb{R}^{l \times n}$, $M \in \mathbb{R}^{l \times n}$, $N \in \mathbb{R}^{l \times m}$ e $A \in \mathbb{R}^{p \times m}$ e os conjuntos J e L satisfazem $J \cap L = \emptyset$ e $J \cup L \subseteq \{1, \dots, n\}$.

Esse algoritmo necessita de uma solução do GLCP_{JL} que corresponde ao seu conjunto de restrições. No capítulo 1 mostrámos que se $J = L = \emptyset$, E é a matriz identidade e M é *PSD* ou uma matriz P , então uma solução do GLCP correspondente pode ser obtida a partir de um ponto estacionário de $z^T w$ no conjunto das restrições lineares desse GLCP. Nesta secção iremos investigar se esse resultado se mantém ou não válido para o GLCP_{JL}. Consideremos então esse problema e o programa quadrático associado

$$\begin{aligned}
 QP_{GLCP_{JL}} \quad & \text{Minimizar} && z^T(q + Mz + Ny) \\
 & \text{sujeito a} && (Mz + Ny + q)_i \geq 0, \quad i \notin J \\
 & && Ay + b = 0 \\
 & && (Mz + Ny + q)_i = 0, \quad i \in J \\
 & && z_i \geq 0, \quad i \notin L \\
 & && z_i = 0, \quad i \in L.
 \end{aligned}$$

Um ponto estacionário para este programa tem de satisfazer as condições

$$q + (M + M^T)z + Ny = M^T\alpha + \beta \quad (4.10)$$

$$N^T z = N^T\alpha + A^T\mu \quad (4.11)$$

$$\alpha_i \geq 0, \quad \alpha_i(Mz + Ny + q)_i = 0 \quad \forall i \notin J$$

$$\alpha_i \text{ sem restrição de sinal} \quad \forall i \in J$$

$$\beta_i \geq 0, \quad \beta_i z_i = 0 \quad \forall i \notin L$$

$$\beta_i \text{ sem restrição de sinal} \quad \forall i \in L$$

$$\mu \text{ sem restrição de sinal}$$

Então verifica-se o seguinte resultado.

Teorema 4.1 *Sejam $M \in PSD$ e E a matriz identidade.*

Se $(z, y, \alpha, \mu, \beta)$ é um ponto estacionário do programa quadrático $QP_{GLCP_{JL}}$ e $\alpha^T \beta \geq 0$ então (z, w, y) é uma solução do $GLCP_{JL}$.

Demonstração:

Para mostrar que (z, w, y) é uma solução do GLCP, temos que provar que

$$z^T (q + Mz + Ny) = 0.$$

Multiplicando as restrições (4.10) por $(\alpha - z)^T$, obtemos

$$(\alpha - z)^T (q + Mz + Ny) = (\alpha - z)^T M^T (\alpha - z) + (\alpha - z)^T \beta$$

Como $z^T \beta = 0$ então

$$(\alpha - z)^T \beta = \alpha^T \beta$$

e

$$\alpha^T (q + Mz + Ny) = \sum_{i \in J} \alpha_i (q + Mz + Ny)_i + \sum_{i \notin J} \alpha_i (q + Mz + Ny)_i = 0$$

Logo

$$-z^T (q + Mz + Ny) = (\alpha - z)^T M^T (\alpha - z) + \alpha^T \beta.$$

Como por hipótese, $\alpha^T \beta \geq 0$ e $M \in PSD$, então $(\alpha - z)^T M^T (\alpha - z) \geq 0$, e

$$z^T (q + Mz + Ny) = 0$$

Consequentemente, (z, y) é uma solução do $GLCP_{JL}$. $\diamond$

Como consequência deste teorema podemos estabelecer o seguinte resultado:

Teorema 4.2

1. Se $M \in PD$ e E é a identidade, então

$$z^T w = 0 \Leftrightarrow \alpha^T \beta \geq 0 \Leftrightarrow \alpha = z, \quad \beta = w \quad e \quad \alpha^T \beta = 0$$

2. Se M é hemi-simétrica ($M = -M^T$), então

$$\alpha^T \beta \geq 0 \Leftrightarrow z^T w = 0$$

Demonstração:

1. Se $\alpha^T \beta \geq 0$, então $z^T w = 0$ e $(\alpha - z)^T M^T (\alpha - z) \leq 0$ implica que $\alpha = z$. Portanto por (4.10), $\beta = q + Mz + Ny = w$ e $\alpha^T \beta = 0$.
2. Se M é hemi-simétrica então $(\alpha - z)^T M^T (\alpha - z) = 0$ e $z^T w = -\alpha^T \beta$. $\diamond$

Seguidamente são apresentadas algumas condições suficientes para que $\alpha^T \beta \geq 0$ e que por isso garantem que um ponto estacionário do $QP_{GLCP_{JL}}$ seja solução do $GLCP_{JL}$.

Teorema 4.3 *Sejam $M \in PSD$ e E a matriz identidade.*

Se $(z, y, \alpha, \mu, \beta)$ é um ponto de estacionário do programa quadrático $QP_{GLCP_{JL}}$ e se se verificar uma das seguintes condições

1. $J \cup L = \emptyset$
2. $z_i + w_i > 0 \quad \forall i \in J \cup L$
3. $\sum_{i \in J \cup L} \alpha_i \beta_i \geq 0$.

então $\alpha^T \beta \geq 0$ e (z, w, y) é solução do $GLCP_{JL}$.

Demonstração:

Como

$$\alpha^T \beta = \sum_{i \in J \cup L} \alpha_i \beta_i + \sum_{i \notin J \cup L} \alpha_i \beta_i$$

e

$$\alpha_i \geq 0 \text{ e } \beta_i \geq 0, \quad \forall i \notin J \cup L,$$

então

$$\alpha^T \beta \geq \sum_{i \in J \cup L} \alpha_i \beta_i.$$

Donde

1. Se $J \cup L = \emptyset$, então $\sum_{i \in J \cup L} \alpha_i \beta_i = 0$. Logo $\alpha^T \beta \geq 0$.

2. Se $z_i + w_i > 0 \ \forall i \in J \cup L$, então

$$\begin{cases} i \in J & \Rightarrow z_i > 0 & \Rightarrow \beta_i = 0 \\ i \in L & \Rightarrow w_i > 0 & \Rightarrow \alpha_i = 0 \end{cases}$$

Donde,

$$\sum_{i \in J \cup L} \alpha_i \beta_i = 0 \Rightarrow \alpha^T \beta \geq 0$$

3. Se $\sum_{i \in J \cup L} \alpha_i \beta_i \geq 0$, obviamente $\alpha^T \beta \geq 0$.

◇

O próximo exemplo mostra que é possível encontrar um ponto estacionário do $QP_{GLCP_{JL}}$ que não forneça uma solução do $GLCP_{JL}$. Consideremos o $GLCP_{JL}$

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 3 \end{bmatrix} y$$

$$z_1 = 0$$

$$w_1 \geq 0, w_2 \geq 0, z_2 \geq 0, y \geq 0$$

$$w_2 z_2 = 0$$

Então $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \in PD$ e portanto é PSD . Além disso $J = \emptyset$ e $L = \{1\}$. Este $GLCP$ tem solução. De facto, se considerarmos $z_2 = 0$, obtemos uma infinidade de soluções da forma $(w_1, w_2, z_1, z_2, y) = (-2 + y, 2 + 3y, 0, 0, y)$, com $y \geq 2$.

Um ponto estacionário $(w_1, w_2, z_1, z_2, y, \alpha_1, \alpha_2, \mu, \beta_1, \beta_2)$ para o programa quadrático associado a este $GLCP$ tem que satisfazer as seguintes condições:

$$\begin{aligned}
\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} &= \begin{bmatrix} -2 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 3 \end{bmatrix} y \\
z_1 &= 0 \\
\begin{bmatrix} -2 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 3 \end{bmatrix} y &= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \\
\begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} &= \begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \mu \\
\alpha_1 &\geq 0, \alpha_1 w_1 = 0 \\
\alpha_2 &\geq 0, \alpha_2 w_2 = 0 \\
\beta_2 &\geq 0, \beta_2 z_2 = 0 \\
\mu &\geq 0, \mu y = 0 \\
w_1 &\geq 0, w_2 \geq 0, z_2 \geq 0, y \geq 0 \\
\beta_1 &\text{ sem restrição de sinal}
\end{aligned}$$

Para mostrar que existe um ponto estacionário com $z^T w > 0$ temos que encontrar uma solução com $z_2 > 0$ e $w_2 > 0$. Pelo teorema anterior o ponto estacionário tem de ser degenerado, isto é, $w_1 = 0$. Assim, pelas restrições anteriores temos que, $\beta_2 = 0$ e $\alpha_2 = 0$, respectivamente. Consequentemente, o sistema anterior é equivalente a

$$\left\{ \begin{array}{l} 0 = -2 + z_2 + y \\ w_2 = 2 + z_2 + 3y \\ z_1 = 0 \\ -2 + y = \alpha_1 + \beta_1 \\ 2 + 3y + 2z_2 = \alpha_1 \\ 3z_2 = \alpha_1 + \mu \\ w_1 = 0 \\ \alpha_1 \geq 0 \\ \alpha_2 = 0 \\ \beta_2 = 0 \\ \mu \geq 0, \mu y = 0 \\ w_2 \geq 0, z_2 \geq 0, y \geq 0 \\ \beta_1 \text{ sem restrição de sinal} \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} w_1 = 0 \\ w_2 = 4 \\ z_1 = 0 \\ z_2 = 2 \\ y = 0 \\ \alpha_1 = 6 \\ \alpha_2 = 0 \\ \beta_1 = -8 \\ \beta_2 = 0 \\ \mu = 0 \end{array} \right.$$

Donde $z = (0, 2)$ é um ponto estacionário com $z^T w > 0$. É evidente que existe um ponto estacionário de $QP_{GLCP_{JL}}$ que fornece uma solução do GLCP. Com efeito se escolhermos $z_2 = 0$, então tem-se

$$\left\{ \begin{array}{l} 0 = -2 + y \\ w_2 = 2 + 3y \\ z_1 = 0 \\ z_2 = 0 \\ w_1 = 0 \\ -2 + y = \alpha_1 - \alpha_2 + \beta_1 \\ 2 + 3y = \alpha_1 + \alpha_2 + \beta_2 \\ 0 = \alpha_1 + 3\alpha_2 + \mu \\ \alpha_1 \geq 0 \\ \alpha_2 \geq 0, \alpha_2 w_2 = 0 \\ \beta_2 \geq 0 \\ \mu \geq 0, \mu y = 0 \\ w_2 \geq 0, y \geq 0 \\ \beta_1 \text{ sem restrição de sinal} \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} w_1 = 0 \\ w_2 = 8 \\ z_1 = 0 \\ z_2 = 0 \\ y = 2 \\ \alpha_1 = 0 \\ \alpha_2 = 0 \\ \beta_1 = 0 \\ \beta_2 = 8 \geq 0 \\ \mu = 0 \end{array} \right.$$

e, $z = (0, 0)$, $y = 2$ é um ponto estacionário com $z^T w = 0$. Notar que $\alpha_1 \beta_1 = 0$.

Como iremos ver na última secção deste capítulo, a experiência computacional com MPECs associados a problemas de dois níveis e a LCPs indica que em geral a solução do $GLCP_{JL}$ se obtém a partir do ponto estacionário de $QP_{GLCP_{JL}}$ associado. Se tal não acontecer, então o $GLCP_{JL}$ deve ser resolvido usando o método enumerativo EASET descrito no capítulo 2. É de notar que esse algoritmo começa por determinar um ponto estacionário do $GLCP_{JL}$ no nó 1 da árvore binária que vai gerando. Na prática o método de Ramificação e Limitação é implementado de forma a que os nós para os quais a obtenção de uma solução do $GLCP_{JL}$ necessite de uma pesquisa em árvore são apenas investigados no fim do processo. Resultados computacionais apresentados na última secção deste capítulo mostram que dessa forma todos os $GLCP_{JL}$ investigados pelo método de Ramificação e Limitação são resolvidos simplesmente a partir da determinação de um ponto estacionário do $QP_{GLCP_{JL}}$ correspondente.

4.4 Algoritmo Sequencial Complementar

Nesta secção iremos descrever de um modo sucinto o Algoritmo Sequencial Linear Complementar (SLCP) desenvolvido por Júdice e Faustino [64] para a resolução do MPEC linear, isto é, para o caso em que a função f é linear da forma

$$f(z, y) = c^T z + d^T y$$

Neste método é introduzido um parâmetro λ e a função objectivo é substituída pela restrição

$$c^T z + d^T y \leq \lambda \quad (4.12)$$

obtendo-se o seguinte problema linear complementar

$$\begin{aligned} &GLCP(\lambda) \\ &\begin{bmatrix} w \\ 0 \\ \mu \end{bmatrix} = \begin{bmatrix} q \\ b \\ \lambda \end{bmatrix} + \begin{bmatrix} M & N \\ 0 & A \\ -c^T & -d^T \end{bmatrix} \begin{bmatrix} z \\ y \end{bmatrix} \\ &w, \mu, z, y \geq 0 \\ &w^T z = 0 \end{aligned} \quad (4.13)$$

O algoritmo sequencial complementar consiste em encontrar uma solução do $GLCP(\bar{\lambda})$ onde $\bar{\lambda}$ é o menor valor de λ , para o qual o $GLCP(\lambda)$ tem uma solução. Para encontrar essa solução, o método começa por resolver o $GLCP(\lambda_0)$ obtido do $GLCP(\lambda)$ (4.13), omitindo a restrição $c^T z + d^T y \leq \lambda$. Seja (z^0, y^0) a solução do $GLCP(\lambda_0)$ e

$$\lambda_0 = f(z^0, y^0) = c^T z^0 + d^T y^0$$

O algoritmo resolve então uma sequência de GLCPs(λ_k), onde $\{\lambda_k\}$ é uma sucessão decrescente definida por

$$\lambda_k = f(z^{k-1}, y^{k-1}) - \gamma |f(z^{k-1}, y^{k-1})|$$

com (z^{k-1}, y^{k-1}) solução do $GLCP(\lambda_{k-1})$ e γ um número positivo pequeno. O processo termina numa iteração k para a qual o $GLCP(\lambda_k)$ não tem solução. Quando isso acontece, a solução (z^{k-1}, y^{k-1}) do $GLCP(\lambda_{k-1})$ satisfaz

$$0 \leq f(z^{k-1}, y^{k-1}) - f(\bar{z}, \bar{y}) \leq \gamma |f(z^{k-1}, y^{k-1})|$$

onde $(\bar{w}, \bar{z}, \bar{y})$ é solução óptima do MPEC (3.4).

Assim, se o conjunto de soluções admissíveis do MPEC é não vazio e limitado, o algoritmo SLCP encontra uma solução ϵ -óptima do MPEC com

$$\epsilon = \gamma |f(z^{k-1}, y^{k-1})| \quad (4.14)$$

Na prática, se γ for suficientemente pequeno, a solução (z^{k-1}, y^{k-1}) do penúltimo GLCP(λ_{k-1}) é normalmente o mínimo global do MPEC. Os passos do algoritmo SLCP são os seguintes:

Algoritmo Sequencial Complementar (SLCP)

Passo 0 Faça $k = 0$.

Passo 1 Resolva o GLCP(λ_k). Se GLCP(λ_k) não tem solução, vá para o Passo 2.

Caso contrário, seja (z^k, w^k, y^k) a solução do GLCP(λ_k). Faça

$$\lambda_{k+1} = f(z^k, y^k) - \gamma |f(z^k, y^k)|$$

onde γ é um número positivo pequeno previamente fixado. Faça $k = k + 1$ e repita o passo.

Passo 2 Se $k = 0$, o MPEC não tem solução. Caso contrário, (z^{k-1}, y^{k-1}) é uma solução ϵ -óptima do MPEC, com ϵ dado por (4.14).

É de notar que, quando se pretende resolver o GLCP(λ_k), a solução do GLCP(λ_{k-1}) já é conhecida e pode por isso ser usada como solução inicial para o GLCP(λ_k). A escolha desta solução básica inicial, apesar de não ser admissível, conduz normalmente a uma grande redução no esforço computacional do método enumerativo que é usado para resolver o GLCP. De facto, o algoritmo requer normalmente um reduzido número de operações pivotais para encontrar a solução do GLCP(λ_k) [63, 64]. Para redução do número de GLCPs a resolver, o método SLCP incorpora o Algoritmo das Restrições Activas Modificado ASETM para obter em cada iteração k uma solução do GLCP(λ_k), com valor da função objectivo do MPEC inferior à da anteriormente obtida pelo método enumerativo.

Como discutimos anteriormente, o método enumerativo é normalmente muito eficaz para

determinar uma solução do GLCP. Por isso o algoritmo SLCP consegue obter normalmente a solução óptima do MPEC num esforço computacional razoável. Contudo, para ter a certeza de que o mínimo global do MPEC foi obtido é necessário mostrar que o último GLCP não tem solução. Como afirmámos no capítulo 2, o método enumerativo necessita de uma pesquisa exaustiva para esse efeito. Como conclusão final podemos afirmar que o método SLCP é eficaz para obter um mínimo global mas tem grande dificuldade para assegurar que essa solução foi obtida. Experiência computacional descrita neste trabalho e em [64, 65] comprova exactamente este comportamento.

A implementação do Algoritmo Sequencial Complementar é uma adaptação simples do processo descrito para o Algoritmo de Restrições Activas Modificado e apresentado na secção 3.6. Esta nova versão do Algoritmo Sequencial Linear Complementar, permite também resolver MPECs cuja função é não linear convexa. Nesse caso a solução do $GLCP_{JL}$ é determinada usando o método enumerativo, que incorpora um processo para a determinação de um ponto estacionário de uma função quadrática num conjunto convexo definido por restrições lineares e uma restrição convexa

$$f(z, y) \leq \lambda_k$$

4.5 Experiência Computacional

Nesta secção, é apresentada alguma experiência computacional realizada com os dois algoritmos de optimização global discutidos neste capítulo. Estas experiências foram realizadas num Pentium IV 2.4GHz com 256 MB de RAM. Os problemas teste são os apresentados no capítulo anterior.

(I) Algoritmo de Ramificação e Limitação

Os MPECs considerados são associados aos problemas de dois níveis lineares (LBP) e quadráticos (QBP) do gerador descrito em [18, 19] e LCPs associados a problemas de mochila (MOCHILA) discutidos nos capítulos anteriores. Em todos os casos é conhecido o valor óptimo da função do MPEC. Como esse valor óptimo é conhecido, então utilizou-se em cada nó k esse valor como limite inferior $LB(k)$. Os resultados das experiências dessa versão simplificada do

método de ramificação e limitação são apresentados na Tabela 4.1 e indicam que o processo é em geral bastante eficiente para determinar o mínimo global do MPEC. É de notar que em todos os problemas teste o $GLCP_{JL}$ em cada nó foi sempre resolvido a partir do ponto estacionário do $QP_{GLCP_{JL}}$ referido na secção 4.3. Isso significa que nunca houve necessidade de usar o método enumerativo EASET para a resolução desses GLCPs e da consequente determinação do ponto estacionário do $MPEC_{JL}$ em cada nó.

Na segunda experiência, resolveram-se os MPECs associados aos programas de dois níveis LBP e QBP sem prévio conhecimento dos seus valores óptimos. Cortes disjuntivos foram então implementados, no processo de Ramificação e Limitação de acordo com o procedimento discutido na secção 4.2. Os resultados da aplicação do método de ramificação e limitação na resolução desses problemas teste são apresentados na tabela a seguir, onde CT designa o número máximo de cortes incorporados em cada nó.

Os resultados indicam que o uso dos cortes disjuntivos no algoritmo de ramificação e limitação parece ser uma ideia interessante pelo menos para os problemas em causa. Com efeito o algoritmo determina a solução óptima em muitos casos sem necessidade de gerar a árvore binária. Como veremos em capítulos posteriores a eficácia dos cortes disjuntivos não é sempre assim notória, particularmente quando a dimensão dos MPECs aumenta. No entanto cortes disjuntivos devem ser em geral recomendados na prática e usados na implementação do método de ramificação e limitação.

(II) Método Sequencial Complementar

Nesta subsecção são apresentados os resultados obtidos pelo Método Sequencial Complementar (SLCP) na resolução de MPECs referidos anteriormente. Como discutimos no capítulo 2, é possível resolver o $GLCP(\lambda_k)$ usando uma das versões EMRG ou EASET. Assim iremos apresentar os resultados das experiências com as seguintes versões do método Sequencial Complementar:

- SLCP/EMRG, que incorpora a versão do método enumerativo baseada no Algoritmo de Gradiente Reduzido Modificado (MRG);
- SLCP/EASET, que incorpora a versão do método enumerativo baseada no Algoritmo de Restrições Activas.

	PROB	NI	T	ND	OP.OBJ.
L B P	Lv1	14	0.02	2	6.000
	Lv1M	10	0.01	1	6.000
	Lv2	47	0.03	15	16.000
	Lv2M	21	0.02	1	16.000
	Lv3	182	0.05	114	30.200
	Lv3M	87	0.06	5	30.200
	Lv4	35	0.02	2	6.000
	Lv5	37	0.02	2	14.000
	Lv6	65	0.06	3	7.500
	Lv7	81	0.03	3	23.500
	Lv8	283	0.02	125	50.500
Q B P	Nlv1	12	0.02	2	0.313
	Nlv2	25	0.02	4	0.593
	Nlv3	28	0.00	2	0.790
	Nlv4	65	0.02	2	1.040
	Nlv5	135	0.08	16	2.436
M O C H I L A	1NA1	66	0.02	2	0.000
	1NA2	11894	2.71	280	0.000
	1NA3	206	0.06	2	0.000
	1NA4	528	0.17	2	0.000
	1NB1	304	0.19	18	0.000
	1NB2	72	0.05	2	0.000
	1NB3	415	0.17	2	0.000
	1NB4	205	0.09	2	0.000
	1NC1	163	0.09	6	0.000
	1NC2	217	0.06	2	0.000
	1NC3	208	0.11	2	0.000
	1NC4	162	0.11	2	0.000
	1IA1	66	0.05	2	0.000
	1IA2	11886	2.85	280	0.000
	1IA3	206	0.06	2	0.000
	1IA4	528	0.27	2	0.000
M O C H I L A	1IB1	304	0.06	18	0.000
	1IB2	72	0.02	2	0.000
	1IB3	415	0.19	2	0.000
	1IB4	205	0.09	2	0.000
	1IC1	163	0.05	6	0.000
	1IC2	217	0.06	2	0.000
	1IC3	208	0.08	2	0.000
	1IC4	162	0.09	2	0.000
	2NA1	191	0.06	2	0.000
	2NA2	443	0.14	2	0.000
	2NA3	974	0.41	2	0.000
	2NA4	593	0.27	2	0.000
	2NB1	1357	0.25	46	0.000
	2NB2	638	0.14	2	0.000
	2NB3	256	0.09	2	0.000
	2NB4	682	0.36	2	0.000
	2NC1	29	0.08	2	0.000
	2NC2	1126	0.33	22	0.000
	2NC3	1269	0.33	2	0.000
	2NC4	733	0.36	2	0.000
M O C H I L A	3PA1	328	0.11	27	0.000
	3PA2	3457	0.56	113	0.000
	3PA3	1516	0.36	3	0.000
	3PA4	1102	0.36	2	0.000
	3PB1	34	0.03	2	0.000
	3PB2	154	0.05	2	0.000
	3PB3	726	0.17	2	0.000
	3PB4	3864	1.23	8	0.000
	3PC1	31	0.06	1	0.000
	3PC2	72	0.02	1	0.000
	3PC3	209	0.08	2	0.000
	3PC4	343	0.11	2	0.000

Tabela 4.1: Algoritmo de Ramificação e Limitação para MPECs com valor óptimo conhecido

	PROB	NI	T	ND	CT	OBJ
L B P	Lv1	22	0.00	1	11	6.000
	Lv1M	197	0.02	5	12	6.000
	Lv2	53	0.01	1	24	16.000
	Lv2M	62	0.01	1	17	16.000
	Lv3	102	0.03	1	51	30.200
	Lv3M	116	0.02	1	35	30.200
	Lv4	43	0.02	1	11	6.000
	Lv5	59	0.00	1	23	14.000
	Lv6	74	0.01	1	16	7.500
	Lv7	105	0.03	1	38	23.500
	Lv8	170	0.14	1	74	50.500
Q B P	NLv1	22	0.00	3	2	0.313
	NLv2	45	0.00	5	4	0.593
	NLv3	78	0.02	9	4	0.790
	NLv4	164	0.05	16	5	1.040
	NLv5	4807	1.16	515	13	2.436

Tabela 4.2: Algoritmo de Ramificação e Limitação com Cortes Disjuntivos

As Tabelas 4.3 e 4.4 incluem o resumo dos resultados obtidos na resolução de MPECs lineares correspondentes aos programas de dois níveis LBP e aos LCPs associados a problemas de mochila. Nestas e nas restantes tabelas, IT designa o número de GLCPs(λ_k) resolvidos pelo algoritmo SLCP (incluindo o GLCPs(λ_0)), NI e ND os números totais de operações pivotais e de nós gerados pelo processo. Relacionados com a obtenção da última solução incumbente são ainda usados os parâmetros NIS, TS e NDS que representam, respectivamente, o número de iterações (operações pivotais), tempo CPU em segundos e o número de nós pesquisados até obtenção dessa solução.

	SLCP/EMRG				SLCP/EASET			
	NI	T	IT	ND	NI	T	IT	ND
PIOR	18307	10.06	10	16541	5581	1.42	13	1650
MEDIA	1442	0.48	5	602	1159	0.37	5	263
MELHOR	38	0.00	1	1	31	0.00	1	1

Tabela 4.3: Resolução dos problemas da Mochila pelo SLCP

Como o valor óptimo dos MPECs associados a qualquer LCP é igual a zero, o método SLCP termina quando se obtém um valor λ_k inferior ou igual ao valor óptimo do problema em questão.

	SLCP/EMRG							SLCP/EASET						
	Ni	T	IT	ND	Nis	Ts	Nds	Ni	T	IT	ND	Nis	Ts	Nds
PIOR	44781	5.27	7	20834	185	0.02	73	41457	22.55	10	39272	360	0.08	286
MÉDIA	4631	0.50	4	2220	75	0.00	22	4329	2.22	5	4140	89	0.02	43
MELHOR	26	0.00	2	18	12	0.00	1	21	0.00	2	20	10	0.00	1

Tabela 4.4: Resolução dos Problemas de dois níveis lineares pelo SLCP

Esses resultados mostram que as duas versões são comparáveis e que ambas as versões do algoritmo determinam as soluções óptimas de todos os problemas em tempo aceitável. Essa experiência e a relatada em [64, 65] indicam que o algoritmo SLCP é muito recomendável para a determinação do mínimo global associado a Problemas de Dois Níveis e a LCPs.

O algoritmo SLCP tem dificuldades em provar que se obteve o mínimo global dos MPECs de valor óptimo desconhecido. Com efeito, os resultados mostram que o método enumerativo tem nesse caso de efectuar uma pesquisa muito grande para estabelecer que o último GLCP não tem solução. Isso confirma a análise feita em [64, 65].

Uma vantagem da versão baseada no método das restrições activas é a sua aplicabilidade a MPECs não lineares. Em particular o algoritmo pode ser usado para determinar o mínimo global dos Problemas de Dois Níveis com função quadrática BQP referidos anteriormente. O esforço computacional para esses problemas é apresentado na Tabela 4.5. Notar que os valores

		SLCP/EASET				OP.OBJ.
PROB		Ni	T	IT	ND	
B Q P	Nlv1	15	0.00	3	2	0.313
	Nlv2	30	0.03	3	2	0.593
	Nlv3	31	0.00	1	2	0.790
	Nlv4	104	0.08	3	3	1.040
	Nlv5	159	0.08	3	3	2.436

Tabela 4.5: SLCP para MPECs associados a Problemas de Dois Níveis Quadráticos

óptimos desses problemas são conhecidos, pelo que não houve necessidade de tratar o último GLCP. Esses resultados indicam que o processo é muito eficiente para a determinação de um mínimo global desses MPECs não lineares. Outros estudos computacionais são necessários para aferir da validade do método para a resolução de MPECs não lineares.

Capítulo 5

PROBLEMA DE AJUSTE DE FUNÇÕES DOBRADIÇAS

Este capítulo é dedicado à resolução do problema de ajuste de funções dobradiças, que aparece na teoria de aproximação [15, 99]. Como é discutido em [102] esse problema pode ser reformulado como um MPEC linear ou quadrático, dependendo da norma escolhida ser l_1 ou l_2 . A aplicabilidade dos métodos de otimização global para o MPEC discutidos no capítulo 4 é investigada e são apresentados resultados computacionais que atestam a validade dessas técnicas para a resolução desses problemas.

5.1 Definição do Problema

O problema de ajuste de uma função a um conjunto de dados obtidos normalmente por medição é comum em várias áreas da ciência. O processo mais usado para efectuar este tipo de ajuste é o Método dos Mínimos Quadrados e tem importantes aplicações, nomeadamente em Métodos Estatísticos, Sistemas Dinâmicos e Processamento de Sinal.

Outros tipos de métodos têm sido propostos para a resolução do problema de ajuste de funções quer lineares quer não lineares. Esses processos baseiam-se em Otimização Discreta ou Contínua, Redes Neurais, Onduletas (Wavelets) ou na Geometria Computacional [11, 23, 59]. As escolhas de uma função para medição dos erros e de funções básicas estão

subjacentes ao problema de ajuste de funções não lineares. Na medição dos erros está implícita a minimização de uma certa norma das diferenças entre a função original e a aproximação desta nos pontos dados. Consoante a norma utilizada diferentes objectivos de aproximação são definidos. As normas mais correntes são as l_1 , l_2 e l_∞ , sendo a norma euclideana l_2 a mais usada.

A escolha das funções básicas depende da classe de funções a serem aproximadas e são normalmente funções indicadoras, polinómios, aproximações segmentadas, funções trigonométricas ou sigmóides.

Em [15] foi considerado o ajustamento de n pontos em $\mathbb{R}^m$ através de funções dobradiças. Este problema, designado por Ajuste de Funções Dobradiças, aparece nas aplicações de redes neuronais, regressão não linear e na classificação de dados. Um estudo mais detalhado sobre este tipo de problemas aparece em [15, 99, 102].

O modelo de hiperplanos dobrados (HD) recebeu a atenção na literatura recente [15, 99, 102] como uma alternativa às funções sigmóides dos modelos de redes neuronais para ajustamento de funções não lineares. O modelo HD usa funções dobradiças, ou seja, funções da forma

$$h(a) = \max \{x^T a + x_m, y^T a + y_m\} \quad \text{ou} \quad h(a) = \min \{x^T a + x_m, y^T a + y_m\},$$

como funções básicas em expansões tais como

$$f^k(a) = \sum_{i=1}^k h_i(a),$$

onde $f^k(\cdot)$ é uma aproximação de uma função $f(\cdot)$, cuja expressão é desconhecida.

Em [15] foi mostrado que se a função $f(\cdot)$ for suficientemente diferenciável, então existe um $\alpha > 0$ tal que para qualquer $k \in \mathbb{N}$, existem k funções dobradiças $h_1, \dots, h_k$, que satisfazem

$$\|f - \sum_{i=1}^k h_i\|^2 \leq \frac{\alpha}{k}.$$

Um esquema de aproximação foi proposto em [15], que consiste em ajustar os valores da função f dados a uma dobradiça h_1 , depois reajustar a diferença $f - h_1$ a uma dobradiça h_2 , e então reajustar a diferença $f - h_1 - h_2$ a uma dobradiça h_3 , e assim sucessivamente. Ao fim de k funções dobradiças, pode redefinir-se, de modo ordenado, cada uma das funções h_j como a solução do problema de ajuste de dobradiças aos valores de $(f - \sum_{i \neq j} h_i)$. É

efectuado sucessivamente um refinamento à aproximação até que não se obtenha qualquer melhoria significativa.

Em [15] foi também proposto um processo iterativo para o problema de ajustamento de funções dobradiças, chamado Algoritmo para Obtenção de Dobradiça (HFA - Hinge Finding Algorithm). Trata-se de um método de ponto-fixo que particiona os dados em dois conjuntos e resolve independentemente um problema de regressão linear para cada conjunto. As duas funções lineares obtidas permitem definir uma nova partição dos dados originais, que é usada como dados para a próxima iteração. Diz-se que o algoritmo converge se as partições em duas iterações seguidas são idênticas. Quando o algoritmo converge, a solução obtida é um minimizador local do problema de mínimos quadrados. Além disso, se os dados originais representam medidas exactas de uma função dobradiça e a partição original não está muito distante da correcta, então o algoritmo converge [15]. Apesar da sua simplicidade e fácil implementação, esta abordagem tem algumas desvantagens [99]. Em [99] é proposta uma modificação do método a fim de garantir convergência global para um mínimo local. Esse método HFA amortecido é um processo tipo Newton que incorpora uma técnica de pesquisa em linha para garantir decréscimo da função erro em cada iteração e assim garantir convergência global.

Como acontece frequentemente ao aplicar-se métodos de optimização local para problemas não-convexos, a qualidade dos mínimos locais obtidos depende fortemente da solução inicial. A grande variação observada na qualidade dos mínimos locais em experiências realizadas, levou à consideração de um método de optimização global para o problema de ajuste de funções dobradiças. Este assunto é tratado em [102] explorando duas reformulações do problema como MPECs. O problema de ajuste de dobradiças com erro dado pela norma euclideana (Ajuste Euclideano de Dobradiças) é formulado como um MPEC quadrático, e o ajuste de dobradiças com erro dado pela norma l_1 (Ajuste Linear de Dobradiças) reduz-se a um MPEC linear. Para o mínimo global do MPEC quadrático foi proposto um método enumerativo, polinomial no tamanho n da amostra, para dimensão m fixa do espaço [102]. Foram realizadas algumas experiências computacionais com esse algoritmo enumerativo, para problemas de média dimensão que indicaram que a qualidade dos ajustamentos é melhorada substancialmente em relação à dos obtidos pelo algoritmo HFA amortecido. Para o problema de Ajuste Linear de Dobradiças, foi ainda sugerida a versão MRG do algoritmo sequencial complementar discutido no capítulo 4. A experiência computacional apresentada em [102]

mostra que este método processa eficientemente os mesmos problemas de média dimensão reformulados com a norma l_1 . Além disso o método SLCP mostrou ser mais eficiente do que o algoritmo enumerativo para determinar o mínimo global do MPEC linear associado a esse problema de ajuste linear.

Neste capítulo é investigada a utilização dos algoritmos de otimização global descritos no capítulo 4 na resolução dos MPECs lineares e quadráticos mencionados anteriormente.

5.2 O Problema do Ajuste Euclidiano de Funções Dobradiças

Definição 5.1 *Sejam $(a^i, b_i) \in \mathbb{R}^{m-1} \times \mathbb{R}$, para $i = 1, \dots, n$, as medidas dadas de uma função f tal que $b_i = f(a^i)$. O Problema do Ajuste Euclidiano de Funções Dobradiças consiste em aproximar f por uma função dobradiça h tal que para cada $a = a^i$ e $b = b_i$, $h(a) = \max \{a^T \bar{x} + \bar{x}_m, a^T \bar{y} + \bar{y}_m\}$, onde $(\bar{x}, \bar{x}_m)$ e $(\bar{y}, \bar{y}_m)$ minimizam a norma euclidiana dos erros da aproximação $\|h(a) - b\|_2$.*

No sentido de obter uma formulação MPEC do problema, sejam $A_i = [(a^i)^T, 1]$, $i = \dots, n-1$ e notemos por $x \in \mathbb{R}^m$ e $y \in \mathbb{R}^m$ os vectores

$$\begin{bmatrix} x \\ x_m \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} y \\ y_m \end{bmatrix}$$

respectivamente. Então o problema de ajuste euclidiano pode ser escrito na forma

$$\min_{x,y} \sum_{i=1}^n (\max \{A_i x, A_i y\} - b_i)^2$$

Introduzindo uma variável adicional σ , é fácil de concluir que esse problema se reduz ao MPEC

$$\left. \begin{array}{ll} \min & \sum_{i=1}^n \sigma_i^2 \\ \text{s.a.} & \sigma_i \geq A_i x - b_i \\ & \sigma_i \geq A_i y - b_i \\ & (\sigma_i - A_i x + b_i)(\sigma_i - A_i y + b_i) = 0, \end{array} \right\} \quad i = 1, \dots, n,$$

ou equivalentemente

$$\begin{aligned}
FDE \quad & \min \quad \sigma^T \sigma \\
& \text{s.a.} \quad r = \sigma - Ax + b \\
& \quad \quad s = \sigma - Ay + b \\
& \quad \quad r, s \geq 0 \\
& \quad \quad r^T s = 0
\end{aligned}$$

Como a função objectivo do problema FDE é convexa, podemos aplicar o método de Ramificação e Limitação para determinar um seu mínimo global. A estrutura especial do problema torna a solução do GLCP_{JL} em cada nó muito simples. Com efeito, se $(\bar{x}, \bar{y}, \bar{r}, \bar{s})$ é a solução óptima do programa relaxado RMPEC_{JL}, então é fácil de obter uma solução do GLCP a partir das seguintes transformações:

$$\begin{aligned}
r_i &= r_i - \min \{r_i, s_i\} \\
s_i &= s_i - \min \{r_i, s_i\} \\
\sigma_i &= \sigma_i - \min \{r_i, s_i\}.
\end{aligned}$$

A solução assim obtida só poderá ser usada como solução inicial para a resolução do MPEC_{J_tL_t}, no Passo 4 do algoritmo, se for devidamente alterada de acordo com a *Modificação 1*, descrita no capítulo 3. Nesta modificação são efectuadas operações pivotais ou é reduzido o conjunto das variáveis superbásicas, de modo a obter uma solução do GLCP em que para cada par de variáveis complementares, pelo menos uma dessas variáveis é não básica.

5.3 O Problema do Ajuste Linear de Funções Dobradiças

Uma formulação alternativa do problema de Ajuste de Dobradiças baseia-se no uso da norma l_1 para medição dos erros. Se introduzirmos a variável $\gamma \in \mathbb{R}^m$ tal que

$$\gamma_i = |\max \{A_i x, A_i y\} - b_i|, \quad i = 1, \dots, m$$

então da definição da norma l_1 é possível formular o Problema de Ajuste Linear de Dobradiças como o seguinte MPEC

$$\begin{aligned}
\min \quad & e^T \gamma \\
\text{s.a.} \quad & r = \sigma - Ax + b \\
& s = \sigma - Ay + b \\
& \gamma \geq \sigma \\
& \gamma \geq -\sigma \\
& r, s \geq 0 \\
& r^T s = 0
\end{aligned}$$

onde $e = (1, \dots, 1)^T \in \mathbb{R}^n$. Substituindo a variável σ por variáveis não negativas α e β ,

$$\sigma = \beta - \alpha$$

com $\alpha^T \beta = 0$, então [102] o problema anterior pode ser reformulado como o seguinte MPEC

$$\begin{aligned}
LHF \quad \min \quad & e^T (\alpha + \beta) \\
\text{s.a.} \quad & r = \beta - \alpha - Ax + b \\
& s = \beta - \alpha - Ay + b \\
& \beta, \alpha, r, s \geq 0 \\
& \beta^T \alpha = r^T s = 0
\end{aligned}$$

É ainda possível estabelecer o seguinte resultado.

Teorema 5.1 *A restrição $\beta^T \alpha = 0$ em LHF é redundante*

Demonstração:

Se

$$(\bar{\alpha}, \bar{\beta}, \bar{x}, \bar{y}, \bar{r}, \bar{s})$$

é solução ótima de LHF sem a restrição $\beta^T \alpha = 0$, então

$$\bar{\beta}^T \bar{\alpha} = 0$$

Com efeito, se $\bar{\beta}^T \bar{\alpha} > 0$ então existe pelo menos um i tal que

$$\bar{\beta}_i > 0, \quad \bar{\alpha}_i > 0$$

Consideremos, sem perda de generalidade, que $\bar{\beta}_i > \bar{\alpha}_i$ e sejam

$$\begin{cases} \tilde{\alpha}_i = 0 \\ \tilde{\beta}_i = \bar{\beta}_i - \bar{\alpha}_i \end{cases}$$

Então

$$\begin{aligned} \tilde{\beta}_i - \tilde{\alpha}_i &= \bar{\beta}_i - \bar{\alpha}_i \\ \tilde{\beta}_i &\geq 0, \quad \tilde{\alpha}_i \geq 0 \quad \tilde{\alpha}_i \tilde{\beta}_i = 0 \\ \tilde{\beta}_i + \tilde{\alpha}_i &= \bar{\beta}_i - \bar{\alpha}_i < \bar{\beta}_i + \bar{\alpha}_i \end{aligned}$$

Isso é impossível, pois $(\bar{\alpha}, \bar{\beta}, \bar{x}, \bar{y}, \bar{r}, \bar{s})$ é a solução ótima de LHF. Se usarmos o mesmo raciocínio para todos os pares (α_i, β_i) , facilmente se conclui que $\alpha_i \beta_i = 0$ para todo $i = 1, \dots, n$. $\diamond$

Devido a este teorema, a condição de complementaridade $\alpha^T \beta = 0$ pode ser omitida de LHF, o que permite escrever a formulação do problema de ajuste linear de dobradiças como o seguinte MPEC

$$\begin{aligned} LHF \quad & \min \quad e^T(\alpha + \beta) \\ \text{s.a.} \quad & r = \beta - \alpha - Ax + b \\ & s = \beta - \alpha - Ay + b \\ & \beta, \alpha, r, s \geq 0 \\ & r^T s = 0 \end{aligned}$$

O Algoritmo de Ramificação e Limitação pode ser utilizado para a resolução deste MPEC. Tal como no caso quadrático associado ao problema de ajuste euclidiano, o GLCP_{JL} em cada nó pode ser resolvido facilmente por transformações simples. Seja $(\bar{r}, \bar{s}, \bar{x}, \bar{y}, \bar{\alpha}, \bar{\beta})$ a solução ótima do programa linear relaxado RMPEC_{JL}. Então para cada i tal que $\bar{r}_i \bar{s}_i > 0$, faça

$$\begin{aligned} \bar{r}_i &= \bar{r}_i - \min \{\bar{r}_i, \bar{s}_i\} \\ \bar{s}_i &= \bar{s}_i - \min \{\bar{r}_i, \bar{s}_i\} \\ \bar{\alpha}_i &= \bar{\alpha}_i + \min \{\bar{r}_i, \bar{s}_i\} \end{aligned}$$

permanecendo $\bar{\beta}$, $\bar{x}$ e $\bar{y}$ sem alteração. Então para cada i nas condições anteriores uma das variáveis $\bar{r}_i$ e $\bar{s}_i$ fica nula, enquanto que α_i é positiva. Então α_i pode passar a básica por troca com a variável complementar que passa a ser nula e não básica. Assim se obtém uma solução básica para o GLCP_{JL} .

Como o MPEC associado a este Problema de Ajustes é linear, então o Método Sequencial Complementar também pode ser usado para o resolver. É de notar que nesse caso apenas o primeiro GLCP é fácil de resolver. Com efeito o processo para a resolução do GLCP anterior não é possível de aplicar aos $\text{GLCP}(\lambda_k)$ com $k > 1$.

5.4 Experiência Computacional

A experiência computacional relatada nesta secção foi realizada num Pentium IV 2.4GHz com 256 MB de RAM.

A Tabela 5.1 apresenta os resultados obtidos na resolução do Problema Euclideano de Ajuste de Funções Dobradiças FDE com o algoritmo de Ramificação e Limitação.

PROB	m	n	NV	NRL	NI	T	Nd	NIs	Ts	Nds	OBJ.
1	5	30	100	60	14597	1.77	227	11972	1.47	174	1.2024
2	10	35	125	70	271215	42.08	1855	200661	31.93	1347	0.5307
3	15	40	150	80	2910363	596.30	15974	1529475	309.62	8252	0.1226
4	20	45	175	90	>10000000	2968.49	44255	3019314	807.54	12660	0.0029
5	25	50	200	100	54826	19.0	44	54826	19.0	44	0.0000
6	30	55	225	110	632	0.25	1	632	0.25	1	0.0000
7	35	60	250	120	2268	1.03	2	2268	1.03	2	0.0000
8	40	65	275	130	698	0.38	1	698	0.38	1	0.0000
9	45	70	300	140	589	0.41	1	589	0.41	1	0.0000
10	50	75	325	150	823	0.52	1	823	0.52	1	0.0000

Tabela 5.1: Algoritmo de Ramificação e Limitação para o Problema de Ajuste Euclideano de Funções Dobradiças

Nessa tabela e nas restantes desta secção, NI é o número total de iterações executadas, ND representa o número de nós pesquisados ao longo do processo, T é o tempo total de CPU em segundos e OBJ é o valor obtido para a função objectivo. Além disso, os parâmetros m e n representam, respectivamente, a dimensão do espaço e o número de pontos do problema

original, enquanto que NV e NRL representam, respectivamente, o número de variáveis e de restrições lineares do MPEC associado. Relacionados com a obtenção da última solução incumbente são ainda usados os parâmetros NIS, TS e NDS que representam, respectivamente, o número de iterações (operações pivotais), tempo CPU em segundos e o número de nós pesquisados até obtenção dessa solução.

Em todos os problemas teste é considerado para minorante do valor da função objectivo o valor zero, pois nesta função é minimizado o quadrado de uma expressão.

Os resultados apresentados na Tabela 5.1 mostram que quando o valor da função objectivo é igual ao minorante (0.000) o algoritmo obtém facilmente o mínimo global para o problema, mostrando-se pouco sensível ao incremento da dimensão do problema. Nos restantes casos a dificuldade na obtenção da solução óptima aumenta com a dimensão do problema. No problema 4 não se chega a ter a garantia de que a solução óptima tenha sido obtida, uma vez que foi atingido o limite máximo imposto de 10000000 operações pivotais.

A comparação do esforço computacional exigido para a obtenção da solução óptima e o total mostra que essa diferença não é muito significativa na maior parte dos exemplos testados.

De um modo geral os valores da função objectivo são iguais aos apresentados em [102]. Exceptuam-se o Problema 4, onde o valor da reformulação FDE é cerca de 83% melhor do que o obtido em [102], e o Problema 3 onde o valor obtido pela reformulação FDE é ligeiramente melhor do que o valor 0.1256 obtido em [102].

É ainda interessante notar que a incorporação de cortes disjuntivos no algoritmo de Ramificação e Limitação não trouxe qualquer melhoria no desempenho do processo. Por isso não é incluída nesta secção a experiência computacional com essa versão do método. Assim estes problemas confirmam a frase apresentada no capítulo anterior de que os cortes disjuntivos nem sempre são uma boa técnica para incorporar no processo de Limitação e Ramificação.

As Tabelas 5.2 e 5.3 incluem os resultados computacionais da resolução dos problemas de ajuste lineares com os algoritmos de Ramificação e Limitação (LBB) e Sequencial Completar baseado nas Restrições Activas (SLCP), respectivamente.

Em todos os problemas é considerado para minorante do valor da função objectivo o valor zero, pois neste problema é minimizada a soma de variáveis não negativas.

De um modo geral, as conclusões apresentadas para o problema quadrático são extensíveis ao problema linear. Com efeito os resultados da Tabela 5.2 indicam que quando o valor da função objectivo é igual ao minorante (0.000) rapidamente o processo encontra a solu-

PROB	NV	NRL	Ni	T	Nd	Nis	Ts	Nds	OBJ
1	130	60	119039	10.47	5879	1973	0.19	22	4.3627
2	160	70	438178	28.84	11673	1878	0.17	25	2.6287
3	190	80	2264946	189.04	54638	117415	10.01	2728	1.1231
4	220	90	>10000000	1242.56	242161	1267456	140.76	29178	0.1457
5	250	100	432	0.09	2	432	0.09	2	0.0000
6	280	110	193	0.06	1	193	0.06	1	0.0000
7	310	120	221	0.09	1	221	0.09	1	0.0000
8	340	130	233	0.13	1	233	0.13	1	0.0000
9	370	140	222	0.16	1	222	0.16	1	0.0000
10	400	150	255	0.19	1	255	0.19	1	0.0000

Tabela 5.2: Algoritmo de Ramificação e Limitação para o Problema de Ajuste Linear de Funções Dobradiças

ção ótima enquanto que nos quatro primeiros problemas teste a dificuldade na obtenção da solução ótima aumenta de uma forma proporcional à dimensão do problema. No problema 4 não se chega a ter a garantia de que o mínimo global foi obtido, uma vez que foi atingido o limite máximo imposto de 10000000 operações pivotais. Para estes primeiros problemas, o algoritmo determina eficazmente a solução ótima, tendo no entanto grande dificuldade em garantir a sua optimalidade. Tal como nos problemas anteriores, a incorporação de cortes disjuntivos não trouxe melhoria ao funcionamento do algoritmo, pelo que os resultados com essa versão do método não são apresentados.

PROB	NV	NRL	IT	Ni	T	Nd	Nis	Ts	Nds	OBJ
1	130	60	4	34188	4.38	3756	1794	0.24	122	4.3627
2	160	70	7	45481	7.95	4650	15619	2.67	1133	2.6287
3	190	80	8	154898	35.93	17200	4674	1.06	276	1.1231
4	220	90	15	3451188	2387.54	6631313	574	0.30	52	0.1457
5	250	100	8	714	0.31	54	714	0.31	54	0.0000
6	280	110	1	86	0.05	1	86	0.05	1	0.0000
7	310	120	1	122	0.06	1	122	0.06	1	0.0000
8	340	130	1	112	0.08	1	112	0.08	1	0.0000
9	370	140	1	131	0.11	1	131	0.11	1	0.0000
10	400	150	1	121	0.14	1	121	0.14	1	0.0000

Tabela 5.3: Algoritmo Sequencial Linear Complementar para o Problema de Ajuste Linear de Funções Dobradiças

A Tabela 5.3 apresenta os resultados da aplicação do algoritmo Sequencial Complementar aos MPECs lineares associados ao Problema de Ajuste Linear. Nessa tabela, as siglas Ni, T, ND, NIS, TS e NDS têm os significados anteriores, enquanto que IT representa o número de GLCPs(λ_k) necessários para obter o mínimo global do MPEC.

A análise desses resultados mostra que o SLCP resolve de uma forma bastante robusta todos os problemas, encontrando em geral a solução ótima rapidamente, num número de iterações e GLCPs relativamente pequeno. Apesar de, para os 4 primeiros problemas teste, o esforço computacional exigido pelo SLCP ser elevado, a maior parte desse esforço é realizado na verificação de que o último GLCP é inadmissível.

Comparando os resultados obtidos na Tabela 5.3 com os da Tabela 5.2, pode concluir-se que o SLCP tem um comportamento mais robusto e eficiente do que o Processo de Ramificação e Limitação para a resolução destes MPECs associados a Problemas de Ajuste Lineares. Além disso ambos os algoritmos foram capazes de obter os mínimos globais de todos os problemas.

Capítulo 6

PROBLEMA COMPLEMENTAR DE VALORES PRÓPRIOS

Neste capítulo é discutido o Problema Complementar de Valores Próprios (EiCP). Alguns resultados de existência relacionados com a classe da matriz do problema são primeiramente apresentados. A resolução do EiCP quando as suas matrizes são simétricas foi amplamente estudado em [101]. Algumas reformulações e métodos de resolução do EiCP para o caso não simétrico são discutidos. Em particular é estudada uma reformulação do EiCP como um MPEC e a aplicação do Método de Restrições Activas Modificado nesse caso. Um algoritmo de ramificação e limitação é também introduzido neste capítulo. Experiência computacional com os vários algoritmos é apresentada e atesta a validade dessas técnicas para a resolução do EiCP.

6.1 Definição do Problema Complementar de Valores Próprios

Este problema teve origem no estudo de estados de equilíbrio estático de sistemas mecânicos com atrito unilateral [24]. Para descrição do modelo, considere-se um sistema elástico não linear de dimensão finita sujeito a forças planares constantes, com atrito constante sobre um obstáculo achatado fixo. O deslocamento $u(t)$ e as reacções $r(t)$ satisfazem as condições de contacto unilateral de Signorini

$$u_N(t) \leq 0, \quad r_N(t) \leq 0, \quad u_N(t)^T r_N(t) = 0$$

e a lei de atrito de Coulomb

$$\|r_T(t)\| + \mu r_N(t) \leq 0, \quad r_T(t) \dot{u}_T(t) - \mu r_N(t) \|\dot{u}_T(t)\| = 0$$

onde $\mu > 0$ é o coeficiente de atrito, N e T são conjuntos de índices das componentes normais e tangenciais, respectivamente e $(\cdot)$ indica a derivada em ordem a t . Sob a aplicação de forças f^0 constantes, a evolução do sistema é governada pelas equações de equilíbrio de momento

$$M\ddot{u}(t) + Ku(t) = f^0 + r(t)$$

onde $(\ddot{\cdot})$ representa a segunda derivada em ordem a t e M e K são matrizes simétricas positivas definidas representando a massa e a rigidez, respectivamente. Um estado de equilíbrio do sistema [24] é caracterizado por um vector de deslocamentos u^0 e um vector de reacções r^0 que satisfazem

$$\left\{ \begin{array}{l} Ku^0 = f^0 + r^0 \\ u_N^0 \leq 0 \\ r_N^0 \leq 0 \\ (u_N^0)^T r_N^0 = 0 \\ \|r_T^0\| + \mu r_N^0 \leq 0. \end{array} \right.$$

Pode mostrar-se [24] que, para alguma vizinhança $[\tau, \tau + \Delta\tau[$, existem soluções dinâmicas da forma

$$u(t) = u^0 + \alpha(t)v$$

$$v(t) = v^0 + \beta(t)w,$$

se e só se existe uma solução do sistema

$$\left\{ \begin{array}{l} w = (\lambda^2 M^* + K^*)v \\ w_F = 0, \quad v_N = 0, \quad v_D = 0 \\ v_T \text{ sinal}(r_T^0) \leq 0 \\ w_T \text{ sinal}(r_T^0) + \mu w_N \leq 0 \\ (v_T \text{ sinal}(r_T^0))^T (w_T \text{ sinal}(r_T^0) + \mu w_N) = 0. \end{array} \right.$$

onde $\lambda \neq 0$, $v \neq 0$ e v, w são taxas de deslocamento e de reacção constantes, $\text{sin}(\mathbf{r}_T^0)$ é o vector com componentes ± 1 de acordo com o sinal das componentes de $\mathbf{r}_T^0$, F e D são os conjuntos correspondentes às partículas sem contacto e com taxa de deslocamento zero, respectivamente e M^* e K^* são matrizes obtidas por transformações lineares nas matrizes M e K , respectivamente.

Como é discutido em [101], este modelo pode ser reduzido ao seguinte Problema Complementar de Valores Próprios

Definição 6.1 *Dadas as matrizes $A, B \in \mathbb{R}^{n \times n}$, com B positiva definida, o Problema Complementar Generalizado de Valores Próprios (GEiCP_J) consiste em encontrar um escalar $\lambda > 0$ e um vector $x \in \mathbb{R}^n \setminus \{0\}$ tais que*

$$\begin{aligned} w &= (\lambda B - A)x \\ w_J &\geq 0, x_J \geq 0, x_J^T w_J = 0 \\ w_{\bar{J}} &= 0 \end{aligned} \tag{6.1}$$

onde $J \subseteq \{1, \dots, n\}$ e $\bar{J} = \{1, \dots, n\} \setminus J$.

Convém salientar que qualquer solução com $w = 0$ e as variáveis x_i sem restrição de sinal, corresponde a uma solução do Problema Generalizado de Valores Próprios (GEiP) [52] definido da seguinte forma:

Definição 6.2 *Dadas as matrizes $A, B \in \mathbb{R}^{n \times n}$, o Problema Generalizado de Valores Próprios (GEiP) consiste em encontrar um escalar $\lambda \in \mathbb{R}$, e um vector $x \in \mathbb{R}^n \setminus \{0\}$ tais que*

$$Ax = \lambda Bx$$

Um outro caso especial do GEiCP_J ocorre quando $J = \{1, \dots, n\}$ e constitui o chamado Problema Complementar de Valores Próprios (EiCP), que pode ser formulado da seguinte forma

Definição 6.3 Dadas as matrizes $A, B \in \mathbb{R}^{n \times n}$, com B positiva definida, o **Problema Complementar de Valores Próprios (EiCP)** consiste em encontrar um escalar $\lambda > 0$ e um vector $x \in \mathbb{R}^n \setminus \{0\}$ tais que

$$w = (\lambda B - A)x$$

$$w \geq 0, \quad x \geq 0$$

$$x^T w = 0$$

Para qualquer solução (λ, x) do GEiCP_J , o valor de λ é designado por *Valor Próprio Complementar* de (A, B) e x é um *Vector Próprio Complementar* correspondente. Uma vez que o conjunto dos vectores próprios complementares associados a um dado valor próprio é um cone, não existe perda de generalidade em considerar apenas as soluções que satisfazem $\|x\|_2 = p$, com $p > 0$. Estas restrições evitam que $x = 0$. No caso do EiCP a restrição $\|x\|_2 = p$ pode ser substituída pela restrição linear $\|x\|_1 = e^T x = p$, uma vez que $x \geq 0$. É fácil ver que qualquer solução do EiCP com $w = 0$ é um valor próprio positivo de (A, B) com um vector próprio correspondente satisfazendo a restrição de sinal. Em geral, verifica-se a seguinte propriedade:

Teorema 6.1 Para qualquer solução (λ, x) de GEiCP_J , existe um conjunto I satisfazendo $\bar{J} \subseteq I \subseteq \{1, 2, \dots, n\}$, tal que λ é um valor próprio positivo de (A_{II}, B_{II}) e x_I é um vector próprio correspondente satisfazendo $x_{J \cap I} \geq 0$.

Demonstração:

Seja $I = \{i : w_i = 0\}$. Então pela complementaridade $\bar{J} \subseteq I$ e $\bar{x}_{\bar{I}} = 0$. Portanto $A_{II}x_I = \lambda B_{II}x_I$. $\diamond$

Para o EiCP, este resultado significa que dada uma solução (λ, x) , λ é um valor próprio de (A_{II}, B_{II}) e x_I é um vector próprio correspondente não negativo. Como corolário deste teorema o número de soluções do EiCP e do GEiCP_J é finito. Além disso, em [101] é provado que o GEiCP_J tem no máximo $(n - |J| + 2)2^{|J|} - |J| - 2$ soluções. Como o EiCP é um caso particular do GEiCP_J com $J = \{1, 2, \dots, n\}$, facilmente se pode demonstrar que o seu número de soluções é no máximo $2^{n+1} - n - 2$.

O conjunto completo de soluções do GEiCP_J pode ser obtido através de uma enumeração completa. Embora esse método não seja prático, fornece uma condição necessária e suficiente para a admissibilidade de um GEiCP_J . No entanto, foi demonstrado em [101], que a

admissibilidade do GEiCP_J é um problema de decisão NP-completo, uma vez que qualquer solução (λ, x) do GEiCP_J tem que satisfazer $x_J \geq 0$ e $x^T A x > 0$, que por sua vez é um problema NP-completo [22]. Consequentemente, o GEiCP_J é de um modo geral um problema difícil de resolver. No entanto a classe da matriz A tem um papel preponderante na resolução do GEiCP_J . De facto, para algumas classes de matrizes, a admissibilidade do GEiCP_J pode ser estabelecida facilmente. Assim, verificam-se os seguintes resultados [101]:

Teorema 6.2 *Se A é uma matriz negativa semi-definida, o GEiCP_J é inadmissível para qualquer $J \subseteq \{1, \dots, n\}$.*

Teorema 6.3 *Se A é positiva ($a_{ij} > 0, \forall i, j$), e $B = I$, então o EiCP tem uma solução $\bar{\lambda} > 0$, com um vector próprio correspondente $\bar{x} > 0$. Além disso, se A é simétrica, então a solução $\bar{\lambda}$ é única.*

Teorema 6.4 *Se existe algum índice j tal que $a_{jj} > 0$ e $a_{ij} \leq 0, \forall i \neq j$, então (λ, x) é uma solução do EiCP , com $\lambda = a_{jj}$ e $x = e^j$, onde e^j é a j -ésima coluna da matriz identidade.*

6.2 O Problema Complementar Simétrico de Valores Próprios

Nesta secção consideramos o EiCP simétrico em que ambas as matrizes A e B são simétricas. Nesse caso o EiCP está estritamente relacionado com o Problema de Valores Próprios clássico. A condição de complementaridade $x^T w = 0$ pode ser reescrita como $x^T (\lambda B x - A x) = 0$, uma vez que $x \neq 0$ então

$$\lambda(x) = \frac{x^T A x}{x^T B x}$$

e $\lambda(x)$ é a função quociente generalizado de Rayleigh [52]. Uma vez que o gradiente desta função é

$$\nabla \lambda(x) = \frac{2}{x^T B x} [A - \lambda(x) B] x.$$

e $\nabla \lambda(x) = 0$ se e só se $[A - \lambda(x) B] x = 0$, então é de esperar que qualquer ponto estacionário $(x, \lambda(x))$ do quociente generalizado de Rayleigh no octante não negativo com $\lambda(x) > 0$ forneça

uma solução do EiCP. Com efeito se considerarmos o problema de otimização

$$\text{OEiCP} \quad \begin{cases} \text{maximizar} & \lambda(x) \\ \text{sujeito a} & x \geq 0 \\ & e^T x = p, \end{cases}$$

onde $e = (1, \dots, 1)^T$ e $p > 0$, então é possível estabelecer o seguinte resultado [101]:

Teorema 6.5

- x é um ponto estacionário de OEiCP com $\lambda(x) > 0$ se e só se $(\lambda = \lambda(x), x)$ é solução do EiCP.
- O EiCP tem solução se e só se existe $0 \neq x \geq 0$ tal que $x^T A x > 0$.
- O EiCP simétrico é NP-difícil.
- O EiCP tem no máximo $2^n - 1$ valores próprios.

Tendo em consideração este teorema, facilmente se consegue estabelecer o seguinte processo de resolução para o EiCP simétrico:

1. Encontre um vector $\bar{x}$ tal que $0 \neq \bar{x} \geq 0$, $\bar{x}^T A \bar{x} > 0$ e $e^T \bar{x} = p$.
2. Determine um ponto estacionário de

$$\begin{aligned} & \text{maximizar} && \lambda(x) \\ & \text{sujeito a} && x \geq 0 \\ & && e^T x = p \end{aligned}$$

utilizando $\bar{x}$ como solução inicial.

Convém salientar, que o último passo deste processo é bastante simples de efectuar, pois qualquer algoritmo local de Optimização Não Linear consegue obter um ponto estacionário, caso este exista. Esta simplicidade, normalmente, não é verificada no primeiro passo, uma vez que o problema subjacente é de um modo geral NP-completo [22]. Não obstante, na

Tabela 6.1 são apresentadas algumas classes de matrizes para as quais uma solução inicial para o OEiCP pode ser facilmente obtida [101]. Portanto o EiCP simétrico tem solução quando a matriz A pertence a essas classes. Em particular o EiCP simétrico tem solução se A é uma matriz PD .

CLASSE DE A	SOLUÇÃO INICIAL
$\exists i : a_{ii} > 0$	$x = e^i$
$A \geq 0, A \neq 0$	$x = e$
$\exists i, j : a_{ii} = 0, a_{jj} \leq 0 \text{ e } a_{ij} > 0$	$x_l = \begin{cases} 1, & l = j \\ \frac{1 - a_{jj}}{2a_{ij}}, & l = i \\ 0, & l \neq i, j \end{cases}$
$\exists x \geq 0 : Ax > 0$	x é solução do LP: $\begin{cases} \max & y \\ \text{s.a.} & Ax - ye \geq 0 \\ & e^T x = p \\ & x \geq 0 \\ & y \in \mathbb{R} \end{cases}$

Tabela 6.1: Solução inicial para a resolução do OEiCP

6.3 O Problema Complementar Não Simétrico de Valores Próprios

A expressão do gradiente do quociente generalizado de Rayleigh, apresentada na secção anterior só é válida quando A é simétrica. Se $A \neq A^T$ e B é simétrica PD , a expressão desse gradiente é:

$$\nabla \lambda(x) = \frac{1}{x^T B x} [A + A^T - 2\lambda(x)B] x,$$

e a relação entre os pontos estacionários e as soluções do EiCP não se verifica. Deste modo, é necessário pensar em outras metodologias para a resolução do EiCP não simétrico. Tal como

é comum em muitos problemas de otimização, uma dessas vias consiste em transformar o EiCP em problemas de otimização conhecidos. Seguidamente apresentamos duas dessas reformulações do EiCP.

(I) Problema Não Linear Complementar

Como os valores próprios são positivos, o EiCP pode ser formulado da seguinte forma:

$$\begin{aligned} \text{EiCP} \quad & w = Bx - x_{n+1}Ax \\ & e^T x = p \\ & x_{n+1} > 0 \\ & x \geq 0, w \geq 0, x^T w = 0 \end{aligned}$$

onde $x_{n+1} = \frac{1}{\lambda}$ e $B \in PD$. Se nesse problema for introduzida uma nova variável w_{n+1} não negativa tal que

$$w_{n+1} = -p + e^T x \quad \text{e} \quad w_{n+1}x_{n+1} = 0,$$

então obtém-se o Problema Complementar Não Linear

$$\begin{aligned} \text{NCP} \quad & w = Bx - x_{n+1}Ax \\ & w_{n+1} = -p + e^T x \\ & x \geq 0, w \geq 0, x_{n+1} \geq 0, w_{n+1} \geq 0 \\ & x^T w = x_{n+1}w_{n+1} = 0 \end{aligned} \tag{6.2}$$

Para verificar que o EiCP e o NCP são equivalentes, se $x_{n+1} = 0$ numa determinada solução do NCP, então

$$\begin{aligned} & w = Bx \\ & x \geq 0, w \geq 0 \\ & x^T w = 0 \\ & e^T x \geq p \end{aligned}$$

Isso é impossível por $B \in PD$ [26]. Então em qualquer solução do NCP tem-se $x_{n+1} > 0$ e $w_{n+1} = 0$. Donde

$$e^T x = p \quad \Rightarrow \quad x \neq 0$$

e

$$w = Bx - x_{n+1}Ax$$

$$w \geq 0, \quad x \geq 0$$

$$x^T w = 0$$

Portanto $(\lambda = \frac{1}{x_{n+1}}, x)$ é uma solução do EiCP. A implicação inversa é óbvia.

Como o NCP (6.2) não é monótono, os algoritmos de complementaridade [35] têm grande dificuldade em o resolver. O algoritmo PATH [16] incluído na colecção GAMS é seguramente o de maior aplicação entre esses processos. Na última secção deste capítulo é apresentada alguma experiência computacional que evidencia a dificuldade do algoritmo PATH em resolver o NCP e assim de processar o EiCP.

(II) Problema de Programação Matemática com Restrições de Equilíbrio

Com o intuito de reformular o EiCP num Problema de Programação Matemática com Restrições de Equilíbrio (MPEC) é introduzida uma nova variável não negativa y no EiCP tal que $y = x_{n+1}x$. Tendo em conta essa igualdade e multiplicando a condição $e^T x = p$ por x_{n+1} facilmente se conclui que qualquer solução do EiCP tem de satisfazer

$$\begin{aligned} w &= Bx - Ay \\ e^T x &= p \\ 0 &= e^T y - px_{n+1} \\ x &\geq 0, \quad w \geq 0, \quad y \geq 0, \quad x_{n+1} \geq 0 \\ x^T w &= 0 \\ \|y - x_{n+1}x\|_2 &= 0 \end{aligned}$$

Consideremos então o problema

$$\begin{array}{ll} \text{MPEC} & \text{Minimizar} \quad (y - x_{n+1}x)^T (y - x_{n+1}x) \\ & \text{sujeito a} \end{array} \quad \left[\begin{array}{l} w = Bx - Ay \\ e^T x = p \\ e^T y - px_{n+1} = 0 \\ x \geq 0, \quad w \geq 0, \quad y \geq 0, \quad x_{n+1} \geq 0 \\ x^T w = 0 \end{array} \right] \quad \text{GLCP} \quad (6.3)$$

É imediato que (λ, x) é solução do EiCP se e só se $(x_{n+1} = \frac{1}{\lambda}, x, y = x_{n+1}x)$ é solução do MPEC com valor ótimo igual a zero. Então é possível desenvolver um processo para a determinação de uma solução para o EiCP que consiste em:

1. Resolver o GLCP.
2. Aplicar um dos algoritmos descritos nos capítulos 3 ou 4 para encontrar um Ponto Estacionário ou um Mínimo Global para o MPEC.

Uma solução para o GLCP pode ser facilmente obtida se for tida em consideração a seguinte propriedade:

Teorema 6.6 *Se o EiCP tem solução então existe $0 \neq \bar{y} \geq 0$ tal que $(A\bar{y})_i > 0$, para um dado i .*

Demonstração:

Se $A \leq 0$, então o EiCP

$$w = Bx - A(x_{n+1}x)$$

$$x \geq 0, w \geq 0$$

$$x^T w = 0$$

$$x_{n+1} > 0, x \neq 0$$

não tem solução. Com efeito se $0 \neq x \geq 0$ e $x_{n+1} > 0$ então

$$A(x_{n+1}x) = x_{n+1}Ax \leq 0$$

Portanto

$$w \geq Bx \Rightarrow w^T x \geq x^T Bx > 0$$

e o EiCP não tem solução. Então A tem pelo menos um elemento $a_{rs} > 0$ e fazendo $\bar{y} = e^s$ tem-se

$$(A\bar{y})_r = \sum_{j=1}^n a_{rj}e_j^s = a_{rs} > 0.$$

◇

Seja então $\bar{y}$ tal que $0 \neq \bar{y} \geq 0$, $(A\bar{y})_i > 0$ para certo i e considere-se o $LCP(-A\bar{y}, B)$:

$$LCP \begin{cases} w = -A\bar{y} + Bx \\ x \geq 0, \quad w \geq 0 \\ x^T w = 0 \end{cases}$$

Como $B \in PD$, então o LCP tem uma única solução $(\bar{w}, \bar{x})$ [26]. Por outro lado, $0 \neq \bar{x} \geq 0$ pois se $\bar{x} = 0$, vem

$$\bar{w}_i = (-A\bar{y})_i < 0$$

Se agora considerarmos os vectores $\tilde{w}$, $\tilde{x}$, $\tilde{y}$ e $\tilde{x}_{n+1}$ tais que

$$\tilde{w} = \frac{p\bar{w}}{e^T \bar{x}}, \quad \tilde{x} = \frac{p\bar{x}}{e^T \bar{x}}, \quad \tilde{y} = \frac{p\bar{y}}{e^T \bar{x}}, \quad \tilde{x}_{n+1} = \frac{e^T \tilde{y}}{p}$$

então $(\tilde{x}, \tilde{y}, \tilde{w}, \tilde{x}_{n+1})$ é solução do GLCP.

Convém salientar que se B é a matriz identidade I , então o LCP é resolvido simplesmente por

$$\begin{cases} (A\bar{y})_i > 0 \Rightarrow \bar{x}_i = (A\bar{y})_i \\ (A\bar{y})_i \leq 0 \Rightarrow \bar{x}_i = 0 \end{cases} \quad \text{para todo } i = 1, \dots, n$$

Por outro lado, se $B \neq I$, então o LCP pode ser resolvido eficientemente por um dos métodos discutidos no capítulo 1.

Como a solução do GLCP é fácil de obter, então o algoritmo de restrições activas modificado ASETM pode ser usado para determinar um ponto estacionário $(\bar{x}, \bar{y}, \bar{\lambda})$ do MPEC associado ao EiCP. Se esse ponto estacionário satisfizer $\|\bar{x} - \bar{x}_{n+1}\bar{y}\|_2 = 0$, então $(\bar{\lambda} = \frac{1}{\bar{x}_{n+1}}, \bar{x})$ é solução do EiCP. Se tal não acontecer, então é necessário determinar um mínimo global do MPEC. Como a função objectivo desse MPEC não é convexa, não é possível aplicar os algoritmos de optimização global discutidos no capítulo 4. Na próxima secção é introduzido um algoritmo para obtenção desse mínimo global.

6.4 Uma Técnica de Reformulação e Linearização para o Problema Complementar de Valores Próprios

Nesta secção é estudada a resolução do EiCP a partir da chamada Técnica de Reformulação-Linearização (RLT) [109]. Nesta técnica são considerados dois passos, nomeadamente a Linearização e a Reformulação. Na fase de reformulação é definido um conjunto de factores variáveis não negativas e são formados produtos com esses factores e as restrições originais gerando várias restrições não lineares. Na fase de linearização é usada uma técnica apropriada de substituição de variáveis para linearizar estas restrições não lineares.

Consideremos a seguinte formulação do EiCP

$$\begin{aligned} \text{GOP} \quad & \text{Minimizar} \quad \sum_{i=1}^n w_i x_i \\ & \text{sujeito a} \quad w = -x_{n+1}Ax + Bx \end{aligned} \quad (6.4)$$

$$e^T x = p \quad (6.5)$$

$$x_{n+1} \geq 0, \quad x \geq 0, \quad w \geq 0 \quad (6.6)$$

Então como vimos na secção anterior, $(\bar{x}, \bar{w})$ é solução do *EiCP* se e só se

$$\min_{(x, w, x_{n+1}) \in S} w^T x = 0$$

com S o conjunto das restrições do GOP. Como é comum em algoritmos de optimização global, há vantagem em restringir a pesquisa da solução óptima do GOP para x_{n+1} pertencente a um intervalo fechado. Então iremos considerar o problema

$$\begin{aligned} \text{GOP} \quad & \text{Minimizar} \quad \sum_{i=1}^n w_i x_i \\ & \text{sujeito a} \quad w = -x_{n+1}Ax + Bx \end{aligned} \quad (6.7)$$

$$e^T x = p \quad (6.8)$$

$$\epsilon \leq x_{n+1} \leq \Delta \quad (6.9)$$

$$x \geq 0, \quad w \geq 0 \quad (6.10)$$

onde $\Delta > 0$ é suficientemente grande e $\epsilon > 0$ é muito pequeno.

O processo de resolução utilizado para resolver o GOP é um algoritmo de Ramificação e Limitação. Como uma solução admissível do GOP é solução de EiCP se e só se o valor da sua função objectivo é nulo, então um nó será podado sempre que o correspondente limite inferior for positivo. Além disso, sempre que for encontrada uma solução admissível do GOP com valor nulo da função objectivo, o algoritmo termina com uma solução para o EiCP. Para cada nó k da árvore binária dois tipos de ramificações podem ser considerados:

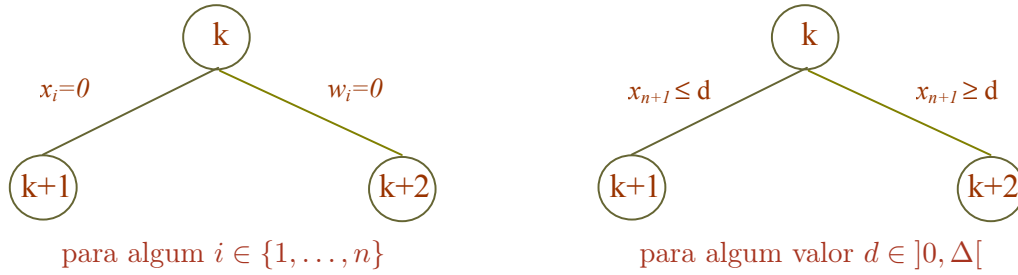


Figura 6.1: Esquema de ramificação do algoritmo

Deste modo, para qualquer subproblema associado a um nó, é necessário impor as restrições que foram adicionadas ao longo dos ramos da cadeia que liga esse nó ao nó de origem. Suponhamos que essas restrições são dadas por

$$\begin{aligned} l &\leq x_{n+1} \leq u \\ w_i &= 0 \quad \forall i \in I_0 \\ x_i &= 0 \quad \forall i \in J_0 \end{aligned}$$

onde $0 \leq l < u \leq \Delta$ e $I_0 \cap J_0 = \emptyset$. Consideremos o conjunto

$$K_0 = I_0 \cup J_0$$

e seja

$$\bar{I}_0 = \{1, \dots, n\} \setminus I_0, \quad \bar{J}_0 = \{1, \dots, n\} \setminus J_0, \quad \bar{K}_0 = \{1, \dots, n\} \setminus K_0$$

O subproblema no nó corrente k tem a forma

$$\begin{aligned}
\text{GOP}(k) \quad & \text{Minimizar} \quad \sum_{i \in \bar{K}_0} w_i x_i \\
& \text{sujeito a} \quad w = -x_{n+1}Ax + Bx
\end{aligned} \tag{6.11}$$

$$e^T x = p \tag{6.12}$$

$$l \leq x_{n+1} \leq u \tag{6.13}$$

$$(x, w) \geq 0 \tag{6.14}$$

$$w_i = 0, \quad \forall i \in I_0, \quad x_i = 0, \quad \forall i \in J_0 \tag{6.15}$$

A técnica RLT é usada para determinar um limite inferior nesse nó. Para isso efectuam-se algumas transformações ao problema de forma a obter um programa linear. Essas modificações são apresentadas a seguir.

- Na Fase de Reformulação as restrições

– (6.11) são multiplicadas por $x_i, \forall i \in \bar{J}_0$.

– (6.12) são multiplicadas por:

$$\begin{cases} x_{n+1} \\ x_i & \forall i \in \bar{J}_0 \\ x_{n+1}x_i & \forall i \in \bar{J}_0 \end{cases}$$

– (6.13) são multiplicadas por $x_i, \forall i \in \bar{J}_0$.

- Na Fase de Linearização são introduzidas as seguintes variáveis

– $z_i = w_i x_i, \quad \forall i \in \bar{K}_0 \quad (z_i = 0, \quad \forall i \in K_0)$.

– $y_i = x_{n+1} x_i, \quad \forall i \in \bar{J}_0 \quad (y_i = 0, \quad \forall i \in J_0)$.

– $v_{ij} = x_{n+1} x_i x_j, \quad \forall i, j \in \bar{J}_0, \quad i \leq j \quad (v_{ij} = 0 \text{ para todo } (i, j) \text{ com } i \in J_0 \text{ ou } j \in J_0)$.

– $s_{ij} = x_i x_j, \quad \forall i, j \in \bar{J}_0, \quad i \leq j \quad (s_{ij} = 0 \text{ para todo } (i, j) \text{ com } i \in J_0 \text{ ou } j \in J_0)$.

A aplicação da técnica RLT dá então origem ao seguinte programa linear

$$LP(k) \quad \text{Minimizar} \quad \sum_{i \in \bar{K}_0} z_i$$

sujeito a

$$w = -Ay + Bx \quad (6.16)$$

$$z_i = -\sum_{j=1}^n a_{ij}v_{ij} + \sum_{j=1}^n b_{ij}s_{ij} \quad \forall i \in \bar{J}_0 \quad (6.17)$$

$$e^T x = p \quad (6.18)$$

$$e^T y = px_{n+1} \quad (6.19)$$

$$\sum_{j=1}^n s_{ij} = px_i \quad \forall i \in \bar{J}_0 \quad (6.20)$$

$$\sum_{j=1}^n v_{ij} = py_i \quad \forall i \in \bar{J}_0 \quad (6.21)$$

$$lx_i \leq y_i \leq ux_i \quad \forall i \in \bar{J}_0 \quad (6.22)$$

$$z_i = 0, \quad \forall i \in K_0, \quad x_i = 0, \quad \forall i \in J_0, \quad w_i = 0, \quad \forall i \in I_0,$$

$$y_i = 0, \quad \forall i \in J_0, \quad v_{ij} = 0 \quad \text{se } i \in J_0 \text{ ou } j \in J_0,$$

$$s_{ij} = 0 \quad \text{se } i \in J_0 \text{ ou } j \in J_0 \quad (6.23)$$

$$\gamma = (z, w, x, x_{n+1}, y, v, s) \geq 0 \quad (6.24)$$

onde $z = (z_i) \in \mathbb{R}^n$, $y = (y_i) \in \mathbb{R}^n$, $v = (v_{ij}) \in \mathbb{R}^{n^2}$ e $s = (s_{ij}) \in \mathbb{R}^{n^2}$.

Notar que a dimensão deste problema é bastante elevada e pode ser reduzida se se tirar partido do facto de $s_{ij} = s_{ji}$ e $v_{ij} = v_{ji}$ para todos $i, j = 1, \dots, n$.

Seja $\bar{\gamma} = (\bar{z}, \bar{w}, \bar{x}, \bar{x}_{n+1}, \bar{y}, \bar{v}, \bar{s})$ uma solução do problema LP(k) com $\bar{\tau}$ o valor da função objectivo. Se $\bar{\tau} > 0$ então o nó k deve ser podado. De outro modo, sejam

$$\theta_1 = \max_{i \in \bar{K}_0} \{\bar{w}_i \bar{x}_i\}, \quad \theta_2 = \max_{i \in \bar{J}_0} |\bar{y}_i - \bar{x}_{n+1} \bar{x}_i|$$

O próximo resultado é consequência imediata da definição do problema GOP e fornece um critério de paragem para o algoritmo.

Teorema 6.7 *Se $\theta_1 = \theta_2 = 0$, então $(\lambda = \frac{1}{\bar{x}_{n+1}}, \bar{x})$ é solução do EiCP.*

Por outro lado, se $\theta = \max\{\theta_1, \theta_2\} \neq 0$ então a ramificação processa-se da seguinte forma:

- Se $\theta_1 > \theta_2$, ramifique na dicotomia $\{w_i = 0\} \vee \{x_i = 0\}$ para i tal que $\theta_1 = \bar{w}_i \bar{x}_i$.

- Caso contrário, $\theta_2 \geq \theta_1$ com $\theta_2 > 0$, ramifique da seguinte forma

$$l \leq x_{n+1} \leq \tilde{x}_{n+1} \quad \vee \quad \tilde{x}_{n+1} \leq x_{n+1} \leq u \quad (6.25)$$

onde

$$\tilde{x}_{n+1} = \begin{cases} \bar{x}_{n+1}, & \text{se } \min\{(\bar{x}_{n+1} - l), (u - \bar{x}_{n+1})\} \geq 0.1(u - l) \\ \frac{(u + l)}{2}, & \text{de outro modo.} \end{cases}$$

A convergência do algoritmo mencionado anteriormente para uma solução do EiCP é demonstrada no seguinte teorema.

Teorema 6.8 *Considere o algoritmo de ramificação e limitação com a regra de ramificação indicada. Então ou o processo termina num número finito de passos com uma solução do EiCP, incluindo possivelmente a indicação de que a solução não existe, ou então, uma árvore de ramificação e limitação infinita é gerada tal que ao longo de um ramo infinito dessa árvore, qualquer ponto de acumulação das soluções óptimas dos correspondentes $LP(k)$ fornece uma solução do EiCP.*

Demonstração:

O caso da convergência finita é óbvia. Suponhamos então que uma árvore infinita é gerada e consideremos um qualquer ramo infinito. Seja um ponto de acumulação das soluções óptimas dos $LP(k)$ ao longo desse ramo que produz os seguinte limites

$$\{\gamma^k\} \rightarrow \gamma^*, \quad [l_k, u_k] \rightarrow [l^*, u^*]$$

onde $l_k \leq x_{n+1}^k \leq u_k$, $k = 1, 2, \dots$ representam os limites em x_{n+1} . Notar que só se pode ter $\theta_1 > \theta_2$ num número finito de vezes, pois existe $k_1 \in \mathbb{N}$ tal que $k \geq k_1$ $w_i^k x_i^k = 0$ para todo $i = \dots, n$. Por este motivo no limite tem-se

$$w_i^* x_i^* = 0 \quad \forall i = 1, \dots, n \quad (6.26)$$

Além disso, uma vez que todos os nós abertos têm limite inferior nulo, tem-se $z^k = 0$ para todo o nó k aberto e o valor objectivo τ^k de $LP(k)$ é também zero para todo k . Então

$$\tau^* = 0, \quad \text{e} \quad z_i^* = 0 \quad \forall i = 1, \dots, n$$

Devido ao facto do ramo da árvore ser infinito e à forma como os intervalos $[l_k, u_k]$ são gerados, então

$$x_{n+1}^* = l^* \quad \vee \quad x_{n+1}^* = u^* \quad (6.27)$$

podendo ter-se $l^* = u^*$. Iremos apenas considerar que $x_{n+1}^* = l^*$ em (6.27), pois o caso de $x_{n+1}^* = u^*$ é semelhante.

Aplicando limites em (6.22), vem

$$y_i^* \geq l^* x_i^* \quad \forall i \quad (6.28)$$

Além disso, por (6.19)

$$\sum_{i=1}^n y_i^* = pl^*. \quad (6.29)$$

Se alguma inequação em (6.28) for satisfeita estritamente, então usando (6.28) e (6.18), vem

$$\sum_{i=1}^n y_i^* > l^* \sum_{i=1}^n x_i^* = pl^*,$$

o que contradiz (6.29). Donde

$$y_i^* = l^* x_i^* = x_{n+1}^* x_i^* \quad \forall i$$

Consequentemente, de (6.16), obtém-se

$$w^* = -Ay^* + Bx^* = -x_{n+1}^* Ax^* + Bx^* \quad (6.30)$$

Mas (6.18), (6.19) e (6.22) implicam que $l \leq x_{n+1}^* \leq u$ e também

$$0 < x_{n+1}^* \leq \Delta \quad (6.31)$$

Assim por (6.18), (6.24), (6.30) e (6.31), (w^*, x^*, x_{n+1}^*) é admissível para o GOP com $x_{n+1}^* > 0$. Mas por (6.26) o valor da função objectivo associado a essa solução admissível é nulo, e portanto $(\lambda = \frac{1}{x_{n+1}^*}, x^*)$ é solução do EiCP. $\diamond$

A demonstração apresentada mostra que as variáveis v_{ij} e s_{ij} e as restrições (6.17), (6.20) e (6.21) podem ser suprimidas do problema LP(k) sem prejudicar a convergência teórica do algoritmo de ramificação e limitação. A vantagem das suas utilizações tem a ver com a possibilidade de obter limites inferiores positivos, que permitem reduzir a pesquisa em árvore. Contudo, tem a óbvia desvantagem de aumentar drasticamente o número de restrições e variáveis dos programas LP(k) a resolver em cada nó k .

6.5 Escalonamento no Problema Complementar de Valores Próprios

O escalonamento consiste em tratar previamente o EiCP de modo a que todos os elementos das matrizes A e B tenham valores absolutos inferiores ou iguais à unidade. Considere-se

$$\begin{aligned} \text{EiCP} \quad & w = Bx - x_{n+1}Ax \\ & x \geq 0, \quad w \geq 0 \\ & x^T w = 0 \\ & e^T x = p \\ & x_{n+1} \geq 0 \end{aligned}$$

com $p > 0$, $B \in PD$ (simétrica ou não simétrica) de ordem $n \times n$, $A \in \mathbb{R}^{n \times n}$, $x, w \in \mathbb{R}^n$ e $e \in \mathbb{R}^n$ um vector de uns.

Sejam

$$\begin{cases} \beta = \max_{(i,j)} |b_{i,j}| > 0 \\ \alpha = \max_{(i,j)} |a_{i,j}| > 0 \end{cases}$$

Então

$$\begin{aligned} w &= Bx - x_{n+1}Ax \\ \Leftrightarrow w &= \left(\frac{1}{\beta}B\right)(\beta x) - \left(\frac{x_{n+1}\alpha}{\beta}\right)\left(\frac{1}{\alpha}A\right)(\beta x) \end{aligned}$$

Além disso:

$$\begin{aligned} x_{n+1} > 0 &\Leftrightarrow \frac{x_{n+1}\alpha}{\beta} > 0 \Leftrightarrow \frac{\beta x_{n+1}}{\alpha} > 0 \\ \beta x \geq 0 &\Leftrightarrow x \geq 0 \\ w^T x = 0 &\Leftrightarrow (\beta x)^T w = \beta (x^T w) = 0 \\ e^T x = p &\Leftrightarrow e^T (\beta x) = \beta (e^T x) = \beta p \end{aligned}$$

e

$$\begin{aligned} \forall_{(i,j)} \quad & \left| \frac{1}{\beta} b_{ij} \right| = \frac{1}{\beta} |b_{ij}| \leq 1, \quad \left| \frac{1}{\alpha} a_{ij} \right| = \frac{1}{\alpha} |a_{ij}| \leq 1 \\ \frac{1}{\beta} B &= \left(\frac{1}{\sqrt{\beta}} I_n \right) B \left(\frac{1}{\sqrt{\beta}} I_n \right) \in PD \text{ (PSD)} \Leftrightarrow B \in PD \text{ (PSD)} \end{aligned}$$

com I_n a matriz identidade de ordem $n \times n$. Notar também que

$$A \in PD \ (PSD) \Leftrightarrow \frac{1}{\alpha}A \in PD \ (PSD)$$

Portanto o EiCP é equivalente a

$$\begin{aligned} EiCP_{esc} \quad w &= \left(\frac{1}{\beta}B\right)x - x_{n+1} \left(\frac{1}{\alpha}A\right)x \\ x &\geq 0, \ w \geq 0 \\ x^T w &= 0 \\ e^T x &= \beta p \end{aligned}$$

e tem-se

$$\left\{ \begin{array}{l} \bar{x} = (\bar{x}_i) \in \mathbb{R}^n \\ \bar{x}_{n+1} > 0 \end{array} \right. \text{ é solução do } EiCP_{esc} \Leftrightarrow \left\{ \begin{array}{l} \tilde{x} = \left(\frac{\bar{x}_i}{\beta}\right) \in \mathbb{R}^n \\ \tilde{x}_{n+1} = \frac{\beta \bar{x}_{n+1}}{\alpha} > 0 \end{array} \right. \text{ é solução do EiCP}$$

6.6 Experiência Computacional

Nesta secção é apresentada alguma experiência computacional que atesta a eficiência dos algoritmos referidos neste capítulo para a resolução de EiCPs simétricos e não simétricos. Esta experiência computacional foi realizada num Pentium IV 2.4GHz com 256MB de RAM.

(I) EiCPs simétricos

Nesta subsecção iremos relatar a experiência computacional realizada com os seguintes algoritmos:

1. MINOS para a resolução do Programa Não Linear OEiCP.
2. PATH para a resolução do NCP (6.2).
3. ASETM para a determinação de um ponto estacionário do MPEC (6.3).

Nos problemas teste utilizados a matriz $A \in \mathbb{R}^{n \times n}$ é gerada aleatoriamente de forma a que os elementos sejam uniformemente distribuídos no intervalo $[0, 1]$ ou $[-50, 50]$, originando deste modo dois tipos de problemas teste, que designamos por S_1 e S_2 , respectivamente. A matriz B é a identidade. De forma a identificar os problemas teste, são utilizadas as letras s ou n

consoante a matriz A é simétrica ou não simétrica. O número apresentado a seguir a essas letras representa a ordem das matrizes A e B .

Na Tabela 6.2 é apresentado o comportamento dos algoritmos MINOS e ASETM na resolução dos EiCPs simétricos.

	Prob	MINOS				ASETM					
		$n \times m$	x_{n+1}	Ni	T	$n \times m$	x_{n+1}	Ni	T	OBJ	RES
$\mathcal{S}_I$	s200(*)	1×200	0.0100	1537	20.92	201×201	0.0100	397	0.31	2.38E-15	8.31E-08
	s100	1×100	0.0197	286	1.01	101×101	0.0197	197	0.03	1.73E-13	1.58E-06
	s50	1×50	0.0396	154	0.15	51×51	0.0396	98	0.02	1.79E-16	3.64E-08
	s40	1×40	0.0495	125	0.09	41×41	0.0495	78	0.00	7.06E-15	2.10E-07
	s30	1×30	0.0665	97	0.12	31×31	0.0665	59	0.00	4.76E-19	1.21E-09
	s20	1×20	0.0987	69	0.04	21×21	0.0987	39	0.02	2.51E-16	1.96E-08
	s10	1×10	0.2065	32	0.03	11×11	0.2065	20	0.00	3.01E-18	1.79E-09
	s6	1×6	0.3447	19	0.03	7×7	0.3447	11	0.00	6.88E-14	1.58E-07
	s3	1×3	0.7878	8	0.03	4×4	0.7878	6	0.00	2.63E-14	5.12E-08
$\mathcal{S}_S$	s200	1×200	0.0018	317	3.90	201×201	0.0019	2508	1.53	3.51E-08	2.35E-03
	s100	1×100	0.0024	163	0.05	101×101	0.0029	1023	0.19	4.57E-07	5.89E-03
	s50	1×50	0.0037	78	0.09	51×51	0.0037	195	0.02	8.34E-16	1.48E-07
	s40	1×40	0.0041	66	0.06	41×41	0.0045	219	0.03	2.03E-06	8.79E-03
	s30	1×30	0.0057	49	0.03	31×31	0.0068	117	0.02	5.07E-05	2.73E-02
	s20	1×20	0.0053	43	0.02	21×21	0.0053	57	0.02	3.73E-16	6.17E-08
	s10	1×10	0.0139	10	0.04	11×11	0.0157	27	0.03	9.05E-15	2.23E-07
	s6	1×6	0.0235	3	0.04	7×7	0.0235	2	0.03	6.65E-17	8.78E-09
	s3	1×3	0.0210	7	0.03	4×4	0.0210	6	0.02	9.99E-17	1.22E-08

Tabela 6.2: Resolução do EiCP simétrico com os algoritmos MINOS e ASETM

Nesta tabela, assim como nas restantes, $n \times m$ é a dimensão do problema resultante das respectivas reformulações, x_{n+1} é o inverso do valor próprio do EiCP, Ni é o número total de operações pivotais executadas, T é o tempo total de CPU em segundos necessário para resolver o problema, OBJ é o valor obtido para a função objectivo do MPEC e RES é a norma do resíduo definida por

$$\|w - Bx + x_{n+1}Ax\|_2.$$

Ao longo da experiência computacional realizada nesta subsecção, o parâmetro p na restrição

$e^T x = p$ tomou sempre o valor 1, com a exceção do problema s200 onde foi usado $p = 10$. A notação (*) é usada para indicar que o algoritmo MINOS não conseguiu obter uma solução para o EiCP com $p = 1$.

Na experiência computacional realizada com o algoritmo MINOS foi utilizada como solução inicial o vector $x = e^i$, coluna i da matriz identidade associada a um elemento diagonal positivo a_{ii} da matriz A .

Na experiência computacional com o algoritmo ASETM os problemas teste S_2 foram escalonados de acordo com o procedimento descrito na secção 6.5. Além disso, como $B = I$, a determinação de uma solução para o GLCP associado ao MPEC decorre do procedimento descrito a seguir ao teorema 6.6. Nesse processo $\bar{y}$ foi escolhido como o vector e^i correspondente a um elemento diagonal positivo a_{ii} de A .

Os resultados apresentados nesta tabela mostram que o algoritmo ASETM consegue obter uma solução para todos os EiCP simétricos S_1 com um esforço computacional significativamente inferior ao do programa comercial MINOS. No entanto, este comportamento inverte-se nos problemas teste S_2 , onde o MINOS é mais eficiente. Nesses problemas o processo ASETM encontra por vezes um ponto estacionário para o MPEC que não é uma solução para o EiCP. No entanto é interessante verificar que o resíduo é sempre muito pequeno.

Foi também testada a eficiência e robustez do método PATH na resolução do EiCP através da resolução do Problema Complementar Não Linear (6.2). Devido ao fraco desempenho deste algoritmo na resolução desses problemas, não são apresentados os respectivos resultados. No entanto, convém referir que o PATH não conseguiu resolver cerca de 63% dos problemas teste e, de um modo geral, quando consegue obter uma solução, o esforço computacional exigido é bastante superior ao do algoritmo ASETM.

Apesar do algoritmo ASETM não garantir uma solução do EiCP, o seu comportamento mostrou ser, de um modo geral, bastante eficiente e robusto quando comparado com metodologias baseadas em outras abordagens.

(II) EiCPs não simétricos

Nesta subsecção são comparados os desempenhos dos algoritmos PATH e ASETM para a resolução de EiCPs não simétricos. Recorde-se que PATH procura resolver o NCP (6.2) associado ao EiCP e ASETM determina um ponto estacionário do MPEC (6.3) equivalente

ao EiCP.

A matriz A foi gerada aleatoriamente de forma análoga à dos EiCPs simétricos. Para este caso foram também construídos dois tipos de problemas teste que designamos por N_1 e N_2 , respectivamente. Além disso B é novamente a matriz identidade.

Na Tabela 6.3 é apresentado o comportamento dos dois algoritmos na resolução desses problemas teste.

	Prob	n×m	PATH				ASETm				
			x_{n+1}	N1	T	Res	x_{n+1}	N1	T	Obj	Res
N_1	n200	201×201	****	521	4.63	1.00E+00	0.0010	397	0.33	1.70E-16	6.74E-08
	n100	101×101	0.0199	202	1.01	8.15E-07	0.0199	197	0.05	3.71E-14	6.46E-07
	n50	51×51	****	202	1.06	1.00E+00	0.0399	97	0.00	8.52E-13	2.07E-06
	n40	41×41	0.0493	133	0.75	1.62E-10	0.0493	78	0.00	1.75E-15	9.49E-08
	n30	31×31	0.0662	123	0.77	1.07E-08	0.0662	58	0.00	3.73E-17	9.18E-09
	n20	21×21	0.1006	114	0.78	5.18E-12	0.1006	40	0.00	2.91E-14	1.62E-07
	n10	11×11	0.1881	90	0.79	1.58E-07	0.1881	19	0.00	3.47E-14	1.86E-07
	n6	7×7	0.3268	90	0.79	3.04E-07	0.3268	11	0.00	1.13E-13	2.64E-07
	n3	4×4	****	122	1.12	1.00E+00	0.6182	7	0.00	1.38E-23	2.20E-12
N_2	n200	201×201	****	690	4.67	1.00E+00	0.0038	3173	2.05	7.84E-09	9.14E-04
	n100	101×101	****	234	1.39	1.00E+00	0.0047	974	0.19	9.34E-14	7.20E-06
	n50	51×51	****	130	1.12	1.00E+00	0.0077	303	0.03	2.56E-06	7.49E-03
	n40	41×41	****	383	1.08	1.00E+00	0.0064	231	0.00	1.32E-15	1.28E-07
	n30	31×31	****	141	1.09	1.00E+00	0.0125	178	0.02	9.04E-06	1.08E-02
	n20	21×21	****	130	1.10	1.00E+00	0.0112	66	0.02	2.13E-13	1.47E-06
	n10	11×11	****	130	1.12	1.00E+00	0.0110	21	0.02	1.78E-02	2.23E-07
	n6	7×7	****	517	1.98	1.00E+00	0.2233	5	0.03	1.78E-19	5.97E-10
	n3	4×4	****	119	1.13	1.00E+00	0.0258	3	0.05	6.91E-05	7.06E-03

Tabela 6.3: Resolução do EiCP não simétrico com os algoritmos PATH e ASETM

Nesta tabela, (****) é usada para indicar que o algoritmo PATH, com os parâmetros opcionais por defeito, não conseguiu obter uma solução para o EiCP. Este facto ocorreu em 67% dos problemas teste utilizados nesta experiência computacional.

Os resultados apresentados na Tabela 6.3 mostram que o algoritmo ASETM é capaz de resolver todos os problemas teste N_1 de uma forma mais eficiente e robusta do que o algoritmo PATH, uma vez que para este tipo de problemas o PATH só consegue obter solução para 75% dos problemas.

Este panorama agrava-se drasticamente para os problemas N_2 , pois o PATH não consegue resolver qualquer problema. O comportamento do ASETM é análogo ao observado para os S_2 só conseguindo obter solução para 65% dos casos. Convém ainda realçar que o esforço computacional exigido pelo ASETM na resolução do EiCP é muito reduzido quando comparado com o do PATH e que os valores dos resíduos das soluções obtidas são muito pequenos.

(III) Resolução de EiCPs simétricos e não simétricos com a Técnica RLT

Nesta subsecção são apresentados os resultados obtidos pelo algoritmo de ramificação e limitação baseado na técnica RLT (BBRLT), descrito na secção 6.4, na resolução dos problemas teste S_2 e N_2 . Não é relatada a experiência com os problemas S_1 e N_1 pois o algoritmo ASETM teve um comportamento bastante eficiente e robusto nesses casos.

Os dois tipos de problemas teste S_2 e N_2 foram escalonados de acordo com o procedimento descrito na secção 6.5.

A Tabela 6.4 inclui os resultados obtidos pelo algoritmo BBRLT na resolução dos problemas teste acima mencionados.

	Prob	$n \times m$	$[a, b]$	p	x_{n+1}	N_I	N_D	N_R	N_{SD}	T
S_2	s30	182×1051	$[0, 3]$	3	0.0054	31217	387	175	42	5.48
	s20	122×501	$[0, 0.5]$	3	0.0053	5483	162	77	11	0.80
	s10	62×151	$[0, 1]$	1	0.0135	335	39	13	16	0.14
	s6	38×67	$[0, 2]$	1	0.0290	396	87	20	27	0.17
	s3	20×25	$[0, 1]$	3	0.0210	38	43	3	19	0.11
N_2	n30	182×1051	$[0, 3]$	3	0.0461	58673	4791	2142	289	123.61
	n20	122×501	$[0, 1]$	3	0.0112	32502	767	443	78	5.53
	n10	62×151	$[0, 3]$	1	0.0423	6286	539	217	74	1.36
	n6	38×67	$[0, 3]$	1	0.2233	302	93	22	31	0.61
	n3	20×25	$[0, 3]$	3	0.0311	23	3	2	0	0.39

Tabela 6.4: Resolução do EiCP simétrico e não simétrico com o algoritmo BBRLT

Nesta tabela, para além dos parâmetros utilizados nas tabelas desta secção, $[a, b]$ indica o intervalo onde vai ser restringida a pesquisa da solução do EiCP, p é o valor utilizado nas restrições do problema $LP(k)$, N_D o número de nodos pesquisados, N_R o número de ramificações efectuadas e N_{SD} o número de subdivisões efectuadas.

A análise aos resultados da Tabela 6.4 indica que o algoritmo de ramificação e limitação é capaz de resolver eficientemente os problemas de dimensão pequena. No entanto o esforço computacional do algoritmo tende a aumentar drasticamente com o aumento da dimensão n do EiCP. Esses resultados são de certo modo esperados, tendo em conta as características da metodologia e o número elevado de restrições que é usado em cada nó para a determinação dos limites inferiores requeridos pelo algoritmo. A supressão de algumas dessas restrições e a combinação com o método ASETM poderá fornecer um processo mais eficiente para a resolução do EiCP. Essa deverá ser uma área de investigação futura.

Capítulo 7

OPTIMIZAÇÃO DE FORMAS ESTRUTURAIS

Neste capítulo é apresentado um modelo para a selecção de formas estruturais desenvolvidas em espaços bidimensionais. São consideradas as diversas condições associadas aos estados limites previstos na regulamentação de segurança aplicável. Os estados limites últimos são abordados com base na Teoria da Plasticidade, enquanto que os estados limites de utilização recorrem à Teoria da Elasticidade. O espaço das soluções admissíveis é discretizado em elementos finitos. A cada elemento está associada uma variável $0 - 1$, de forma a indicar se esse elemento está ou não contido na solução.

A formulação matemática do modelo conduz a um programa não linear inteiro misto com função objectivo bilinear e restrições bilineares e lineares. A sua reformulação como um problema de programação linear inteira mista é efectuada recorrendo à técnica de reformulação-linearização (RLT). Finalmente, é apresentada a redução deste problema a um MPEC e é investigada a resolução do problema correspondente usando os algoritmos discutidos em capítulos anteriores.

Alguma experiência computacional é incluída de forma a realçar a importância da formulação e da eficiência dos algoritmos para a resolução do MPEC equivalente ao programa inteiro correspondente.

7.1 Introdução à Optimização Estrutural

No projecto de qualquer obra de Engenharia Civil assume particular importância a fase da concepção da estrutura. Conhecidas as funções que a estrutura vai desempenhar e as condicionantes que a natureza da obra impõe, é possível definir um domínio num espaço bidimensional (2D) ou tridimensional (3D) onde a estrutura se pode desenvolver. Após a definição desse domínio, é possível enquadrar dentro dos seus limites uma infinidade de soluções. Para que uma solução se possa considerar admissível é necessário que a estrutura correspondente seja capaz de desempenhar as funções para que é concebida e de suportar as acções que possam ocorrer durante a sua construção e durante a sua vida útil.

Apesar do engenheiro poder projectar uma estrutura concebendo-a de acordo com a sua sensibilidade, um programa que lhe permita obter a forma óptima do esqueleto da estrutura que suporta as acções aplicadas pode tornar-se numa ferramenta bastante útil no desenvolvimento de um bom projecto. Esta tem sido uma área de intensa investigação em optimização [5, 9, 14, 54, 55, 56, 60, 75, 104, 105, 106, 113, 114, 116].

A maior parte dos trabalhos desenvolvidos nesta área usam estruturas reticuladas que, devido à sua simplicidade e apesar de não triviais, constituem a estrutura ideal para estudar algumas propriedades e características associadas às estruturas óptimas. A sua utilização pode ser observada em coberturas, pontes, torres de alta tensão entre outras obras. Algumas dessas aplicações podem ser observadas na Figura 7.1.



Figura 7.1: Estruturas Reticuladas

Algumas técnicas baseadas nos métodos de optimização clássica têm sido desenvolvidas para encontrar estruturas reticuladas óptimas. A maior parte destas técnicas pode ser classificadas em três categorias principais: Optimização de Secções Transversais, Optimização da Configuração e Optimização Topológica.

Na Optimização de Secções Transversais as áreas das secções transversais das barras são

consideradas como variáveis de projecto e as coordenadas dos nós e a matriz de ligação entre as várias barras são consideradas fixas, o que origina um problema de programação não linear [100].

Na Optimização da Configuração as coordenadas nodais são tidas como variáveis de projecto e movem-se durante o processo de optimização de forma a obter um esboço óptimo para a estrutura. Na maior parte dos estudos efectuados, a optimização de secções transversais e da configuração são usadas simultaneamente [106]. O problema resultante é mais complexo, obtendo-se também neste caso um programa não linear com a área da secção das barras e as coordenadas dos nós como variáveis.

Na Optimização Topológica são removidos elementos de uma estrutura inicial de modo a obter um subconjunto óptimo de elementos. O modelo estrutural vai variando ao longo do processo, o que torna o problema correspondente num dos mais aliciantes da optimização estrutural. Uma outra característica importante destes modelos é o facto de ser obrigatório que a matriz de ligação das barras da subestrutura resultante tenha característica completa (por linhas). Essa condicionante não é possível de ser representada analiticamente como uma restrição, o que torna o problema de optimização ainda mais complexo de resolver.

Alguns trabalhos têm sido desenvolvidos nas últimas quatro décadas na área do projecto estrutural óptimo, sendo a maior parte relacionados com a optimização de secções transversais e da forma da estrutura. Devido à sua complexidade poucas contribuições têm sido dedicadas à topologia óptima. No entanto, é reconhecida a sua importância em optimização estrutural, uma vez que, poderá permitir um ganho substancial de material e um melhoramento significativo no projecto.

A optimização topológica de estruturas discretas foi inicialmente atribuída a Dorn et al. (1964) [33], que aplicou um método de programação linear para otimizar a topologia das estruturas reticuladas. Desde então vários trabalhos têm surgido nesta área [105, 116]. Na maior parte destes estudos, a área da secção transversal de cada barra pode tomar o valor zero e deste modo a barra é eliminada da estrutura. Além disso, é assumido que estas variáveis são contínuas, de forma a tornarem mais fácil o processo de obtenção da solução. Como é mostrado em [32, 112] algumas dessas barras não podem ser eliminadas da estrutura.

Um outro grande obstáculo para estes estudos é a possibilidade da topologia óptima corresponder a um ponto singular do domínio admissível do projecto. Este facto é consequência da descontinuidade de algumas restrições quando a área da secção transversal é nula. A

singularidade da topologia óptima em estruturas reticuladas foi pela primeira vez mostrada em [112] e desde então tem sido objecto de intenso estudo [54, 55, 60, 75, 104].

A substituição de variáveis discretas $x_i \in \{0, 1\}$ por variáveis contínuas $0 \leq x_i \leq 1$ é demasiado forte. Assim para que esses modelos possam ser devidamente utilizados na prática, as variáveis associadas à área das secções transversais de cada barra devem ser consideradas como discretas. Apesar do problema associado ser NP-difícil, a dificuldade associada à descontinuidade de algumas restrições é completamente ultrapassada. Vários métodos têm sido desenvolvidos para a resolução deste tipo de problemas, alguns dos quais podem ser encontrados em [5, 9, 56, 113, 114]. Mais recentemente, Bollapragada et al. [14] apresentaram um modelo de optimização topológica e de secções transversais cuja formulação matemática resulta num programa linear inteiro misto [47].

Um outro aspecto que distingue os diferentes trabalhos realizados nesta área é o modo como são abordados os estados limites definidos pela regulamentação de segurança estrutural e que correspondem aos estados para além dos quais a estrutura deixa de satisfazer as exigências de projecto. São considerados os estados limites últimos e os estados limites de utilização. Os primeiros correspondem ao colapso ou a outras formas de rotura estrutural, enquanto que os últimos reportam-se a situações para além das quais os requisitos exigidos à estrutura em situações de serviço deixam de ser cumpridos. Os requisitos de utilização referem-se a deformações e deslocamentos excessivos, vibrações que provoquem desconforto ou prejuízo para a sua utilização ou ainda a danos que afectem o aspecto, a durabilidade ou o uso da estrutura.

Assim, enquanto que nos estados limites últimos os efeitos das acções de cálculo, esforços ou tensões instaladas não devem exceder as resistências do cálculo das estruturas, esforços máximos ou tensões máximas, nos estados limites de utilização os efeitos não devem exceder os critérios de bom desempenho, que se podem traduzir, de uma forma simplificada, pela limitação dos deslocamentos verificados a valores máximos admissíveis.

A optimização estrutural refere-se normalmente a modelos plásticos ou a modelos elásticos. No primeiro caso, os estados limites últimos são abordados com base na Teoria da Plasticidade e os estados limites de utilização não são considerados [111, 115]. Os modelos elásticos tratam os estados limites de utilização e os estados limites últimos através da Teoria de Elasticidade [14, 118]. Neste capítulo será apresentado um modelo designado por elastoplástico, que é baseado na Teoria da Elasticidade para os estados limites de utilização e na Teoria da

Plasticidade para os estados limites  ltimos. O modelo elastopl stico   considerado o mais apropriado pelos actuais c digos de seguran a [1].

Analogamente ao descrito em outros trabalhos [32, 33, 54, 55, 60, 75, 104, 105, 112, 116], o estudo desenvolvido neste cap tulo tem como base estruturas reticuladas articuladas. O dom nio das solu  es admiss veis   discretizado por um certo n mero de elementos finitos (barras). Qualquer solu  o admiss vel   caracterizada por um vector x , cujas componentes est o associadas a cada uma dessas barras, sendo iguais a 1 ou 0, conforme a correspondente barra est  ou n o contida na solu  o considerada. O crit rio de escolha da melhor solu  o corresponde ao do m nimo uso de material estrutural. A formula  o matem tica conduz a um programa n o linear inteiro misto. Utilizando a t cnica de reformula  o-lineariza  o (RLT) [109]   poss vel reduzir este problema a um programa linear inteiro misto 0 – 1.

Este problema de optimiza  o   transformado num MPEC, para o qual   proposto o algoritmo Sequencial Complementar descrito no cap tulo 4. Experi ncia computacional com v rias estruturas indica o interesse do modelo de optimiza  o estrutural e a efici ncia do algoritmo para a resolu  o do MPEC correspondente.

7.2 Um Modelo de Optimiza  o Topol gica

O meio estrutural admiss vel   referenciado num sistema cartesiano Oxy , no qual v rias solu  es alternativas para o problema em estudo podem ser descritas. Uma discretiza  o por elementos finitos [123] desse meio   ent o considerada, na qual a malha   composta por barras biarticuladas ligadas entre si nas suas extremidades (pontos nodais).

O meio estrutural   solicitado por diversas ac  es definidas na Regulamenta  o de Seguran a [1], como sejam o pr prio peso da estrutura, as sobrecargas de utiliza  o, as ac  es do vento, de sismos e outras. Estas ac  es podem conduzir a diferentes combina  es l , e cada uma delas   representada por for as nodais

$$f^l = \begin{Bmatrix} f_x^l \\ f_y^l \end{Bmatrix}$$

Algumas destas for as assumem o significado de reac  es r^l quando os pontos em causa est o ligados ao exterior. Em correspond ncia com as for as nodais podem definir-se os

deslocamentos nodais u

$$u^l = \begin{Bmatrix} u_x^l \\ u_y^l \end{Bmatrix}$$

O estado de tensão dentro de cada elemento finito (barra) i para a combinação de forças l pode ser descrito pelo seu esforço axial e_i^l , enquanto o estado de deformação se representa pela deformação axial d_i^l .

As condições fundamentais a impor nos estados limites de utilização são as relações de equilíbrio, as relações de compatibilidade, as condições fronteira e as relações constitutivas do material estrutural de acordo com a Teoria da Elasticidade. Devem ainda ser consideradas condições relacionadas com a limitação de deslocamentos.

As condições de equilíbrio são definidas nos nós da estrutura e relacionam os esforços elásticos instalados nas barras e_e^l com as reacções nos apoios r_e^l e com as forças nodais f^l aplicadas a partir de

$$C^T e_e^l - B r_e^l - f^l = 0 \quad (7.1)$$

onde C é a matriz dos cossenos directores que relacionam as direcções das barras com as direcções nodais e B é a matriz de incidência das direcções nos nós de apoio.

As condições de compatibilidade impõem deslocamentos iguais para todas as extremidades das barras que concorrem no mesmo nó e podem exprimir-se por

$$d^l = C u^l \quad (7.2)$$

onde d^l representa o vector das deformações das barras, u^l é o vector dos deslocamentos dos nós e C é a matriz de ligação referida anteriormente.

Os esforços e_e^l instalados nas barras da estrutura em situações de serviço estão relacionados com as deformações d^l das barras pelas relações constitutivas correspondentes a um comportamento do material que se considera elástico e linear. Este comportamento é regido pela Lei de Hook

$$e_e^l = K D_A d^l \quad (7.3)$$

onde $D_A = \text{diag} \{A_i\}$, com A_i uma variável discreta associada à área da secção transversal da barra i e $K = \text{diag} \{E_i h_i^{-1}\}$, com $E_i > 0$ o módulo de elasticidade da barra i e h_i o seu

comprimento.

De (7.1) , (7.2) e (7.3) resulta que

$$C^T K D_A C u^l - B r_e^l - f^l = 0 \quad (7.4)$$

As condi  es de liga  o da estrutura ao exterior s o dadas por

$$u_{jm}^l = 0 \quad (7.5)$$

para os pontos nodais ligados ao exterior com deslocamentos nulos. Os deslocamentos nodais n o devem ultrapassar os limites superiores e inferiores especificados na regulament  o de seguran a

$$u_{min_j} \leq u_j^l \leq u_{max_j} \quad (7.6)$$

Os estados limites  ltimos de resist ncia podem ser considerados com base na Teoria da Plasticidade. De acordo com o Teorema Est tico da Teoria da Plasticidade, as condi  es fundamentais a impor s o o equil brio e as condi  es de plasticidade, que respeitam as condi  es de apoio. As condi  es de equil brio podem exprimir-se por rela  es semelhantes a (7.1)

$$C^T e_p^l - B r_p^l - \lambda f^l = 0 \quad (7.7)$$

onde e_p^l representa o vector dos esfor os, designados por esfor os pl sticos, r_p^l o vector das reac  es pl sticas e λ   um coeficiente parcial de seguran a que majora as for as nodais que caracterizam as ac  es aplicadas e que   predefinido nos c digos estruturais de seguran a [1, 2].

As condi  es de plasticidade podem exprimir-se por

$$e_{min} \leq e_p^l \leq e_{max} \quad (7.8)$$

onde e_{min} e e_{max} s o os valores m nimo e m ximo admiss veis para os esfor os previstos na regulament  o [2].

As condi  es (7.4)-(7.8) s o satisfeitas para v rias solu  es onde algumas barras t m esfor o nulo. Um vector x adicional   introduzido no modelo, onde cada vari vel x_i est  associada   barra i e assume o valor 1 ou 0 consoante a barra i est  ou n o contida na solu  o. Assim,

o esforço instalado na barra i pode ser substituído pelo produto $x_i e_{p_i}^l$ conduzindo a esforços nulos em barras inexistentes. Esta condicionante é equivalente a considerar que os esforços devem satisfazer a seguinte condição

$$D_x e_{min} \leq e_p^l \leq D_x e_{max} \quad (7.9)$$

onde

$$D_x = \text{diag}(x_i) \quad (7.10)$$

Além disso, a matriz diagonal D_A é substituída por $D_A D_x$.

O objectivo do modelo é obter uma estrutura que utiliza o menor volume de material estrutural V possível. Se A_i for a área da secção transversal da barra i e h_i o seu comprimento, a função objectivo apresenta a seguinte forma

$$\text{Minimizar } V = \sum_i x_i A_i h_i \quad (7.11)$$

De acordo com a descrição das restrições e da função objectivo apresentadas nesta secção, o modelo pode ser formulado como o seguinte programa inteiro misto

$$\begin{aligned} \text{OPT} \quad & \text{Minimizar } V = \sum_{i=1}^{nb} x_i A_i h_i \\ & \text{sujeito a } C^T K D_A D_x d^l - B r_e^l - f^l = 0 \end{aligned} \quad (7.12)$$

$$d^l = C u^l \quad (7.13)$$

$$u_{min_j} \leq u_j^l \leq u_{max_j} \quad (7.14)$$

$$-C^T e_p^l + B r_p^l + \lambda f^l = 0 \quad (7.15)$$

$$D_x D_A t_{min} \leq e_p^l \leq D_x D_A t_{max} \quad (7.16)$$

$$u_{j_m}^l = 0 \quad (7.17)$$

$$A_i \in S_i = \{A_{i1}, \dots, A_{iN_i}\} \quad (7.18)$$

$$x_i \in \{0, 1\} \quad (7.19)$$

com $l \in \{1, \dots, nc\}$, $j \in \{1, \dots, 2nn\}$, $m \in \{1, \dots, na\}$, $j_m \in \{1, \dots, 2nn\}$ e $i \in \{1, \dots, nb\}$. Neste programa de optimização os parâmetros (ou dados) do problema têm os seguintes significados:

nb	número de barras;
na	número de apoios simples;
nn	número de nós;
nc	número de combinações de acções;
N_i	número de áreas discretas possíveis para a secção transversal da barra i ;
A_{ik_i}	k_i -ésima área possível para a barra i ;
S_i	conjunto discreto de áreas possíveis para a secção transversal da barra i ;
C	matriz $nb \times 2nn$ dos cossenos directores que relacionam as direcções das barras com as direcções nodais;
B	matriz $2nn \times na$ dos cossenos directores que relacionam as direcções dos apoios com as direcções nodais;
D_x	matriz diagonal, $[diag(x_i)]$;
D_A	matriz diagonal, $[diag(A_i)]$;
K	matriz diagonal, $diag\left(\frac{E_i}{h_i}\right)$;
E_i	modulo de elasticidade da barra i ;
h_i	comprimento da barra i ;
f_j^l	força nodal aplicada na direcção j para a combinação de forças l ;
λ	factor de segurança;
t_{min_i}	valor mínimo para a tensão de compressão da barra i , onde $t_{min_i} = \frac{e_{min_i}}{A_i}$;
t_{max_i}	valor máximo para a tensão de tracção da barra i , onde $t_{max_i} = \frac{e_{max_i}}{A_i}$;
u_{min_j}, u_{max_j}	valores mínimos e máximos admissíveis para o deslocamento nodal na direcção j .

As variáveis do problema e as suas definições são apresentadas a seguir:

A_i	área da secção transversal da barra i ;
x_i	variável que toma o valor 1 ou 0 conforme a barra i existe ou não;

d_i^l	deformação da barra i para a combinação de acções l ;
$e_{p_i}^l$	esforço da barra i para a combinação de acções l ;
$r_{p_m}^l$	reação plástica no apoio m para a combinação de acções l ;
$r_{e_m}^l$	reação elástica no apoio m para a combinação de acções l ;
u_j^l	deslocamento nodal na direcção j para a combinação de acções l .

Podemos assim concluir que o programa contém $nc \times (nb + 2nn + 2na) + 2nb$ variáveis e $nc \times (4nn + 2nb)$ restrições, algumas das quais bilineares. Além disso a função objectivo é também bilinear. É conveniente salientar que nesta formulação apenas se considera aplicada uma única força nodal. Contudo, a generalização para a aplicação de sistemas de forças mais complexos, contendo o peso próprio da estrutura e a acção sísmica, pode ser efectuada sem dificuldade.

7.3 Um Exemplo de Aplicação do Modelo

Nesta secção é introduzido um pequeno exemplo, que será um dos Problemas Teste (Pt0) utilizados na experiência computacional relatada na última secção deste capítulo.

Considere-se a estrutura reticulada bidimensional, apresentada na Figura 7.2 constituída por 6 barras, numeradas de 1 a 6 e por 4 nós, numerados de I a IV e cujas coordenadas (em metros) estão indicadas entre parêntesis junto a cada nó.

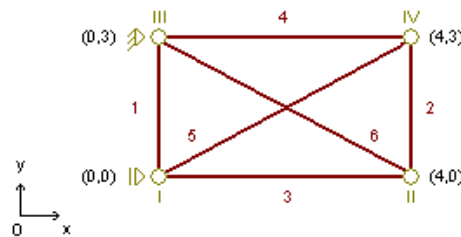


Figura 7.2: Exemplo de uma estrutura

Fazem parte da estrutura dois tipos de apoios, nomeadamente um apoio simples no nó I e um apoio duplo no nó III. O apoio no nó I é simples, uma vez que neste nó apenas está impedido o deslocamento numa das direcções Ox , podendo a estrutura deslocar-se paralelamente a Oy . O apoio no nó III é designado de duplo porque estão impedidos os seus deslocamentos em

qualquer direcção, isto é, o nó está fixo.

A estrutura está sujeita à acção de uma força exterior, f , aplicada no nó IV com a intensidade $65kN$, fazendo um ângulo de 0° com o semieixo Ox . Na Figura 7.3 pode observar-se a força exterior f a ser aplicada na estrutura, a numeração dos deslocamentos nodais, de u_1 a u_8 , assim como a numeração das reacções, de r_1 a r_3 , provocadas pelo facto estarem impedidos deslocamentos nos nós de apoio.

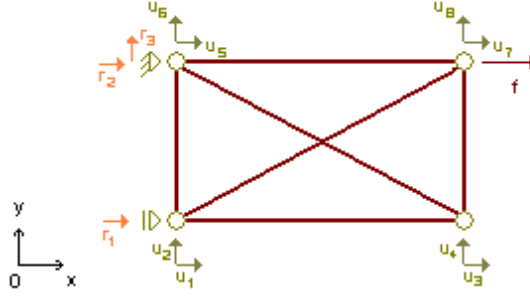


Figura 7.3: Exemplo (direcções nodais e reacções)

A partir das coordenadas dos nós extremos de cada barra, facilmente se calculam os comprimentos das barras e os seus cossenos directores (cossenos dos ângulos que cada barra faz com o semieixos positivos Ox e Oy). Isso permite calcular os elementos das matrizes C e B , que são dadas por

$$C = \begin{bmatrix} 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ -0.8 & -0.6 & 0 & 0 & 0 & 0 & 0.8 & 0.6 \\ 0 & 0 & 0.8 & -0.6 & -0.8 & 0.6 & 0 & 0 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

A matriz $KD_A = \text{diag}(k_i A_i)$ é definida da seguinte forma

$$k_i A_i = \frac{EA_i}{h_i}$$

onde A_i é a área da secção transversal de barra i ($A_i = \hat{A} = 3\text{cm}^2$), E é a constante de elasticidade ou módulo de Young ($E = 2.0 \times 10^5 \text{Mpa} = 2.0 \times 10^8 \text{kN/m}^2$) e h_i é o

comprimento da respectiva barra.

O vector força, acima mencionado, que actua na estrutura e é expresso em kN , é dado por

$$f^T = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 65 \ 0]$$

As variáveis do problema são as seguintes:

$x = (x_1, x_2, x_3, x_4, x_5, x_6)$	existência ou não da barra i ;
$e_p = (e_{p_1}, e_{p_2}, e_{p_3}, e_{p_4}, e_{p_5}, e_{p_6})$	esforços axiais das barras;
$d = (d_1, d_2, d_3, d_4, d_5, d_6)$	deformações das barras;
$r_p = (r_{p_1}, r_{p_2}, r_{p_3})$	reacções plásticas nos apoios;
$r_e = (r_{e_1}, r_{e_2}, r_{e_3})$	reacções elásticas nos apoios;
$u = (u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8)$	deslocamentos nodais;
$A = (A_1, A_2, A_3, A_4, A_5, A_6)$	áreas das secções transversais das barras.

Como na estrutura há apoios e por consequência existem deslocamentos que estão impedidos, as condições fronteira são traduzidas por:

$$u_1 = u_5 = u_6 = 0$$

Considera-se ainda que os limites dos deslocamentos são $u_{max} = -u_{min} = 50cm$ e que o coeficiente de segurança λ é igual a 1.5. Os valores extremos da resistência de cada barra são $e_{max} = D_A t_{max}$ e $e_{min} = D_A t_{min}$, onde t_{max} e t_{min} são os valores das tensões máximas e mínimas admitidas para o material, respectivamente. Como a estrutura considerada neste exemplo é constituída somente por um material (aço), para cada barra i resulta $t_{max_i} = -t_{min_i} = 355kPa$.

Uma vez que assumimos que a área da secção transversal de cada barra é igual a $\hat{A} > 0$, podemos considerar a versão mais simplificada do problema OPT:

$$\begin{aligned}
 \text{Minimizar} \quad & \hat{V} = \sum_{i=1}^n h_i x_i \\
 \text{sujeito a} \quad & C^T (K D_A D_x) d - B r_e - f = 0 \\
 & d = C u \\
 & u_{min} \leq u_j \leq u_{max} \quad j \in \{1, \dots, 8\} \\
 & -C^T e_p + B r_p + \lambda f = 0 \\
 & D_x e_{min} \leq e_p \leq D_x e_{max}
 \end{aligned}$$

$$u_1 = 0, \quad u_5 = 0, u_6 = 0$$

$$x_i \in \{0, 1\} \quad i \in \{1, \dots, 6\}$$

onde $\hat{V} = \frac{V}{A}$.

A estrutura da Figura 7.2 satisfaz as restrições de admissibilidade do problema anterior mas não é a óptima, uma vez que possui vínculos (barras e nós) superabundantes. Este tipo de estruturas designa-se de *Hiperstática*. Em contrapartida, a sua subestrutura constituída apenas pelos nós III e IV e pela barra 4 satisfaz as restrições de admissibilidade e é a solução óptima para o problema. No entanto, esta subestrutura não contém os vínculos suficientes



Figura 7.4: Exemplo duma estrutura hipostática

para garantir a sua total imobilidade, como pode ser constatado na Figura 7.4 sendo, por isso, designada por *Estrutura Hipostática*. A sua aplicação é limitada a estruturas particulares e a carregamentos especiais. É, por exemplo, o caso de estruturas com cabos ou tirantes, onde os esforços instalados para aquelas acções são sistematicamente positivos (de tracção). Na Figura 7.5 os tirantes correspondem ao conjunto de barras radiais, que ligam o tabuleiro da ponte ao mastro.

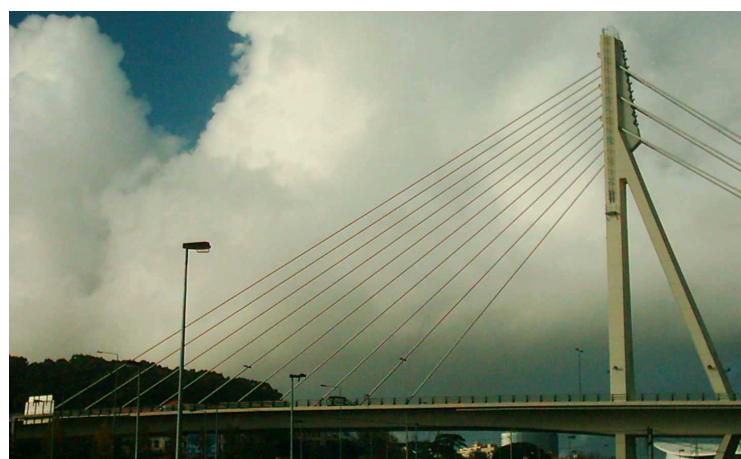


Figura 7.5: Ponte com tirantes

7.4 Estruturas Cinematicamente Instáveis

De um modo geral, a regulamentação de segurança aplicável às estruturas metálicas na área da Engenharia Civil [1, 2] não admite a utilização de estruturas hipostáticas ou cinematicamente instáveis.

Definição 7.1 *Uma estrutura hipostática é uma estrutura que pode apresentar deslocamentos sem que haja a deformação de qualquer barra individual.*

Nestas estruturas podem distinguir-se dois tipos de situações, nomeadamente os mecanismos e as estruturas críticas. Os primeiros possuem um conjunto de ligações que permitem deslocamentos finitos de nós da estrutura (ver deformada na Figura 7.6 a)). As estruturas críticas só admitem deslocamentos infinitesimais (ver deformada na Figura 7.6 b)).



Figura 7.6: Estruturas Hipostáticas: Mecanismo e Estrutura Crítica

O problema de optimização (OPT) descrito na secção 7.2 pode ter uma solução óptima que conduz a um destes tipos de estruturas. Como é descrito em [30], uma maneira de verificar a estabilidade cinemática de uma estrutura é verificar se a matriz de rigidez, que relaciona os deslocamentos com as forças exteriores, é positiva definida. Isto é equivalente a impor que:

a matriz $[-C^T B]$ associada à solução óptima obtida tem característica completa.

O programa matemático utilizado não permite incorporar esta restrição de forma directa. Seguidamente são apresentadas duas técnicas alternativas que implicam que a solução óptima do problema OPT satisfaça essa condição.

(I) Técnica da Perturbação (TP)

Esta técnica tem alguma correspondência com as tolerâncias aceites pelas disposições regulamentares na execução das estruturas de Engenharia Civil. De facto, a regulamentação

admite que as estruturas sejam construídas com erros de execução, desde que não ultrapassem os valores das tolerâncias fixadas na regulamentação. Sendo assim, a estrutura com tais erros de execução tem de ser capaz de suportar as acções previstas. A técnica consiste em analisar a estrutura resultante da perturbação das coordenadas dos seus nós. Na Figura 7.7 pode observar-se o que acontece à estrutura crítica apresentada na Figura 7.6 b), quando se perturbam as coordenadas dos nós com valores Δx e Δy .

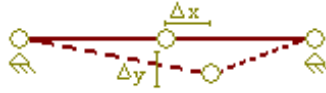


Figura 7.7: Estrutura crítica perturbada

Facilmente se conclui que a nova estrutura não é crítica. Além disso se Δx e Δy são aleatórios, a probabilidade de se obter uma nova situação crítica é nula. Se os vectores Δx e Δy forem

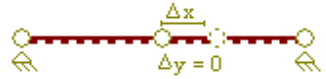


Figura 7.8: Estrutura crítica

gerados de forma aleatória, a matriz $[-C^T \ B]$ associada à malha inicial perturbada continua a ter característica completa. Além disso, é nula a probabilidade de existirem nessa nova malha duas direcções de barras linearmente dependentes, ou direcções de barras incidentes nos nós de aplicação das forças colineares com as direcções das forças.

Como o campo das forças internas (esforços das barras) $\{e_p\}$ pode ser visto como um caminho através do qual as forças aplicadas $\{f\}$ são transmitidas ao longo da estrutura até aos apoios onde são equilibradas pelas reacções $\{r\}$ [57], para que em cada nó haja equilíbrio, é necessário que existam pelo menos duas barras nas quais vão ser instaladas forças internas capazes de compensarem as forças que incidam nesse nó. Deste modo, a probabilidade de uma subestrutura não ter característica completa é nula.

Parece assim garantido que a técnica da perturbação conduz a submatrizes associadas a subestruturas que não são hipostáticas (críticas ou mecanismos). No entanto tais submatrizes poderão conduzir a soluções vizinhas de situações hipostáticas, ou seja, que a subestrutura tenda para uma estrutura hipostática quando o vector $\Delta = \begin{pmatrix} \Delta_x \\ \Delta_y \end{pmatrix}$ tende para zero. Se esse caso ocorrer, fazendo tender Δ para zero, o vector e_p (esforços axiais) possui valores que

tendem para infinito. Com efeito pelas condições de equilíbrio (7.61)

$$[-\tilde{C}^T \tilde{B}] \begin{bmatrix} e_p \\ r_p \end{bmatrix} = -f$$

Suponhamos agora que $[-\tilde{C}^T \tilde{B}]$ tende para uma matriz com característica não completa. Como $f \neq 0$ então para assegurar o equilíbrio terá de haver pelo menos uma componente de e_p a tender para ∞ . Como existem condições que limitam esses esforços, representados por e_p , tais situações não se verificam. Fica assim garantido teoricamente que a técnica das perturbações conduz a soluções que não correspondem a estruturas hipostáticas. No entanto, estas pequenas perturbações em conjugação com os erros de computação podem conduzir a resultados que nem sempre são satisfatórios. Deste modo, uma técnica alternativa mais eficiente foi também tida em consideração neste trabalho.

Convém realçar que a incorporação da técnica das perturbações no problema OPT em nada altera a sua formulação. A única diferença em relação ao modelo apresentado reside nos dados numéricos que constituem as matrizes do problema de optimização.

(II) Técnica da Estabilidade Cinemática (TEC)

Um critério simples para verificar se uma estrutura é ou não um mecanismo consiste na análise dos graus de liberdade da estrutura. Assim, seja

$$DOF = 2nn^* - nb^* - na^* \quad (7.20)$$

o grau de liberdade de uma estrutura com nn^* nós, nb^* barras e na^* apoios simples.

Definição 7.2 [48]

Diz-se que se uma estrutura não é um mecanismo, DOF é não positivo.

Com o objectivo de incorporar este critério no modelo, é introduzido um outro vector z , em que cada variável z_n associada ao nó n da malha inicial toma o valor 1 ou 0, consoante o nó está ou não está incluído na solução, isto é, se existem ou não barras incidentes nesse nó. Deste modo, a existência do nó $n \in \{1, \dots, nn\}$ pode ser formulada da seguinte forma

$$z_n \leq \sum_{i \in I(n)} x_i \leq |I(n)| z_n \quad (7.21)$$

onde $I(n) \subseteq \{1, \dots, nb\}$ é o conjunto dos índices i das barras que incidem no nó n e cujo cardinal $|I(n)|$ é sempre diferente de zero. Assim, se na solução não incidem barras no nó n , $\sum_{i \in I(n)} x_i = 0$ e por (7.21) $z_n = 0$. Caso contrário, $\sum_{i \in I(n)} x_i \geq 1$ e, por (7.21), $z_n = 1$. Associado a estas novas variáveis podemos introduzir o chamado Critério de Grubler [48]

$$2 \sum_{n=1}^{nn} z_n - \sum_{i=1}^{nb} x_i - \sum_{n=1}^{nn} s_n z_n \leq 0 \quad (7.22)$$

onde s_n é o número de apoios simples associados ao nó n . Este critério elimina uma grande parte das soluções associadas a mecanismos, diminuindo consideravelmente o conjunto admissível. Como não é uma condição suficiente para que a estrutura seja cinematicamente estável, tem de se verificar se tem característica completa (por linhas) a matriz C^* de ordem $2nn^* \times (nb^* + na^*)$, com $nb^* = \sum_{i=1}^{nb} x_i$, $nn^* = \sum_{n=1}^{nn} z_n$ e $na^* = \sum_{n=1}^{nn} s_n z_n$, obtida pela eliminação das colunas nulas da matriz $[-C^T D_x \ B]$ e das linhas e colunas associadas a $z_n = 0$. Assim, duas novas restrições semelhantes a (7.7) e (7.8) são incorporadas no modelo. A diferença consiste na carga nodal que é gerada aleatoriamente de modo a que a sua intensidade seja suficientemente pequena e só é aplicada nos nós existentes na solução. Deste modo, as novas restrições têm a seguinte forma:

$$-C^T e_a + Br_a + Zf_a = 0 \quad (7.23)$$

$$D_x e_{min} \leq e_a \leq D_x e_{max} \quad (7.24)$$

onde Z é uma matriz diagonal $2nn \times 2nn$, com z_{jj} igual a z_n do nó n associado à direcção j e f_a é o vector das cargas nodais perturbadas aplicadas em todas as direcções.

Seja f_a^* o vector obtido pela eliminação das componentes associadas a $z_n = 0$ do vector Zf_a . Se a característica da matriz C^* é menor do que $2nn^*$, a característica da matriz aumentada $[C^* \ -f_a^*]$ será superior em uma unidade à de C^* , pois f_a^* é gerada aleatoriamente. Deste modo, as restrições (7.23) seriam inadmissíveis. A probabilidade da matriz C^* ter linhas linearmente dependentes é então nula e a matriz C^* tem que ter característica completa (por linhas) no conjunto das soluções admissíveis. A utilização da técnica de estabilidade cinemática provoca uma modificação em algumas das restrições de programa OPT referido anteriormente. A nova formulação será notada por OPTEC e é apresentada a seguir.

$$OPTEC \quad \text{Minimizar} \quad V = \sum_{i=1}^{nb} x_i A_i h_i$$

$$\text{sujeito a} \quad M(D_A D_x) d^l - B r_e^l - f^l = 0, \quad l = 1, \dots, nc \quad (7.25)$$

$$d^l = C u^l, \quad l = 1, \dots, nc \quad (7.26)$$

$$u_{min_j} \leq u_j^l \leq u_{max_j}, \quad l = 1, \dots, nc, \quad j = 1, \dots, 2nn \quad (7.27)$$

$$u_{j_m}^l = 0, \quad l = 1, \dots, nc, \quad m = 1, \dots, na, \quad j_m \in \{1, \dots, 2nn\} \quad (7.28)$$

$$-C^T e_p^l + B r_p^l + \lambda f^l = 0, \quad l = 1, \dots, nc \quad (7.29)$$

$$D_x D_A t_{min} \leq e_p^l \leq D_x D_A t_{max}, \quad l = 1, \dots, nc \quad (7.30)$$

$$z_n \leq \sum_{i \in I(n)} x_i \leq |I(n)| z_n, \quad n = 1, \dots, nn \quad (7.31)$$

$$2 \sum_{n=1}^{nn} z_n - \sum_{i=1}^{nb} x_i - \sum_{n=1}^{nn} s_n z_n \leq 0 \quad (7.32)$$

$$-C^T e_a + B r_a + Z f_a = 0 \quad (7.33)$$

$$D_x D_A t_{min} \leq e_a \leq D_x D_A t_{max} \quad (7.34)$$

$$A_i \in S_i = \{A_{i1}, \dots, A_{iN_i}\}, \quad i = 1, \dots, nb \quad (7.35)$$

$$x_i \in \{0, 1\}, \quad i = 1, \dots, nb \quad (7.36)$$

$$z_n \in \{0, 1\}, \quad n = 1, \dots, nn \quad (7.37)$$

Neste novo programa há a acrescentar os seguintes parâmetros aos apresentados em OPT:

$I(n)$	conjuntos dos índices das barras que incidem no nó n ;
$ I(n) $	cardinal do conjunto $I(n)$;
s_n	número de apoios simples associados ao nó n ;
M	matriz $\left[C^T \text{diag} \left(\frac{E_i}{h_i} \right) \right]$;
Z	matriz $2nn \times 2nn$ diagonal com z_{jj} igual a z_n do nó n associado à direcção j ;
f_a	carga nodal perturbada aplicada a todas as direcções,

e as seguintes variáveis adicionais:

z_n	variável 0 – 1 que indica se o nó n existe ou não;
e_{a_i}	esforço da barra i para a carga nodal perturbada;
r_{a_n}	reacção plástica nos apoios n para a carga nodal perturbada.

Este problema não linear inteiro misto tem $nc \times (4nn + 2nb) + 4nn + 2nb + 1$ restrições e $nc \times (nb + 2nn + 2na) + 3nb + nn + na$ variáveis.

7.5 Formulação Linear Inteira Mista

As formulações descritas nas secções 7.2 e 7.4 contêm uma função objectivo bilinear nas variáveis x_i e A_i e algumas restrições lineares e bilineares. A existência destas funções bilineares é uma grande limitação à utilização de software comercial existente para a resolução de programas inteiros mistos. Nesta secção essas formulações são reduzidas a um programa linear inteiro misto usando a chamada técnica de reformulação-linearização (RLT) [109].

A técnica RLT é essencialmente constituída por dois passos, nomeadamente a reformulação e a linearização. Na fase de reformulação considera-se um conjunto de variáveis não negativas, baseadas nas restrições de limites, e efectua-se produtos destas variáveis com as restrições originais gerando deste modo restrições não lineares. Na fase de linearização é usada uma técnica apropriada de substituição de variáveis para linearizar estas restrições não lineares. De um modo geral, o problema não linear original e o problema linear resultante não têm as mesmas soluções óptimas. Contudo, existem determinados casos especiais nos quais estes problemas possuem exactamente a mesma solução óptima. Nesta secção é demonstrado que o programa OPT é exactamente um desses casos e que este tipo de técnica permite reduzir os programas não lineares inteiros mistos descritos nas secções anteriores a programas lineares inteiros mistos. Iremos mostrar primeiramente que o programa OPT é equivalente a um programa linear inteiro misto. Em seguida, uma extensão destes resultados ao programa OPTEC é apresentada sucintamente.

Para ser estabelecida a equivalência do programa OPT a um programa linear inteiro misto é necessário considerar que o vector d^l é limitado, isto é, que existem vectores constantes fixos $d_{\min}$ e $d_{\max}$ tais que

$$d_{\min} \leq d^l \leq d_{\max} \quad l = 1, \dots, nc \quad (7.38)$$

A existência desses limites é consequência das condições (7.13), (7.14) e (7.17).

Na fase de Reformulação as restrições (7.38) são multiplicadas por x_i e $(1 - x_i)$, $i = 1, \dots, nb$ e as restrições resultantes

$$\begin{aligned}
x_i d_{\min_i} &\leq x_i d_i^l \leq x_i d_{\max_i} \\
d_{\min_i}(1 - x_i) &\leq d_i^l - x_i d_i^l \leq d_{\max_i}(1 - x_i)
\end{aligned}$$

são incorporadas em OPT. Na fase de Linearização são introduzidas novas variáveis v_i^l que são definidas a partir de

$$v_i^l = x_i d_i^l \quad (7.39)$$

Usando estas duas fases da técnica *RLT*, o problema OPT pode ser formulado do seguinte modo:

$$\begin{aligned}
OPT2 \quad & \text{Minimizar} \quad V = \sum_{i=1}^{nb} x_i A_i h_i \\
& \text{sujeito a} \quad M(D_A)v^l - Br_e^l - f^l = 0
\end{aligned} \quad (7.40)$$

$$d^l = Cu^l \quad (7.41)$$

$$u_{\min_j} \leq u_j^l \leq u_{\max_j} \quad (7.42)$$

$$d_{\min_i} x_i \leq v_i^l \leq d_{\max_i} x_i \quad (7.43)$$

$$d_{\min_i}(1 - x_i) \leq d_i^l - v_i^l \leq d_{\max_i}(1 - x_i) \quad (7.44)$$

$$-C^T e_p^l + Br_p^l + \lambda f^l = 0 \quad (7.45)$$

$$D_x D_{At_{\min}} \leq e_p^l \leq D_x D_{At_{\max}} \quad (7.46)$$

$$u_{j_m}^l = 0 \quad (7.47)$$

$$A_i \in S_i = \{A_{i1}, \dots, A_{iN_i}\} \quad (7.48)$$

$$x_i \in \{0, 1\} \quad (7.49)$$

onde $l = 1, \dots, nc$, $j = 1, \dots, 2nn$, $i = 1, \dots, nb$, $m = 1, \dots, na$, $j_m \in \{1, \dots, 2nn\}$ e $M = C^T K$.

No programa OPT2, para cada barra i , a área da secção transversal A_i toma um valor discreto no conjunto S_i . Consoante os cardinais destes conjuntos, dois casos podem ocorrer que são apresentados a seguir.

- Se $|S_i| = 1 \quad \forall i$, então OPT2 é um programa linear inteiro misto;
- Se existe i tal que $|S_i| > 1$, então OPT2 é um programa não linear inteiro misto.

Portanto basta provar que neste último caso o programa OPT2 é equivalente a um programa linear inteiro misto. Da definição dos conjuntos S_i , podemos escrever

$$A_i x_i = \sum_{k_i=1}^{N_i} A_{ik_i} y_{ik_i} \quad (7.50)$$

onde $A_i \in \{A_{i1}, \dots, A_{iN_i}\}$ e y_{ik_i} são variáveis binárias tais que

$$\sum_{k_i=1}^{N_i} y_{ik_i} = x_i. \quad (7.51)$$

Como x_i é uma variável binária, então tem que se impor que

$$\sum_{k_i=1}^{N_i} y_{ik_i} \leq 1. \quad (7.52)$$

Por conseguinte para cada barra i , $A_i x_i$ é zero ou assume o valor de um só tamanho discreto no conjunto S_i . Além disso,

$$A_i v_i^l = \sum_{k_i=1}^{N_i} A_{ik_i} q_{ik_i}^l \quad (7.53)$$

onde $q_{ik_i}^l$ é a deformação correspondente à k_i -ésima área discreta possível da secção transversal da i -ésima barra para a l -ésima combinação de forças. Os $q_{ik_i}^l$ satisfazem as seguintes expressões

$$\sum_{k_i=1}^{N_i} q_{ik_i}^l = v_i^l \quad (7.54)$$

$$d_{\min_i} y_{ik_i} \leq q_{ik_i}^l \leq d_{\max_i} y_{ik_i} \quad (7.55)$$

Usando as expressões (7.50), (7.51), (7.53) e (7.54) para efectuar substituições e adicionando as restrições (7.52) e (7.55), obtém-se o seguinte programa linear inteiro misto

$$\begin{aligned} PLINT \quad & \text{Minimizar} \quad V = \sum_{i=1}^{nb} \left(\sum_{k_i=1}^{N_i} A_{ik_i} y_{ik_i} \right) h_i \\ & \text{sujeito a} \\ & \sum_{i=1}^{nb} M_{ji} \left(\sum_{k_i=1}^{N_i} A_{ik_i} q_{ik_i}^l \right) - \sum_{m=1}^{na} B_{jm} r_{e_m}^l - f_j^l = 0 \end{aligned} \quad (7.56)$$

$$d^l = C u^l \quad (7.57)$$

$$u_{\min} \leq u^l \leq u_{\max} \quad (7.58)$$

$$d_{\min_i} y_{ik_i} \leq q_{ik_i}^l \leq d_{\max_i} y_{ik_i} \quad (7.59)$$

$$d_{\min_i} \left(1 - \sum_{k_i=1}^{N_i} y_{ik_i} \right) \leq d_i^l - \sum_{k_i=1}^{N_i} q_{ik_i}^l \leq d_{\max_i} \left(1 - \sum_{k_i=1}^{N_i} y_{ik_i} \right) \quad (7.60)$$

$$-C^T e_p^l + B r_p^l + \lambda f^l = 0 \quad (7.61)$$

$$t_{\min_i} \sum_{k_i=1}^{N_i} A_{ik_i} y_{ik_i} \leq e_{p_i}^l \leq t_{\max_i} \sum_{k_i=1}^{N_i} A_{ik_i} y_{ik_i} \quad (7.62)$$

$$u_{j_m}^l = 0 \quad (7.63)$$

$$y_{ik_i} \in \{0, 1\} \quad (7.64)$$

$$\sum_{k_i=1}^{N_i} y_{ik_i} \leq 1 \quad (7.65)$$

onde $l = 1, \dots, nc$, $j = 1, \dots, 2nn$, $k_i = 1, \dots, N_i$, $i = 1, \dots, nb$, $m = 1, \dots, na$ e $j_m \in \{1, \dots, 2nn\}$.

Este programa possui

$$nc \times \left(4nn + 5nb + 2 \sum_{i=1}^{nb} N_i \right) + nb$$

restrições e

$$nc \times \left(2nb + 2nn + 2na + \sum_{i=1}^{nb} N_i \right) + \sum_{i=1}^{nb} N_i$$

variáveis. Esta dimensão é consideravelmente maior que a dimensão do problema OPT. Contudo, todas as funções envolvidas na sua definição são lineares e este facto possibilita o uso de um código de programação linear inteira para a sua resolução.

Seguidamente, provamos a equivalência entre os programas OPT e PLINT.

Teorema 7.1 $(\bar{x}, \bar{A}, \bar{e}_p^l, \bar{r}_p^l, \bar{d}^l, \bar{u}^l, \bar{r}_e^l)$ é uma solução óptima de OPT se e só se $(\bar{x}, \bar{A}, \bar{e}_p^l, \bar{r}_p^l, \bar{v}^l, \bar{d}^l, \bar{u}^l, \bar{r}_e^l)$ é uma solução óptima de OPT2, com $\bar{v}^l$ o vector de componentes $\bar{v}_i^l = \bar{x}_i \bar{d}_i^l$

Demonstração:

Como as funções objectivo dos dois problemas são iguais, então basta provar que há uma correspondência biunívoca entre as soluções admissíveis dos dois problemas. Se introduzirmos o vector $\bar{v}^l$, de componentes $\bar{v}_i^l = \bar{x}_i \bar{d}_i^l$, então da definição das variáveis x_i facilmente se conclui que $(\bar{x}, \bar{A}, \bar{e}_p^l, \bar{r}_p^l, \bar{d}^l, \bar{u}^l, \bar{r}_e^l)$ é uma solução admissível de OPT se e só se $(\bar{x}, \bar{A}, \bar{e}_p^l, \bar{r}_p^l, \bar{v}^l, \bar{d}^l, \bar{u}^l, \bar{r}_e^l)$ é admissível para OPT2. $\diamond$

Teorema 7.2 $(\bar{x}, \bar{A}, \bar{e}_p^l, \bar{r}_p^l, \bar{v}^l, \bar{d}^l, \bar{u}^l, \bar{r}_e^l)$ é uma solução óptima de OPT2 se e só se $(\bar{y}, \bar{e}_p^l, \bar{r}_p^l, \bar{q}^l, \bar{d}^l, \bar{u}^l, \bar{r}_e^l)$ é uma solução óptima de PLINT.

Demonstração:

Na formulação OPT2 cada variável A_i multiplica por x_i sempre que aparece. Assim, quando $x_i = 0$, A_i pode tomar qualquer valor em S_i sem alterar o valor da função objectivo. Deste modo, é suficiente demonstrar que existe uma bijecção entre as soluções admissíveis dos dois programas.

1. Seja $(\bar{x}, \bar{A}, \bar{e}_p^l, \bar{r}_p^l, \bar{v}^l, \bar{d}^l, \bar{u}^l, \bar{r}_e^l)$ uma solução admissível de OPT2, onde

$$\bar{A}_i = A_{ik_i^*}, \forall i : \bar{x}_i = 1. \quad (7.66)$$

Se as variáveis $\bar{y}_{ik_i}$ forem definidas da seguinte forma:

$$\sum_{k_i=1}^{N_i} \bar{y}_{ik_i} = \bar{x}_i \quad (7.67)$$

com $\bar{y}_{ik_i^*} = 1$ se $\bar{x}_i = 1$, então como $\bar{x}_i$ é uma variável binária, as restrições (7.65) são verificadas.

Além disso,

$$\sum_{k_i=1}^{N_i} A_{ik_i} \bar{y}_{ik_i} = \bar{A}_i \bar{x}_i \quad (7.68)$$

Consequentemente, as restrições (7.62) são satisfeitas.

Do mesmo modo, se

$$\sum_{k_i=1}^{N_i} \bar{q}_{ik_i}^l = \bar{v}_i^l \quad (7.69)$$

com

$$d_{\min_i} \bar{y}_{ik_i} \leq \bar{q}_{ik_i}^l \leq d_{\max_i} \bar{y}_{ik_i} \quad (7.70)$$

então (7.59) e (7.60) são verificadas. Para provar que $\sum_{k_i=1}^{N_i} A_{ik_i} \bar{q}_{ik_i}^l = \bar{A}_i \bar{v}_i^l$ tem-se:

- Se $\bar{x}_i = 0$ então por (7.43) $\bar{v}_i^l = 0$. Por outro lado, por (7.67) $\bar{y}_{ik_i} = 0$ para todo i e portanto por (7.70) $\bar{q}_{ik_i}^l = 0$ para todo i .
- Se $\bar{x}_i = 1$ então por (7.66) e (7.67) $\bar{y}_{ik_i^*} = 1$ e $\bar{y}_{ik_i} = 0$ para todo $k_i \neq k_i^*$. Portanto $\bar{q}_{ik_i}^l = 0$ para todo $k_i \neq k_i^*$ e por (7.69) $\bar{q}_{ik_i^*}^l = \bar{v}_i^l$.

Consequentemente, as restrições (7.56) são satisfeitas.

Assim $(\bar{y}, \bar{e}_p^l, \bar{r}_p^l, \bar{q}^l, \bar{d}^l, \bar{u}^l, \bar{r}_e^l)$ é uma solução admissível de PLINT onde $\sum_{k_i=1}^{N_i} \bar{y}_{ik_i} = \bar{x}_i$, com

$$\bar{y}_{ik_i^*} = 1 \text{ se } \bar{x}_i = 1 \text{ e } \sum_{k_i=1}^{N_i} \bar{q}_{ik_i}^l = \bar{v}_i^l, \text{ com } d_{\min_i} \bar{y}_{ik_i} \leq \bar{q}_{ik_i}^l \leq d_{\max_i} \bar{y}_{ik_i}.$$

2. Seja agora $(\bar{y}, \bar{e}_p^l, \bar{r}_p^l, \bar{q}^l, \bar{d}^l, \bar{u}^l, \bar{r}_e^l)$ uma solução admissível de PLINT. Se

$$\bar{x}_i = \sum_{k_i=1}^{N_i} \bar{y}_{ik_i} \quad (7.71)$$

então por (7.64) e (7.65) as restrições (7.49) são satisfeitas. Além disso, se $\bar{A}_i \in \{A_{i1}, \dots, A_{iN_i}\}$ então

$$\bar{A}_i \bar{x}_i = \sum_{k_i=1}^{N_i} A_{ik_i} \bar{y}_{ik_i} \quad (7.72)$$

Para demonstrar essa igualdade tem-se:

- Se $\bar{y}_{ik_i} = 0$ para todo i então por (7.71) $\bar{x}_i = 0$.
- Se existe k_i^* tal que $\bar{y}_{ik_i^*} = 1$ então por (7.71) $\bar{x}_i = 1$. Por outro lado, $\sum_{k_i=1}^{N_i} A_{ik_i} \bar{y}_{ik_i} = A_{ik_i^*}$ e $A_{ik_i^*} \in \{A_{i1}, \dots, A_{iN_i}\}$.

Consequentemente, as restrições (7.46) e (7.48) são satisfeitas.

Por outro lado, se

$$\bar{v}_i^l = \sum_{k_i=1}^{N_i} \bar{q}_{ik_i}^l \quad (7.73)$$

então as restrições (7.44) são verificadas, assim como, as restrições (7.43). Com efeito

- Se $\bar{y}_{ik_i} = 0 \forall k_i$ então por (7.59) $\bar{q}_{ik_i}^l = 0$ para todo k_i e por (7.73) $\bar{v}_i^l = 0$. Por outro lado, por (7.71) $\bar{x}_i = 0$.
- Se existe k_i^* tal que $\bar{y}_{ik_i^*} = 1$ então por (7.71) $\bar{x}_i = 1$. Por outro lado, por (7.59), $\bar{q}_{ik_i} = 0$ para todo $k_i \neq k_i^*$ e $d_{\min_i} \leq \bar{q}_{ik_i^*}^l \leq d_{\max_i}$. Portanto, por (7.73), $\bar{v}_i^l = \bar{q}_{ik_i^*}^l$.

Além disso, $\bar{A}_i \bar{v}_i^l = \sum_{k_i=1}^{N_i} A_{ik_i} \bar{q}_{ik_i}^l$ e as condições (7.40) são verificadas.

Assim $(\bar{x}, \bar{e}_p^l, \bar{r}_p^l, \bar{v}^l, \bar{d}^l, \bar{u}^l, \bar{r}_e^l)$ é uma solução admissível de OPT2, com

$$\bar{x}_i = \sum_{k_i=1}^{N_i} \bar{y}_{ik_i} \quad \text{e} \quad \bar{v}_i^l = \sum_{k_i=1}^{N_i} \bar{q}_{ik_i}^l.$$

Mas os valores das funções objectivo dos problemas OPT2 e PLINT são os mesmos, uma vez que $\sum_{k_i=1}^{N_i} A_{ik_i} \bar{y}_{ik_i} = \bar{A}_i \bar{x}_i$. Como existe uma correspondência biunívoca entre as soluções admissíveis destes dois problemas, então a solução óptima de um dos problemas corresponde à solução óptima do outro.

◇

Tendo em consideração o que foi mencionado nas secções anteriores relativamente à transformação do problema OPT num programa linear inteiro misto, podemos usar o mesmo tipo de desenvolvimento para mostrar que o programa OPTEC é equivalente ao seguinte programa linear inteiro misto

$$\begin{aligned}
 \text{PLINTEC} \quad & \text{Minimizar} \quad V = \sum_{i=1}^{nb} \left(\sum_{k_i=1}^{N_i} A_{ik_i} y_{ik_i} \right) h_i \\
 \text{sujeito a} \quad & \sum_{i=1}^{nb} M_{ji} \left(\sum_{k_i=1}^{N_i} A_{ik_i} q_{ik_i}^l \right) - \sum_{m=1}^{na} B_{jm} r_{e_m}^l - f_j^l = 0 \\
 & d^l = C u^l \\
 & u_{min} \leq u^l \leq u_{max} \\
 & d_{min_i} y_{ik_i} \leq q_{ik_i}^l \leq d_{max_i} y_{ik_i} \\
 & d_{min_i} \left(1 - \sum_{k_i=1}^{N_i} y_{ik_i} \right) \leq d_i^l - \sum_{k_i=1}^{N_i} q_{ik_i}^l \leq d_{max_i} \left(1 - \sum_{k_i=1}^{N_i} y_{ik_i} \right) \\
 & u_{j_m}^l = 0 \\
 & -C^T e_p^l + B r_p^l + \gamma_p f^l = 0 \\
 & t_{min_i} \sum_{k=1}^{N_i} A_{ik} y_{ik} \leq e_{p_i}^l \leq t_{max_i} \sum_{k=1}^{N_i} A_{ik} y_{ik} \\
 & z_n \leq \sum_{i \in I(n)} \sum_{k_i=1}^{N_i} y_{ik_i} \leq |I(n)| z_n \\
 & 2 \sum_{n=1}^{nn} z_n - \sum_{i=1}^{nb} \sum_{k=1}^{N_i} y_{ik} - \sum_{n=1}^{nn} s_n z_n \leq 0 \\
 & -C^T w_a + B r_a + Z f_a = 0 \\
 & t_{min_i} \sum_{k_i=1}^{N_i} A_{ik_i} y_{ik_i} \leq w_{a_i} \leq t_{max_i} \sum_{k_i=1}^{N_i} A_{ik_i} y_{ik_i} \\
 & y_{ik_i} \in \{0, 1\} \\
 & \sum_{k_i=1}^{N_i} y_{ik_i} \leq 1
 \end{aligned}$$

onde $l = 1, \dots, nc$, $j = 1, \dots, 2nn$, $k_i = 1, \dots, N_i$, $n = 1, \dots, nn$, $i = 1, \dots, nb$, $m = 1, \dots, na$ e $j_m \in \{1, \dots, 2nn\}$.

Este problema é um programa linear inteiro misto, com

$$nc \times \left(4nn + 5nb + 2 \sum_{i=1}^{nb} N_i \right) + 3nb + 4nn + 1$$

restrições e

$$nc \times \left(2nb + 2nn + 2na + \sum_{i=1}^{nb} N_i \right) + \sum_{i=1}^{nb} N_i + nn + nb + na$$

variáveis.

7.6 Formulação de Programação Matemática com Restrições de Complementaridade

Como foi referido na secção anterior, o modelo de optimização topológica descrito na secção 7.2 foi formulado como um programa linear inteiro misto, que pode ser apresentado na seguinte forma

$$\begin{aligned} PLI \quad & \text{Minimizar} \quad c^T x + d^T y \\ & \text{sujeito a} \quad Ax + By = q \\ & \quad \alpha \leq x \leq \beta, \quad y \in \{0, 1\}^m \end{aligned} \tag{7.74}$$

Nesta secção é introduzida uma reformulação deste problema a um MPEC recorrendo à Programação de Dois Níveis. Para isso é usado o seguinte resultado apresentado em [6].

Teorema 7.3

Se (x^*, y^*) é uma solução óptima do problema PLI, então $(x^*, y^*, v^* = 0)$ é uma solução óptima do problema de dois níveis

$$\begin{aligned} BL \quad & \text{Minimizar} \quad c^T x + d^T y \\ & \text{sujeito a} \quad Ax + By = q \\ & \quad 0 \leq y \leq e \\ & \quad \alpha \leq x \leq \beta \\ & \quad v = 0 \\ & \quad \text{Minimizar} \quad -e^T v \\ & \quad \text{sujeito a} \quad v \leq y \\ & \quad \quad v \leq e - y \\ & \quad \quad v \geq 0, \end{aligned}$$

onde e é um vector de componentes unitárias. Além disso, se $(x^*, y^*, v^* = 0)$ é uma solução óptima de BL, então (x^*, y^*) é uma solução óptima do problema PLI.

Tendo em conta razões computacionais, há a vantagem em substituir a restrição $v = 0$ por $e^T v = 0$. Tendo em consideração o exposto no capítulo 3, o programa de dois níveis (BL) pode ser transformado no seguinte MPEC equivalente.

$$\begin{aligned}
 \text{MPEC} \quad & \text{Minimizar} \quad c^T x + d^T y \\
 & \text{sujeito a} \quad Ax + By = q \\
 & 0 \leq y \leq e \\
 & \alpha \leq x \leq \beta \\
 & e^T v = 0 \\
 & \begin{bmatrix} \alpha \\ \beta \\ \pi \end{bmatrix} = \begin{bmatrix} 0 & 0 & -I \\ 0 & 0 & -I \\ I & I & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \\ v \end{bmatrix} + \begin{bmatrix} -I \\ I \\ 0 \end{bmatrix} y + \begin{bmatrix} e \\ 0 \\ -e \end{bmatrix} \\
 & \alpha^T \lambda = \beta^T \mu = \pi^T v = 0 \\
 & \alpha, \lambda, \beta, \mu, \pi, v \geq 0
 \end{aligned}$$

Convém salientar que a existência da restrição $e^T v = 0$ torna a resolução deste MPEC bastante complexa uma vez que, as matrizes E e M associadas às variáveis complementares não são quadradas. Por isso cada GLCP_{JL} (mesmo quando $J = L = \emptyset$), necessário à implementação do método de ramificação e limitação, tem de ser processado por um método enumerativo. Isso torna este processo pouco recomendado para resolver o MPEC neste caso. Contudo, o método Sequencial Complementar é baseado num método enumerativo e pode por isso ser utilizado para processar o MPEC. Na próxima secção são apresentados alguns resultados computacionais que atestam a validade desta proposta.

7.7 Experiência Computacional

Nesta secção é apresentada alguma experiência computacional de resolução de alguns problemas estruturais usando as formulações lineares inteira mista PLINT e PLINTEC e as suas reformulações MPEC. Estas experiências foram realizadas num Pentium IV 2.4GHz com 256

MB de RAM. Além disso foi utilizado o programa comercial de programação linear inteira OSL da colecção GAMS [16] para a resolução dos programas lineares inteiros mistos.

(I) Problemas Teste

A malha inicial de cada problema teste é constituída por pontos nodais (nós) e por barras, tendo uma forma semelhante à malha tipo representada na Figura 7.9.

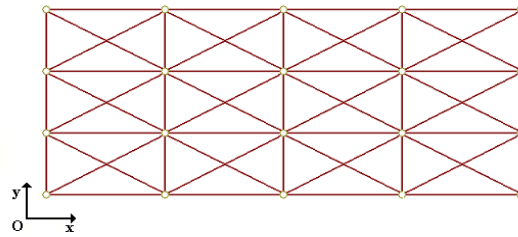


Figura 7.9: Malha inicial

O principal objectivo deste modelo é identificar o conjunto de barras incluídas na forma estrutural óptima, que se obtém a partir dos valores das variáveis x_i na solução óptima do problema.

Com o intuito de tirar conclusões sobre o comportamento das estruturas reticuladas foram tidos em consideração na constituição dos problemas teste diferentes tipos de tamanhos de malhas iniciais assim como de forças nodais aplicadas. Foram ainda considerados quatro tamanhos de malhas iniciais M_i , $i = 0, \dots, 3$, cujas topologias estão apresentadas na Tabela 7.1 e que conduzem a cinco problemas teste Pt0, ..., Pt4, de acordo com as seguintes definições:

- Pt0 - malha M0 e uma única força nodal aplicada ($f_{x_4}^1 = 65$, $f_{y_4}^1 = 0$).
- Pt1, Pt2 - malha M1 e dois tipos de forças nodais aplicadas. Em Pt1 é aplicada uma única força nodal ($f_{x_8}^1 = 0$, $f_{y_8}^1 = -65$), enquanto que em Pt2 são aplicadas duas forças nodais ($f_{x_8}^1 = 0$, $f_{y_8}^1 = -65$, $f_{x_9}^2 = -40$, $f_{y_9}^2 = -40$).
- Pt3 - malha M2 e duas forças nodais aplicadas em simultâneo ($f_{x_3}^1 = 45.9619$, $f_{y_3}^1 = -45.9619$, $f_{x_{12}}^1 = 45.9619$, $f_{y_{12}}^1 = 45.9619$).
- Pt4 - malha M3 e uma única força nodal aplicada ($f_{x_{23}}^1 = 0$, $f_{y_{23}}^1 = -65$).

Nessas definições são usados os seguintes parâmetros:

- $f_{x_n}^l$ força nodal em (kN) aplicada no nó n na direcção Ox para a combinação de forças l ;
- $f_{y_n}^l$ força nodal em (kN) aplicada no nó n na direcção Oy para a combinação de forças l .

	MALHA	h_x	h_y	nal	nb	nn	na	N_i	S_i
GRUPO I	M0	4	3	2×2	6	4	3	1	3
	M1	8	6	3×3	20	9	3	1	3
	M2	6	9	3×4	29	12	8	1	3
	M3	16	12	5×5	72	25	3	1	3
GRUPO II	SM1	8	6	3×3	20	9	3	2	0.5;3
	SM2	8	6	3×3	20	9	3	3	0.5;1;2
	SM3	6	9	3×4	29	12	8	2	0.5;3
	SM4	6	9	3×4	29	12	8	3	0.5;2;3

Tabela 7.1: Malhas dos problemas teste

Na Tabela 7.1 são incluídas as seguintes notações:

- nal dimensão da grelha em termos de número de pontos nodais no eixo Ox e no eixo Oy , respectivamente (na Figura 7.9, $nal = 5 \times 4$);
- h_x comprimento total (em m) no eixo Ox ;
- h_y comprimento total (em m) no eixo Oy ;
- nb número de barras;
- nn número de nós;
- na número de apoios simples;
- S_i conjunto discreto de áreas possíveis para a barra i (em cm^2);
- N_i número de áreas discretas possíveis para a secção transversal da barra i .

No primeiro grupo de problemas teste foram consideradas estruturas para as quais só existe um valor discreto disponível para a área da secção transversal de cada barra, enquanto que no segundo grupo é permitido que cada barra da estrutura assuma um dos valores de um conjunto finito de valores discretos possíveis para a área da sua secção transversal. Este último grupo conduz a quatro problemas teste adicionais, designados por St1, St2, St3 e St4, e cujas malhas iniciais associadas são SM1, SM2, SM3 e SM4, respectivamente. As malhas

SM1 e SM2 têm as mesmas dimensões da malha M1, enquanto que as SM3 e SM4 têm as mesmas dimensões das de M2. As forças nodais aplicadas em ST1 e ST2 são as mesmas que em PT1, enquanto que nos dois problemas restantes são as mesmas que em PT3. O número de restrições (nr) e o número de variáveis (nv) das formulações PLINT e PLINTEC associados a estes problemas teste são apresentados na Tabela 7.2.

		PLINT		PLINTEC	
	PROB	nr	nv	nr	nv
GRUPO I	Pt0	64	38	93	51
	Pt1	196	104	273	136
	Pt2	372	188	449	220
	Pt3	280	156	387	205
	Pt4	676	344	921	444
GRUPO II	St1	236	144	313	176
	St2	276	184	353	216
	St3	338	214	445	263
	St4	396	272	503	321

Tabela 7.2: Dimensões dos problemas teste

Em todos os problemas teste os limites considerados para os deslocamentos e tensões nas barras são $u_{max_j} = -u_{min_j} = 50cm$ e $t_{max_i} = -t_{min_i} = 355MPa$, respectivamente e o factor de segurança λ é igual a 1.5.

(II) Resultados Computacionais com as Formulações Inteiras

A Tabela 7.3 apresenta os resultados correspondentes ao desempenho do código OSL para encontrar a solução óptima de cada um dos problemas teste, introduzidos anteriormente, e associados às formulações PLINT, com a técnica de perturbação tida em consideração e PLINTEC. Esse comportamento é avaliado em termos do número de operações pivotais (NI), nós (Nd) e tempo CPU em segundos (T). Nesta tabela, a notação "> 25000000" é usada para indicar que o programa não conseguiu encontrar a solução óptima do problema teste após 25000000 iterações (operações pivotais). O valor óptimo da função objectivo (OBJ), em dm^3 , obtido pelo código é também incluído. No caso de insucesso (NI > 25000000), OBJ corresponde ao valor da solução incumbente, que pode não ser a solução óptima do problema em causa.

PROB	TP			TEC			OBJ (dm ³)
	NI	T	ND	NI	T	ND	
PT0	35	0.06	13	53	0.04	7	3.60
PT1	5464	1.72	1271	3033	0.69	311	10.80
PT2	13474	4.78	1709	5579	1.77	497	12.90
PT3	362625	183.86	64909	891143	325.64	82075	11.92
PT4	***	***	***	>25000000	15018.78	347541	27.30
ST1	4202	1.43	835	64943	22.80	8132	7.05
ST2	26683	8.22	3011	57473	30.01	10052	(5.05)4.90
ST3	430978	215.14	72962	4788682	3996.54	411084	6.29
ST4	2110439	1284.79	267549	20606789	61486.08	1496081	5.46

Tabela 7.3: Resultados numéricos com as técnicas TP e TEC

(ST2: OBJ_{TP}: 5.05 e OBJ_{TEC} : 4.90)

Os resultados apresentados na Tabela 7.3 mostram que, para malhas de pequena e média dimensão, o código OSL de programação inteira é capaz de encontrar uma solução óptima que conduz a formas estruturais que contêm um número pequeno de nós e barras. No entanto, para malhas de dimensões superiores, de um modo geral, não se tem a garantia que a solução obtida seja a óptima, apesar do valor da função objectivo ser próximo do óptimo.

A utilização da técnica de perturbação das coordenadas dos nós enfrenta alguns problemas. Para estruturas de média e grande dimensões os erros de computação e as perturbações efectuadas podem confundir-se e dificultar a resolução de alguns problemas. É o caso do problema PT4 para o qual o programa OSL não consegue obter uma solução óptima devido à ocorrência de um erro na sua execução (***) . A análise da solução incumbente mostra que esta corresponde a uma estrutura cinematicamente instável.

Noutras situações, se a solução óptima associada à malha inicial não perturbada corresponde a uma estrutura cujos valores de algumas variáveis estão muito próximos dos limites, as perturbações efectuadas vão originar pequenas alterações nos valores das variáveis que podem transformar uma solução admissível em inadmissível. É o caso do problema ST2 para o qual a estrutura óptima, associada ao valor $4.90dm^3$, se torna inadmissível quando se perturba a malha inicial, pois os esforços axiais instalados em algumas dessas barras ultrapassam os limites impostos. Deste modo, é obtida a solução final, com valor $5.05dm^3$, que não é óptima. Na tentativa de eliminar estas dificuldades efectuou-se uma nova reformulação do problema

PLINT, na qual as matrizes associadas ao problema não eram perturbadas mas sim incorporadas duas novas restrições do tipo (7.61) e (7.62), com a matriz C^T correspondente às perturbações nas coordenadas dos nós e com $\lambda = 1$. Com este modelo que tem dimensão superior à de PLINT, e inferior à de PLINTEC, conseguiu-se resolver o problema ST2 mas mantiveram-se as dificuldades ocorridas na resolução de PT4.

A técnica de Estabilidade Cinemática mostrou ser mais robusta do que a de Perturbação, uma vez que o problema de optimização OPTEC fornece sempre uma estrutura cinematicamente estável. Além disso na maior parte dos casos ($N_I < 25000000$), a solução obtida é a óptima. No entanto, a resolução dos problemas por esta técnica, de um modo geral, requer um maior esforço computacional do que pela técnica de perturbação. Na Figura 7.10 estão representadas a malha inicial M2 e as formas óptimas e deformadas obtidas pela técnica TEC para os problemas PT3, ST3 e ST4 associados a essa malha. Entre parêntesis indicam-se os esforços instalados nas barras assim como as reacções quer elásticas quer plásticas nos nós de apoio.

(III) Resultados Computacionais com a Formulação MPEC

Nesta subsecção, é relatada a experiência computacional realizada com os algoritmos de restrições activas modificado ASETM e o Sequencial Complementar SLCPA na resolução dos MPECs correspondentes à formulação MPEC do programa linear inteiro PLINTEC.

As dimensões dos problemas resultantes podem ser observadas na Tabela 7.4, onde nr , nv e nvc representam, respectivamente, o número de restrições, de variáveis e de pares de variáveis complementares.

É notório um aumento significativo das dimensões dos problemas MPEC relativamente às dimensões dos problemas correspondentes PLINTEC, apresentadas na Tabela 7.2.

A primeira experiência computacional foi realizada com o algoritmo ASETM com o intuito de se obter um ponto estacionário para os MPECs associados aos problemas de optimização estrutural e analisar se essa solução está próxima ou não da solução óptima. Na Tabela 7.5 é apresentado o comportamento desse algoritmo na resolução dos problemas teste. Nessa tabela, assim como nas restantes, N_I é o número total de operações pivotais executadas e T é o tempo total de CPU em segundos necessário para resolver o MPEC, OBJ é o valor obtido

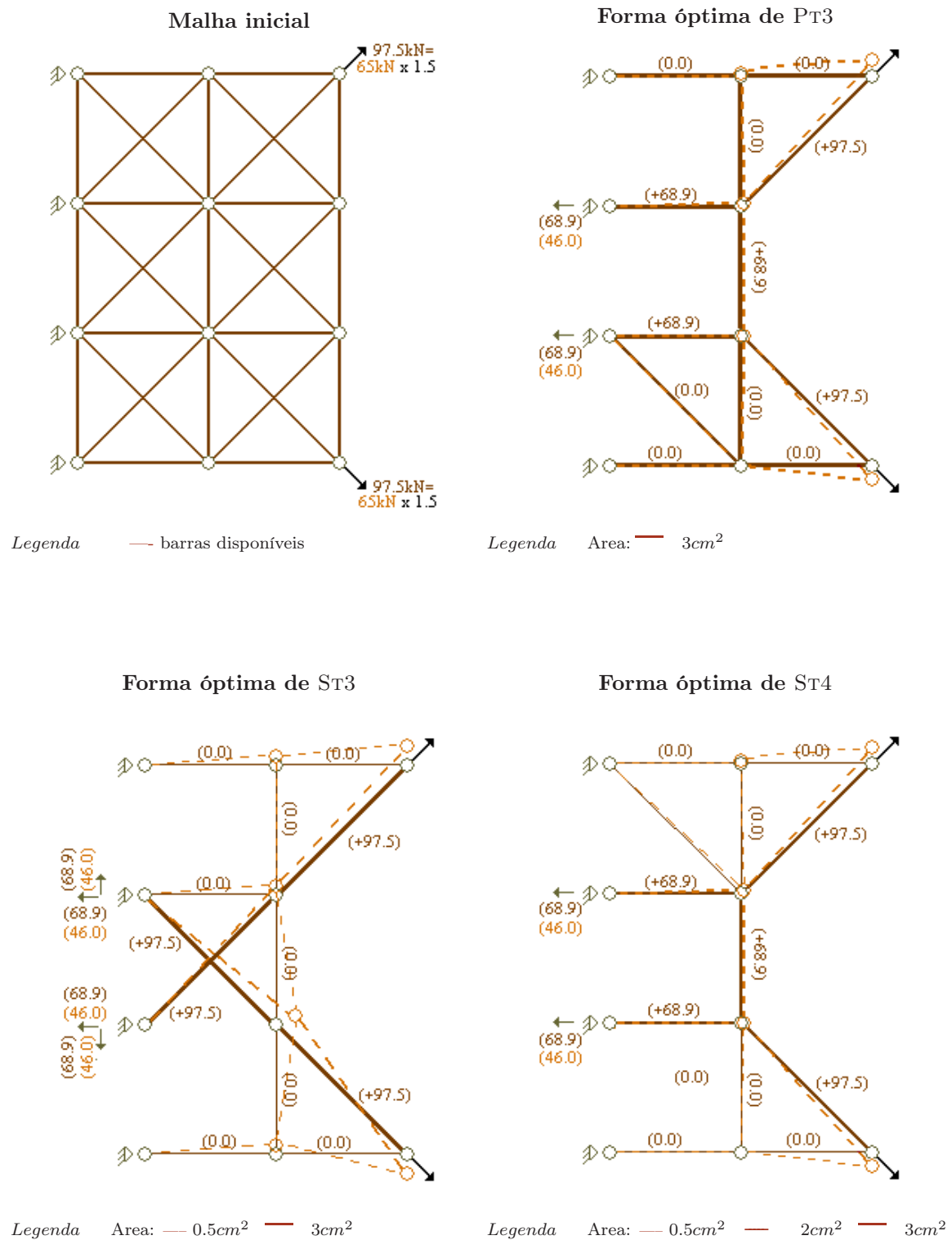


Figura 7.10: Malha inicial e formas estruturais ótimas de PT3, ST3 e ST4

		MPEC		
PROB		<i>nr</i>	<i>nv</i>	<i>nvc</i>
GRUPO I	Pt0	124	111	30
	Pt1	361	310	87
	Pt2	537	394	87
	Pt3	511	451	123
	Pt4	1213	1026	291
GRUPO II	St1	461	470	147
	St2	561	630	207
	St3	656	683	210
	St4	801	915	297

Tabela 7.4: Dimensões dos problemas teste MPEC

	ASETm				
PROB	Ni	T	OBJ	OP.OBJ.	DIFOBJ
PT0	116	0.03	3.60	3.60	0.00
PT1	522	0.24	11.10	10.80	0.03
PT2	961	0.44	14.10	12.90	0.09
PT3	922	0.44	12.66	11.92	0.06
PT4	9859	9.00	27.90	27.30	0.02
ST1	726	0.36	7.10	7.05	0.01
ST2	820	0.56	8.75	4.90	0.79
ST3	1819	1.03	9.40	6.29	0.50
ST4	1731	1.44	9.80	5.46	0.80

Tabela 7.5: Aplicação do algoritmo ASETm em problemas estruturais

para a função objectivo em dm^3 , OP.OBJ. é o valor óptimo conhecido e o parâmetro DIFOBJ é definido por

$$DIFOBJ = \begin{cases} \frac{OBJ - OP.OBJ.}{OP.OBJ.} & \text{se } OP.OBJ. \neq 0 \\ OBJ - OP.OBJ. & \text{se } OP.OBJ. = 0 \end{cases}$$

Os resultados apresentados na Tabela 7.5 mostram que, quando é utilizada apenas uma área possível para a secção transversal de cada barra (problemas do Grupo I), o algoritmo obtém uma solução muito próxima da solução óptima, com um esforço computacional bastante pequeno. No entanto se as áreas possíveis para cada barra são mais do que uma, então aumenta a distância entre o valor óptimo global e o local obtido pelo algoritmo. Notar que, como a solução é admissível e rápida de obter, pode fornecer uma informação bastante útil para a elaboração do projecto, se houver necessidade de obter rapidamente alguma informação

sobre esse assunto.

A Tabela 7.6 inclui os resultados computacionais obtidos pelo algoritmo Sequencial Complementar na resolução dos problemas teste anteriores. Nesta tabela, além dos parâmetros

PROB	SLCP						OBJ.
	Ni	T	Nd	Nis	Ts	Nds	
PT0	227	0.05	49	131	0.02	15	3.60
PT1	5530	2.24	917	750	0.27	40	10.80
PT2	31996	15.48	4054	7212	3.35	732	12.90
PT3	34267	16.49	5202	4811	2.25	476	11.92
PT4(*)	3590182	3471.07	164433	90081	83.35	3654	27.30
ST1	18233	8.15	1826	4134	1.85	406	7.05
ST2	38762	23.75	4492	19356	11.64	1914	4.90
ST3	315894	175.15	24095	135345	73.98	10489	6.29
ST4	914040	614.74	50404	186189	124.28	12357	5.46

Tabela 7.6: Aplicação do algoritmo SLCP em problemas estruturais

usados anteriormente, ND representa o número de nós pesquisados ao longo do processo, enquanto que NIS, Ts e NDS são, respectivamente, o número de operações pivotais, o tempo CPU em segundos e o número de nós pesquisados até à obtenção da solução óptima. Além disso, a notação (*) é usada no problema PT4 para indicar que a resolução do último GLCP foi interrompida pois excedeu o limite máximo de 3500000 iterações (operações pivotais) imposto para a resolução de cada GLCP. Como se pode constatar na Tabela 7.6, a maior parte do esforço computacional realizado pelo método é gasto na resolução do último GLCP, que não tem solução.

Comparando os resultados da Tabela 7.6 com os apresentados na Tabela 7.3 para o programa OSL, pode concluir-se que OSL tem um melhor desempenho do que SLCP para problemas de pequena dimensão, invertendo-se os papéis à medida que a dimensão das malhas aumenta. Desta forma, o algoritmo Sequencial Complementar é um processo alternativo ao programa OSL bastante eficiente e robusto. Além disso, o mínimo global do problema de optimização é obtido de uma forma muito mais rápida do que usando o código de programação inteira. Esses resultados também indicam que a formulação MPEC de um programa inteiro pode ser uma interessante área de investigação a explorar. Com efeito, a aplicação do algoritmo Sequencial Complementar à resolução do MPEC correspondente parece ser uma técnica muito válida para obter um mínimo global de um programa linear inteiro.

CONCLUSÃO

O grande objectivo desta tese foi o desenvolvimento, análise de convergência, implementação e teste de algoritmos para a determinação de mínimos locais e globais de Problemas de Programação Matemática com Restrições de Complementaridade ou de Equilíbrio (MPEC). Nesse sentido desenvolvemos e implementámos um algoritmo local de Restrições Activas Modificado e dois algoritmos globais, nomeadamente um Algoritmo de Ramificação e Limitação e um Método Sequencial Complementar, baseados no algoritmo de restrições activas.

Como o conjunto de restrições do MPEC é um problema linear complementar e a sua resolução é um dos passos mais importantes de qualquer um dos algoritmos mencionados anteriormente, desenvolvemos e implementámos um algoritmo enumerativo também baseado num algoritmo das restrições activas. A eficiência e a robustez deste processo é superior à do algoritmo enumerativo anteriormente desenvolvido em [62], tornando-se ainda mais notória à medida que a dimensão dos problemas aumenta.

O algoritmo local de restrições activas modificado é um processo que tenta obter um valor óptimo para o MPEC utilizando sempre soluções complementares, obtendo assim um ponto estacionário para o MPEC. A sua convergência foi estabelecida tendo por base a hipótese de não degenerescência. Foram sugeridas modificações no algoritmo para também permitir a sua utilização na resolução de problemas degenerados. Embora a convergência não tenha sido estabelecida para este tipo de problemas, na experiência computacional realizada pôde constatar-se que o algoritmo tem um bom desempenho para determinar uma solução para todos os problemas teste degenerados ou não.

Em relação aos algoritmos para a obtenção de um mínimo global do MPEC, o Algoritmo de Ramificação e Limitação mostrou ser um processo eficiente e robusto. Para a determinação de limites inferiores foi sugerido a incorporação no algoritmo de cortes disjuntivos. Esta ideia

revelou-se bastante interessante em alguns dos problemas testados. No entanto, a eficácia dos cortes disjuntivos nem sempre é notória como se pôde observar na resolução dos MPECs associados a Problemas de Ajuste de Funções Dobradiças. Nesta tese foi também proposta uma nova implementação do método sequencial complementar introduzido em [64], que tem por base o método das restrições activas. Esse processo pode ser aplicado na determinação de um mínimo global para MPECs com função objectivo linear e quadrática.

Um dos grandes objectivos desta tese foi a aplicação dos algoritmos referidos à resolução de problemas reais. Nesse sentido foi estudado o Problema Complementar de Valores Próprios, tendo sido desenvolvido para efeito um novo algoritmo de ramificação e limitação baseado na técnica de reformulação-linearização.

Um modelo para a selecção de formas estruturais desenvolvidas em espaços bidimensionais foi ainda introduzido neste trabalho. Um modelo elastoplástico foi adoptado para a optimização estrutural (topológica e de secções), tendo sido consideradas as condições de segurança relacionadas com os estados limites últimos de resistência e os estados limites de utilização. A eliminação de soluções cinematicamente instáveis foi abordada por duas técnicas, nomeadamente a da perturbação (das coordenadas dos pontos nodais) e a da estabilidade cinemática, associada a uma técnica de perturbação de forças nos nós existentes em cada solução.

Os resultados mostraram que a formulação obtida pela segunda técnica fornece uma solução óptima associada a uma estrutura cinematicamente estável para problemas de média dimensão. A aplicação do método sequencial complementar à resolução dos MPECs associados a esses modelos permitiu obter soluções óptimas que fornecem uma substancial melhoria nos gastos de recursos e encoraja a sua utilização em modelos de engenharia estrutural.

Da investigação realizada no âmbito desta tese de doutoramento, surgiram novas ideias que devem merecer a nossa atenção num futuro próximo. Nesse sentido, é de importância fundamental o estudo de novas técnicas para a obtenção de melhores limites inferiores para o algoritmo de ramificação e limitação. Uma nova técnica que combine os algoritmos Sequencial Complementar e de Ramificação e Limitação para a resolução de MPECs com restrições lineares e que tire partido das vantagens dos dois processos deve também ser objecto de investigação. Tal como foi referido no capítulo 6, pensamos ser possível desenvolver um novo processo para a resolução do Problema Complementar de Valores Próprios que explore o MPEC equivalente e que ao mesmo tempo se baseie na técnica de reformulação-linearização.

Finalmente, dada a nossa posição como docente de um Departamento de Engenharia Civil, é de esperar que a nossa atenção incida no futuro no desenvolvimento de outros modelos de optimização estrutural e na análise de algoritmos de optimização eficientes para lidar com esses modelos.

Referências

- [1] Eurocode 1 EN 1991. *Basis of Design and Actions on Structures*. CEN, Brussels, 1998.
- [2] Eurocode 2 EN 1992. *Design of Concrete Structures - Part1: General Rules and Rules for Buildings*. CEN, Brussels, 1999.
- [3] F. Al-Khayyal. An implicit enumeration procedure for the general linear complementarity problem. *Mathematical Programming Studies*, 31:1–20, 1987.
- [4] M. Anistescu. On solving mathematical programs with complementarity constraints as nonlinear programs. Preprint ANL/MCS-P864-1200, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, 2000.
- [5] I. Arora and M. Haung. Methods for optimization of nonlinear problems with discrete variables: a review. *Structural Optimization*, 8:69–85, 1994.
- [6] C. Audet, P. Hansen, B. Jaumard, and G. Savard. Links between the linear bilevel and mixed 0-1 programming problems. *Journal of Optimization Theory and Applications*, 93:273–300, 1997.
- [7] E. Balas. Disjunctive programming. *Annals of Discrete Mathematics*, 5:3–51, 1979.
- [8] J. Bard. Convex two level optimization. *Mathematical Programming*, 40:15–27, 1988.
- [9] J. Bauer. A survey of methods for discrete optimum structural design. *Comput. Assis. Mech. Eng. Sci.*, 1:27–38, 1994.
- [10] M. Bazaraa, H. Sherali, and C. Shetty. *Nonlinear Programming: Theory and Algorithms*. John Wiley & Sons, New York, 1993.

- [11] D. Bertsekas and J. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, Belmont, Massachusetts, 1996.
- [12] W. Bialas and M. Karwan. Two-level linear programming. *Management Science*, 30:1004–1020, 1984.
- [13] S. Billups and K. Murty. Complementarity problems. *Journal of Computational and Applied Mathematics*, 124:303–318, 2000.
- [14] S. Bollapragada, O. Ghattas, and J. Hoocher. Optimal design of truss structures by logic-based branch and cut. *Operations Research*, 49:42–51, 2001.
- [15] L. Breiman. Hinging hyperplanes for regression, classification and function approximation. *IEEE Transactions on Information Theory*, 39:999–1013, 1993.
- [16] A. Brooke, D. Kendrick, A. Meeraus, and R. Raman. GAMS a User's Guide. GAMS Development Corporation, New York, 1998.
- [17] R. Byrd, M. Hribar, and J. Nocedal. An interior point method for large scale nonlinear programming. *SIAM J. Optimization*, 9:877–900, 1999.
- [18] P. Calamai and L. Vicente. Generating linear and linear-quadratic bilevel programming problems. *SIAM J.SCI.Comput*, 14:770–782, 1993.
- [19] P. Calamai and L. Vicente. Generating quadratic bilevel programming test problems. *ACM Transactions on Mathematical Software*, 20:103–119, 1994.
- [20] M. Campelo and S. Scheimberg. A simplex approach for finding local solutions of a linear bilevel program by equilibrium points. a aparecer em *Annals of Operations Research*.
- [21] Y. Chen and M. Florian. The nonlinear bilevel programming problem: Formulations, regularity and optimality conditions. *Optimization*, 32:193–209, 1995.
- [22] S. Chung. NP-completeness of the linear complementarity problems. *Journal of Optimization Theory and Applications*, 60:393–399, 1989.
- [23] A. Cichocki and R. Unbehauen. *Neural Networks for Optimization and Signal Processing*. John Wiley & Sons, New York, 1996.

- [24] A. Costa, J. Martins, I. Figueiredo, and J. Júdice. The directional instability problem in systems with frictional contacts. *Comput. Methods Appl. Mech. Engrg*, 193:357–384, 2004.
- [25] R. Cottle. The principal pivoting method of quadratic programming. In G. Dantzig and A. Veinott Jr., editors, *Mathematics of Decision Sciences*, pages 144–162. American Mathematical Society, Rhode Island, 1968.
- [26] R. Cottle, J. Pang, and R. Stone. *The Linear Complementarity Problem*. Academic Press, New York, 1992.
- [27] G. Dantzig. *Linear Programming and Extensions*. Princeton University Press, Princeton, New Jersey, 1963.
- [28] G. Dantzig and R. Cottle. Positive (semi-)definite programming. In J. Abadie, editor, *Nonlinear Programming*, pages 55–73. Amsterdam, 1967.
- [29] W. Davidon. Variable metric methods for minimization. Technical Report ANL 5990, Argonne National Laboratory, Argonne, 1959.
- [30] K. Deb and S. Gulati. Design of truss-structures for minimum weight using genetic algorithms. *Finite Elements in Analysis and Design*, 37:447–465, 2001.
- [31] S. Dirkse, M. Ferris, and A. Meeraus. Mathematical programs with equilibrium constraints: Automatic reformulation and solution via constraint optimization. Technical Report NA-02/11, Oxford University Computing Laboratory, Oxford, 2002.
- [32] W. Dobbs and L. Felton. Optimization of truss geometry. *J. Struct. Div. ASCE*, 95:2105–2118, 1969.
- [33] W. Dorn, R. Gomory, and H. Greenberg. Automatic design of optimal structures. *J. Mec.*, 3:25–52, 1964.
- [34] F. Facchinei, H. Jiang, and L. Qi. A smoothing method for mathematical programs with equilibrium constraints. *Mathematical Programming*, 85:107–134, 1999.
- [35] F. Facchinei and J. Pang. *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer-Verlag, New York, 2003.

- [36] L. Fernandes, J. Júdice, and J. Patrício. An investigation of interior-point and block pivoting algorithms for large-scale monotone linear complementarity problems. *Computational Optimization and Applications*, 5:49–77, 1994.
- [37] M. Ferris and C. Kanzow. Complementarity and related problems: A survey. Technical Report 98-17, Mathematics Research Center, University of Wisconsin, 1999.
- [38] A. Fischer. A special Newton-type optimization method. *Optimization*, 24:269–284, 1992.
- [39] A. Fischer. An NCP-functional and its use for the solution of complementarity problems. In D. Du, L. Qi, and R. Womersley, editors, *Recent Advances in Nonsmooth Optimization*, pages 88–105. World Scientific Publishers, Singapore, 1995.
- [40] A. Fischer. A Newton-type method for positive semidefinite linear complementarity problems. *Journal of Optimization Theory and Applications*, 86:585–608, 1995.
- [41] R. Fletcher and S. Leyffer. Nonlinear programming without a penalty function. *Mathematical Programming*, 91:239–269, 2002.
- [42] R. Fletcher and S. Leyffer. Numerical experience with solving MPECs as NLPs. Technical Report NA/210, Department of Mathematics, University of Dundee, 2002.
- [43] R. Fletcher, S. Leyffer, and P. Toint. On the global convergence of an SLP-filter algorithm. Technical Report NA/183, Department of Mathematics, University of Dundee, 1998.
- [44] C. Floudas. *Deterministic Global Optimization: Theory, Methods and Applications*. Kluwer Academic Publishers, Boston, 2000.
- [45] J. Fortuny-Amat and B. McCarl. A representation and economic interpretation of a two-level programming problem. *Journal of the Operational Research Society*, 32:783–792, 1981.
- [46] M. Fukushima, Z. Luo, and J. Pang. A globally convergent sequential quadratic programming algorithm for mathematical programs with linear complementarity constraints. *Computational Optimization and Applications*, 10:5–34, 1998.

- [47] O. Ghattas and I. Grossmann. MINLP and MILP strategies for discrete sizing structural optimization problems. In *Proceedings of ASCE 10th Conference on Electronic Computation*. Indianapolis, 1991.
- [48] A. Ghosh and A. Mllik. *Theory of Mechanisms and Machines*. Affiliated East-West Press, New Delhi, 1988.
- [49] P. Gil, W. Murray, and M. Saunders. SNOPT: An SQP algorithm for large-scale constrained optimization. *SIAM J. Optimization*, 12:979–1006, 2002.
- [50] P. Gil, W. Murray, and M. Wright. *Pratical Optimization*. Academic Press, London, 1981.
- [51] R. Glowinski. Finite elements and variational inequalities. Technical Report 1885, Mathematics Research Center, University of Wisconsin, 1978.
- [52] G. Golub and C. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, 1996.
- [53] R. Graves. A principal pivoting simplex algorithm for linear and quadratic programming. *Operations Research*, 15:482–494, 1967.
- [54] X. Guo and G. Cheng. An extrapolation approach for the solution of singular optima. *Struct. Multidis. Optim.*, 19:255–262, 2000.
- [55] X. Guo, G. Cheng, and K. Yamazaki. A new approach for the solution of singular optima in truss topology optimization with tress and local buckling constraints. *Struct. Multidis. Optim.*, 22:364–372, 2001.
- [56] W. Gutkowski. A discrete programming in structural optimization. In M. Kleiber, editor, *Handbook of Computational Solid Mechanics*, chapter V.4. Springer, Berlin, 1998.
- [57] A. Hall and A. Kabaila. *Basic Concepts of Structural Analysis*. John Wiley & Sons, New York, 1977.
- [58] P. Harker and J. Pang. Finite-dimensional variational inequalities and complementarity problems: a survey of theory, algorithms and applications. *Mathematical Programming*, 60:161–220, 1990.

- [59] M. Hassoun. *Fundamentals of Artificial Neural Networks*. MIT Press, Massachusetts, 1995.
- [60] A. Hoback. Optimization of singular problems. *Structural Optimization*, 12:93–97, 1996.
- [61] R. Horst, P. Pardalos, and N. V. Thoai. *Introduction to Global Optimization*. Kluwer Academic Publishers, Boston, 1995.
- [62] J. Júdice and A. Faustino. An experimental investigation of enumerative methods for the linear complementarity problem. *Computers and Operations Research*, 15:417–426, 1988.
- [63] J. Júdice and A. Faustino. The solution of the linear bilevel programming problem by using the complementarity problem. *Investigação Operacional*, 8:77–95, 1988.
- [64] J. Júdice and A. Faustino. A computational analysis of LCP methods for bilinear and concave quadratic programming. *Computers and Operations Research*, 18:645–654, 1991.
- [65] J. Júdice and A. Faustino. A sequential LCP algorithm for bilevel linear programming. *Annals of Operations Research*, 34:89–106, 1992.
- [66] J. Júdice, A. Faustino, and I. Ribeiro. On the solution of NP-hard linear complementarity problems. *TOP-Sociedad de Estadística e Investigación Operativa*, 10, No.1:125–145, 2002.
- [67] J. Júdice and G. Mitra. Reformulation of mathematical programming problems as linear complementarity problems and investigation of their solution methods. *Journal of Optimization Theory and Applications*, 57:123–149, 1988.
- [68] J. Júdice and M. Pires. A block principal pivoting algorithm for large-scale strictly monotone linear complementarity problems. *Computers and Operations Research*, 21:587–596, 1994.
- [69] J. Júdice and L. Vicente. On the solution and complexity of a generalized linear complementarity problem. *Journal of Global Optimization*, 4:415–424, 1994.

- [70] H. Jiang and D. Ralph. Smooth SQP methods for mathematical programs with nonlinear complementarity constraints. University of Melbourne, Department of Mathematics, 1997.
- [71] H. Jiang and D. Ralph. QPECgen, a Matlab generator for mathematical programs with quadratic objectives and affine variational inequality constraints. *Computational Optimization and Applications*, 13:25–59, 1999.
- [72] N. Kamarkar. A new polynomial-time algorithm for linear programming. *Combinatória*, 4:373–395, 1984.
- [73] E. Keller. The general quadratic optimization problem. *Mathematical Programming*, 5:311–337, 1973.
- [74] K. Kirchgässner. Ein verfahren zur maximierung linear funktionen in nichtkonvexen bereichen. *Zeitschrift fur Angewandte Mathematik*, 42:22–24, 1962.
- [75] U. Kirsch. On singular topologies in optimum structural design. *Structural Optimization*, 2:133–142, 1990.
- [76] M. Kojima, N. Megiddo, T. Noma, and A. Yoshise. A unified approach to interior-point algorithms for linear complementarity problems. In *Lecture Notes in Computer Science: 538*. Springer-Verlag, Berlin, 1991.
- [77] C. Kolstad. A review of the literature on bilevel mathematical programming. Technical Report LA-10284-MS, Los Alamos National Laboratory, Los Alamos, 1985.
- [78] C. Kolstad and L. Lasdon. Derivative evaluation and computational experience with large bilevel mathematical programs. Technical Report 1266, College of Commerce and Business Administration, University of Illinois, Urbana-Champaign, 1986.
- [79] H. Konno. Bilinear programming: Part ii - Applications of bilinear programming. Technical report, Department of Operations Research, Stanford University, 1971.
- [80] H. Konno. A cutting-plane algorithm for solving bilinear programs. *Mathematical Programming*, 11:14–27, 1976.
- [81] M. Kostreva. Block pivoting methods for solving the complementarity problem. *Linear Algebra and its Applications*, 21:207–215, 1979.

- [82] H. Kuhn and W. Tucker. Nonlinear programming. In J. Neyman, editor, *Proceedings of the Second Berkeley on Mathematical Statistics and Probability*, pages 481–492. University of California Press, Berkeley, 1951.
- [83] C. Lemke. On complementary pivot theory. In G. Dantzig and A. Veinott Jr., editors, *Mathematics of Decision Sciences*, pages 95–114. American Mathematical Society, Providence, 1968.
- [84] S. Leyffer. The penalty interior point method fails to converge for mathematical programs with equilibrium constraints. Technical Report NA/208, Department of Mathematics, University of Dundee, 2002.
- [85] Z. Luo, J. Pang, and D. Ralph. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, New York, 1997.
- [86] O. Mangasarian. The linear complementarity problem as a separable bilinear program. *Journal of Global Optimization*, 6:153–161, 1995.
- [87] H. Mills. Equilibrium points in finite games. *J. Soc. Indust. Appl. Math*, 8:397–402, 1960.
- [88] B. Murtagh and A. Saunders. A projected lagrangian algorithm and its implementation for sparse nonlinear constraints. *Mathematical Programming, Study 16*, pages 84–117, 1982.
- [89] B. Murtagh and A. Saunders. MINOS 5.0 User’s Guide. Technical Report SOL 83-20, Department of Operations Research, Stanford University, 1983.
- [90] K. Murty. Note on a Bard-type scheme for solving the complementarity problem. *Opsearch*, 11:123–130, 1974.
- [91] K. Murty. *Linear Programming*. Wiley, New York, 1983.
- [92] K. Murty. *Linear Complementarity, Linear and Nonlinear Programming*. Heldermann Verlag, Berlin, 1988.
- [93] K. Murty and J. Júdice. On the complexity of finding stationary points of nonconvex quadratic programming. *Opsearch*, 33:162–166, 1996.

- [94] A. Neves, A. Faustino, F. Pires, and J. Júdice. Elastoplastic analysis of structures and linear complementarity. In J. Coelho and L. Tavares, editors, *O. R. Models on Microcomputers*, pages 217–228. North-Holland, Amsterdam, 1986.
- [95] J. Nocedal and S. Wright. *Numerical Optimization*. Springer-Verlag, New York, 1999.
- [96] H. Onal. A modified simplex approach for solving bilevel linear programs. *European Journal of Operational Research*, 67:126–135, 1993.
- [97] J. Outrata, M. Kocvara, and J. Zowe. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications and Numerical Results*. Kluwer Academic Publishers, Boston, 1998.
- [98] F. Pires and A. Neves. A complementaridade linear na resolução de estruturas com apoios unilaterais. *Investigação Operacional*, 9:89–96, 1989.
- [99] P. Pucar and J. Sjöberg. On the hinge finding algorithm for hinging hyperplanes. *IEEE Transactions on Information Theory*, 44:1310–1318, 1998.
- [100] M. Pyrz and J. Zawidzka. Optimal discrete truss design using improved sequential and genetic algorithm. *Engineering Computations*, 18:1078–1090, 2001.
- [101] M. Queiroz, J. Júdice, and C. Humes Jr. The symmetric eigenvalue complementarity problem. a aparecer em *Mathematics of Computations*.
- [102] M. Queiroz, C. Humes Jr, and J. Júdice. On finding global optima for the hinge fitting problem. *Computers and Operations Research*, 31:101–122, 2004.
- [103] S. Robinson. A quadratically convergent algorithm for general nonlinear programming problems. *Mathematical Programming*, 3:145–156, 1972.
- [104] G. Rozvany. Difficulties in truss topology optimization with stress, local buckling and system stability constraints. *Structural Optimization*, 11:213–217, 1996.
- [105] G. Rozvany, M. Bendsoe, and Kirsch. Layout optimization of structures. *Appl. Mec. Rev.*, 48:41–118, 1995.
- [106] E. Salajegheh and G. Vanderplaats. Optimum design of trusses with discrete sizing and shape variables. *Structural Optimization*, 6:79–85, 1993.

- [107] H. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: Stationarity, optimality and sensitivity. *Mathematics of Operations Research*, 25:1–22, 2000.
- [108] S. Scholtes. Active set methods for inverse linear complementarity problems. Technical Report 28/1999, Department of Engineering, University of Cambridge, 1999.
- [109] H. Serali and W. Adams. *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*. Kluwer Academic Publishers, Boston, 1999.
- [110] H. Serali and C. Shetty. Optimization with disjunctive constraints. In M. Beckman and H. Künzi, editors, *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag, New York, 1980.
- [111] H. Simon. *The Sciences of the Artificial*. MIT Press, Massachusetts, 1969.
- [112] G. Sved and Z. Ginos. Structural optimization under mutiple loading. *Int. J. Mech. Sci.*, 10:803–805, 1968.
- [113] A. Templeman. Heuristic methods in discrete structural optimization. In W. Gutkowski, editor, *Discrete Structural Optimization*, pages 135–165. CISM Courses and Lectures No. 373, Udine, Italy, 1997.
- [114] P. Thander and G. Vanderplaats. Survey of discrete variable optimization for structural design. *Journal of the Structural Engineering*, 121:301–306, 1995.
- [115] A. Toakley. Optimum design using available sections. *Journal of the Structural Division, American Society of Civil Engineers*, 34:1219–1241, 1968.
- [116] B. Topping. Topology design of discrete structures. In M. Bendsoe and C. Soares, editors, *Topology design of structures*, pages 517–534. Kluwer Academic Publishers, Amsterdam, 1993.
- [117] R. Vanderbei and D. Shanno. An interior point algorithm for nonlinear programming. *Computational Optimization and Applications*, 13:231–252, 1999.

-
- [118] G. Vanderplaats and P. Thanedar. A survey of discrete variable optimization for structural design. In *Proceedings of the ASCE Tenth Conference on Electronic Computations*, pages 173–180. Dayton, Ohio, 1991.
- [119] P. Wolfe. The reduced-gradient method. Unpublished Manuscript, 1962.
- [120] S. Wright. An infeasible-interior-point algorithm for linear complementarity problems. *Mathematical Programming*, 67:29–51, 1994.
- [121] S. Wright. *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia, 1997.
- [122] Y. Ye. A fully polynomial-time approximation algorithm for computing a stationary point of the general linear complementarity problem. *Mathematics of Operations Research*, 18:334–345, 1993.
- [123] O. Zienkiewicz. *The Finite Element Method*. MacGraw-Hill, Berkshire, 1997.